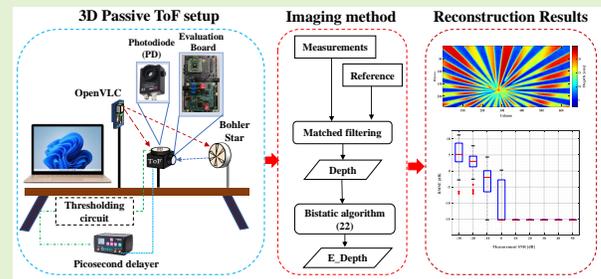


Pseudo-Passive Time-of-Flight Imaging: Simultaneous Illumination, Communication, and 3D Sensing

Faisal Ahmed, Miguel Heredia Conde, Paula López Martínez, Thomas Kerstein, and Bernd Buxbaum

Abstract—3D Time-of-Flight (ToF) cameras have recently received a lot of attention due to their wide range of applications. Despite remarkable advancements in ToF imaging, state-of-the-art ToF cameras are still afflicted by the power hungriness of their illumination sources. To tackle this problem, we exploited existing lighting infrastructure, that ensures the ubiquitous presence of modulated light sources in indoor spaces, which serve as opportunity illuminators. We explored the bistatic geometry for passive imaging using the pulse-based ToF approach. Our work is inspired by the recently introduced visible light communication (VLC) or light-fidelity (Li-Fi) infrastructure. VLC allows the infrastructure to provide indoor simultaneous illumination, communication, and sensing (SICS). To this end, we designed a bistatic geometry for the purpose of attaining passive 3D imaging. Such capabilities are achieved by exploiting the pulse shape of the autocorrelation function of real optical signals generated by VLC/Li-Fi modules (e.g., OpenVLC and LiFiMAX). We demonstrated passive imaging by means of matched filtering. In this work, we also studied different sampling strategies in the time shift domain, including uniform, random, and sparse rulers, which is another step forward towards preserving high depth accuracy with a minimal number of measurements. The proposed methodology achieved successful depth reconstruction with negligible root-mean-square-error (RMSE) for the low signal-to-noise ratio (SNR) of the measurements. Parametric models such as Gaussian and sum-of-sines are used to characterize the cross-correlation functions and allow for robust parametric depth retrieval from a few measurements. Moreover, we attained 20 mm worst-case error for a target at 25 cm. The experiment proved that the bistatic passive depth reconstruction is feasible.

Index Terms—Passive sensing, Time-of-Flight, VLC, OpenVLC, bistatic, sampling, LiFi, depth imaging.



I. INTRODUCTION

OVER the past few years, 3D Time-of-Flight (ToF) imaging technology has evolved significantly and has attained a prominent interest from industry and academic circles due to its wide range of applications, such as robot navigation, indoor sensing, autonomous driving, surface mapping, surveillance, gaming, and human-machine interaction, previously reserved to LiDARs and stereo vision systems [1]–[3]. ToF is an active imaging technique that is able to produce both *intensity* and *depth* maps. A ToF camera computes the distance between the camera and the scene objects in a per-pixel basis by exploiting the time lapse that the photons confront, when the modulated light signals are projected onto the scene and bounce back to

the camera, which is proportional to the distance between the camera and each corresponding scattered point.

In recent years, rapid advances in solid-state technology have profoundly transformed the lighting infrastructure from conventional lamps (e.g., incandescent and halogen) to light-emitting diodes (LEDs). Thus, LEDs have become popular for displays and light sources because of their advantages such as long lifetime, small size, low-cost, energy efficiency, and low switching transient [4]. The switching capability of LEDs enables the visible light channel to be modulated at high frequencies, which are high enough to be imperceptible to the human eye [5]. LED lighting has gained global recognition as a green lighting technology in recent years [6], [7]. LEDs are predicted to replace conventional lights eventually and become the ultimate light source for many applications [8]. This transition has transmogrified the lighting infrastructure into a novel communication paradigm. This paradigm shift has accelerated the development of Visible Light Communication (VLC). VLC is a promising mechanism and an accessible technology, driven by the ubiquitous proliferation of low-cost LED sources in indoor environments. LED-based VLC systems offer numerous advantages, including low-cost front-

This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No: 860370 (MENELAOS^{NT}).

Faisal Ahmed and Miguel Heredia Conde are with the Center for Sensorsystems, University of Siegen, Germany (e-mail: Faisal.Ahmed@uni-siegen.de and heredia@zess.uni-siegen.de).

Paula López Martínez is with CITIUS, University of Santiago de Compostela, Spain.

Thomas Kerstein and Bernd Buxbaum are with pmdtechnologies ag, Siegen, Germany.

ends, immunity to electromagnetic interference, high physical-layer security, and unregulated spectrum resources (400 THz - 900 THz) [9]. It is believed that VLC will play a significant role in the fifth-generation (5G) and sixth-generation (6G) communication [10], [11]. This brings communication and lighting together within a single module. Therefore, the VLC infrastructure can provide multiple services simultaneously, such as communication, illumination, and now, also ToF 3D imaging in indoor environments. In this context, the VLC infrastructure has laid down a solid foundation for pseudo-passive ToF sensing.

Additionally, VLC-enabled front-ends from companies such as pureLiFi, Philips, and Oledcomm are accessible to the general public and have rapidly penetrated into the commercial market, including homes, offices, and industrial buildings¹. The commercialized VLC technology (IEEE 802.15.7) offers a data rate of 150 Mbps and the research prototypes reach 1 Gbps of data rates [12]. Light-based wireless communication, the so-called *VLC*, and *ToF sensing* systems, have been developed independently for two decades after the emergence of optical wireless communication (OWC). The passive imaging method we present transcends the communication-only or sensing-only philosophy and, alternatively, uses a VLC-enabled sensing approach based on the cooperation between VLC and ToF imaging. Thus, this study broadens the reach of VLC systems to synergistically support communication and sensing services.

Active 3D imaging has been a vibrant research topic for many years [13]–[15]. Despite significant progress in ToF imaging, state-of-the-art ToF cameras are still susceptible to the high power consumption of their dedicated illumination units. Unfortunately, this problem remains unaddressed, and most of the recent progress is on the signal processing side. For example, in [16], the illumination system features 91 W of optical power for wide-area ToF imaging. This precludes the applicability of ToF imaging in specific scenarios. To overcome this problem, an alternative approach that uses an opportunity illuminator for sensing is proposed. This makes the built-in dedicated illumination unit futile, thus enabling *passive* ToF 3D sensing. VLC has recently been used for passive ToF sensing without synchronization between the source and the ToF camera. This leads to an unknown depth offset in passive ToF imaging [17]. Nonetheless, this passive modality still needs refining to attain accurate depth recovery.

Thanks to the development of a bistatic sensing geometry and the refinement of the ToF sensing pipeline for passive imaging modality, we solved the synchronization problem by introducing a direct link between the emitter and a reference photodiode. In general, the bistatic configuration uses two parallel channels. One is the reference channel, which transmits signals between the VLC source and the reference photodiode to acquire an external reference signal for the *Photonic Mixer Device* (PMD) camera. Another one captures scene-related reflections. Informative measurements are achieved by cross-correlating the reference and reflected signals. This solution exploits the existing VLC infrastructure [18] to illuminate the scene with modulated light and synchronize the camera

with an externally-provided signal, allowing accurate depth reconstruction.

To evaluate this alternative, we have used an OpenVLC1.3 module with a white LED, and a LiFiMAX module with an infrared (IR) LED for our simulations [19], [20]. OpenVLC is a low-cost and open-source platform with a bandwidth of 1 MHz and supports a throughput of 400 kbps. The LiFiMAX module can provide a data rate of 40 Mbps and 100 Mbps in uplink (UL) and downlink (DL), respectively.

Moreover, in this work, the depth is reconstructed using different sampling methods in the time shift domain. In an attempt to reduce the number of measurements, we compared different sampling approaches, both uniform and non-uniform. The signal processing community extensively relies on uniform sampling in the communication and sensing domains, while non-uniform alternatives often remain unexplored, arguably except from random sampling. Random sampling is used as a basis for compressive sensing to recover sparse signals from a few measurements [21]. In addition, for depth reconstruction, we used matched filtering method.

Matched Filtering: A matched filter (MF) is a well-known signal processing technique for improving signal quality and estimating delays. The temporal shift is observed by correlating a known delayed signal (or template) to an unknown signal [22]–[24]. Matched filtering is often performed in an analog circuit that carries out a correlation operation and then finds the peaks, or in a digital circuit or computer that takes samples of the signal and calculates the discrete correlation function. The matched filter offers accurate results at a low computational cost, thus allowing for high frame rates. In this context, the sampling rate should be high enough.

6G technology promotes communication and sensing simultaneously. In this context, VLC infrastructure can be re-used for multiple services. To the best of the authors' knowledge, this is the first work exploring VLC infrastructure as a drop-in replacement for the illumination unit in ToF depth cameras.

The rest of the paper is structured as follows. Section II provides a summary of related work in the area of pulse-based (PB) ToF and existing passive sensing methods. Section III is devoted to the system model. In Section IV, the depth recovery method and different sampling schemes are briefly discussed. The experimental setup is demonstrated and elaborated in Section V. Section VI provides the simulation performance of the system based on several sampling schemes and the first-ever passive-ToF 3D reconstruction. Finally, Section VII draws conclusions and proposes future lines of work.

II. RELATED WORK

OWC and ToF sensing have achieved unprecedented results independently in the recent past, but without mutual intersection. To date, no prior works have studied the intersection of both technologies for simultaneous communication and 3D sensing. Recently, the emergence of OWC variants, such as VLC, LiFi, and free space optical communication (FSOC), has brought interesting avenues for passive ToF imaging. These variants are frequently used for communications and illumination in indoor and outdoor settings.

¹<http://purelifi.com/case-studies/>

In parallel, ToF cameras have made great strides in depth reconstruction. In general, ToF cameras are segregated into two operational modes, i.e., PB-ToF mode, and *continuous wave* (CW) ToF mode. In PB-ToF mode, the source emits a short pulse that illuminates the scene, and bounce-back signals are received by the ToF camera. ToF pixels perform an integration of the scene reflected light mixed with a *demodulation control signal* (DCS). The measurements are achieved by shifting the DCS with respect to the illumination control signal. In CW mode, measurements are obtained for different phase shifts between the DCS and the illumination signal.

The phased ToF camera was primarily pioneered by Prof. R. Schwarte [25]. His work paved the way for the success of ToF cameras in the computer vision community. Over the years, ToF imaging technology has become ubiquitous in a wide range of 3D imaging applications. The PMD camera is a leading-edge technology for CW-ToF cameras. These devices are able to extract depth from raw data following the phase stepping algorithm [26]. Such devices are frequently endorsed due to their mature processing pipeline and publicly-accessible designs [27]. A related work outlined the fundamental operation of lock-in ToF cameras, their merits and shortcomings, the layout of ToF pixels, and a remedy to practical difficulties that appear when a PMD camera is being used in the presence of background light [26], [28]–[30].

LiFi was initially demonstrated by the German physicist Harald Haas. VLC is a subset of temporally structured lighting [31]. The challenging task is to encode the information in a lighting framework (one or multiple VLC sources). In [32], the authors reported recent advancements in VLC hardware technology. They demonstrated that the blue LEDs and color converters attained 1485 MHz and 470 MHz bandwidth, respectively. These advances not only boost the capacity of VLC channels, but also facilitate passive ToF sensing. In 2019, VLC was used for ranging and vehicular communication. This work is not directly related, but the ToF technique is analogous [33]. The ToF sensing technique is a well-known problem: determining the distance from the reflection of a known signal.

Recent research efforts have been devoted to the development of PB-ToF imaging systems. Typically, rectangular pulse shapes are used in PB-ToF sensors. It is challenging to generate perfectly square pulses, since vertical rising or falling edges would require unlimited bandwidth.

In 2018, Sarbolandi [34] used a PB-ToF camera (Hamamatsu area sensor S11963-01CR) featuring two electronic shutters, known as *gates*, which are used for the accumulation of photogenerated carriers from bounced-back photons. Both gates are triggered sequentially for a few nanoseconds. The distance to an object leads to a shift in the reflected signal compared to the emitted signal. Thereby, the time-shift determines depth. Lang [35] developed a PB-ToF technique for classifying materials based on their unique signatures. The authors of [36] developed an interferometric technique for depth imaging that uses the inherent photon bunching signature of thermal light, which was initially demonstrated by Hanbury Brown and Twiss [37], [38]. This work adds complexity to the system, since the illumination signals should be conditioned before use. This method requires high-bandwidth detectors.

In contrast to our work, other authors employed photometric stereo (PS) and aperture masking interferometry to perform passive imaging [39]–[42]. Furthermore, the need for multiple sources and an appropriate footprint are fundamental problems in such techniques for scene reconstruction using low-coherence interferometry. The aforementioned passive ToF alternatives [36], [39] have significant drawbacks and are far from being practical.

Differently, a VLC-enabled passive ToF system can use a single broadcast signal for multiple purposes, such as communication, illumination, and 3D sensing. This exploits the same spectrum and hardware resources, allowing it to play a significant role in the future wireless network industry.

Unlike previous studies, we make use of the fortunate fact that VLC infrastructure is often found in homes, office settings, industrial zones, and vehicles, i.e., interesting application fields for ToF cameras [43]. We propose using the existing VLC infrastructure for illumination, turning the background light from a disturbance into a useful optical signal for the ToF camera. In light of existing literature, no prior works have shown passive ToF imaging by leveraging opportunity illuminators. This paper demonstrates the VLC-enabled passive ToF system that can provide communication and 3D sensing *for free* in terms of additional power consumption.

III. SYSTEM MODEL

We consider a single-input-single-output (SISO) system that can emit signals aimed at both probing scene objects and enabling communication with DL users. Fig. 1 demonstrates the detailed block diagram of our proposed VLC-enabled passive ToF sensing scheme. In this section, we first model the free-space optical channel, and later on, we define the mathematical model based on the communication and the passive ToF sensing perspectives. Consider a point-to-point VLC-based optical wireless link that broadcasts a DL signal toward the scene and the user. A bistatic setup is designed to synchronize the ToF camera. In this context, the reference signal for the PMD module can be obtained via a direct line-of-sight (LoS) link between the VLC source and the reference photodiode. The VLC source and the ToF camera should not be co-located in contrast to conventional ToF cameras.

A. VLC Channel

The underlying characteristics of a VLC channel are mainly dictated by the optical link configuration [44]. The most influential aspect of VLC communication performance is the quality of the communication channel. Kahn *et al.* [45] demonstrated six different configuration cases for indoor VLC links based on the presence or absence of an LoS link between an optical transmitter (TX) and a receiver (RX). The degree of directionality and orientation between TX and RX determines the link configuration of indoor VLC systems. The direct LoS link is a dominant communication link configuration and is widely used in the literature [46]. The emitter-beam and the Field-of-View (FoV) of the receiver define the optical transmission channel. In this case, the LoS channel facilitates obtaining a higher received light intensity (i.e., a higher SNR

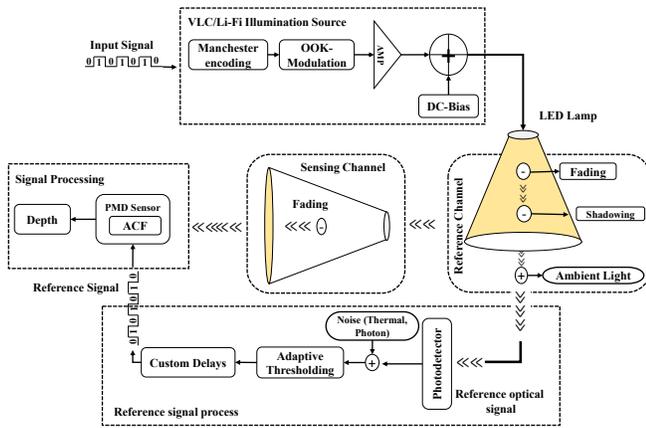


Fig. 1. Block diagram of proposed simultaneous illumination, communication, and 3D ToF sensing.

that can be used to achieve suitable data rates and a long communication range). Propagation of optical signals in an optical wireless channel [47] can be written as:

$$P_R = P_T \cdot G_{OTX} \cdot G_D \cdot G_{ORX} \cdot A_{SL} \quad (1)$$

where P_R denotes the received optical power, P_T is the source transmitted signal power, G_{OTX} is the gain of an optical transmitter, G_{ORX} is the gain of the optical receiver, and G_D denotes the attenuation of light intensity over the propagation distance. A_{SL} accounts for the other system-dependent losses due to the system design and the link misalignment configurations.

Optical transmitters (i.e., LEDs) are the core components of VLC communication. A Lambertian pattern models the intensity profile of LEDs in the spatial domain. In fact, the radiation pattern from the Lambertian source has radial symmetrical profiles and is controlled by its order of emission $m = -\ln(2)/\ln(\cos(\Phi_{1/2}))$. Here $\Phi_{1/2}$ represents the half-power beam-width of the transmitter. The existing literature on optical wireless channels has reported that the Lambertian radiation pattern is widely accepted by the VLC research community (see e.g., [48], [44], and references therein). Considering the Lambertian beam transmitter aperture and received light at the photodiode (PD), then the gain of the optical transceiver in the presence of free-space propagation losses can be expressed [49] as follows:

$$G_{OTX} = \frac{32}{\Phi^2}; G_{ORX} = \left(\frac{\pi D_R}{\lambda}\right)^2; G_D = \left(\frac{\lambda}{4\pi d_{TRX}}\right)^2 \quad (2)$$

where Φ is the full-width transmit beam divergence angle, D_R denotes the optical receiver aperture diameter, d_{TRX} is the distance between TX and RX systems, and λ is the wavelength. Equation (1) can be rewritten as (3) by substituting the gain values.

$$P_R = P_T \left(\frac{2D_R^2}{\Phi^2 d_{TRX}^2}\right) \cdot A_{SL} \quad (3)$$

From (3), we can encapsulate the free-space propagation loss η and this can be given by,

$$\eta = \left(\frac{2D_R^2}{\Phi^2 d_{TRX}^2}\right) \cdot A_{SL} \quad (4)$$

According to (4), the attenuation factor is also affected by the emitter's beam width, the receiver's aperture diameter, and the propagation distance between TX and RX. A wider beam width leads to high attenuation and a large aperture reduces the attenuation coefficient because the receiver may collect more light. In practice, LED generates the far-field pattern with its maximum intensity region. For an LED with a Lambertian emission pattern, the angular intensity distribution is the maximum at 0° , and the theoretical half-power beam width is $\Phi_{1/2} = 120^\circ$. Optical signal propagation loss (optical path loss) can be computed by making use of the LED beam divergence angle and the diameter of the PD aperture.

B. VLC Communication Model

VLC is a novel wireless communication technology supported by existing lighting infrastructure. Fig. 2 presents an indoor VLC link. The VLC infrastructure now serves as an opportunity illuminator in passive ToF sensing. We used two types of light-based wireless communication modules: OpenVLC1.3, which is a research platform, and LiFiMAX module, which is a marketable product. The functionality of the modules is discussed in the sequel.

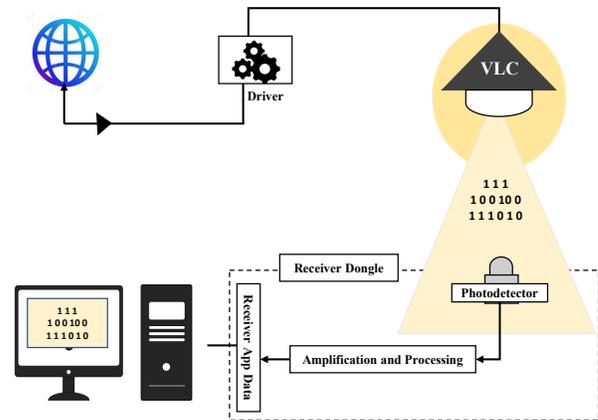


Fig. 2. Indoor VLC communication link is providing illumination and communication services in indoor scenario.

1) *OpenVLC*: This is one of the most common open-source and adaptable platforms available to the VLC research community. It emits a *manchester-coded on-off-keying* (MC-OOK) modulated signal. This translates each bit into two transition levels, such as 0 bit and a 1 bit as 'HIGH' during the first and second half of the symbol period, respectively, and otherwise it is 'LOW', as enunciated in (6). The transmitted data signal for $j \in \{0, 1\}$ can be formulated as [17]:

$$x_{MC-OOK}(t) = \sum_{n=0}^{+\infty} v_j(t - nT) \quad (5)$$

where,

$$v_j(t) = \begin{cases} 1, & t \in [0, T/2] \text{ for } j = 0 \\ 1, & t \in [T/2, T] \text{ for } j = 1 \\ 0, & \text{elsewhere} \end{cases} \quad (6)$$

where, $v_j(t)$ denotes the pulse waveform, and T is the period of each bit. Let us assume that $x_{MC-OOK}(t)$ is the transmitted

data stream. In reality, the emitted optical signal $x_{\text{MC-OOK}}(t)$ does not exactly coincide with (5) due to the low-pass behavior of the LEDs. Therefore, our realistic theoretical model exploits the definition of convolution between the emitted signal (5) and the LED impulse response function, $h_{\text{LED}}(t) = e^{-t/\tau}/\tau$, which is modeled as a first-order low-pass filter. This operation results in a probing signal for the sensing and transmitted signals for communication, as expressed by [46],

$$p_{\text{TX}}^{\text{VLC}}(t) = x_{\text{MC-OOK}}(t) * h_{\text{LED}}(t) \quad (7)$$

The optical signal propagates via the LoS channel model and is received by the DL receiver (i.e., PD). The PD translates the light intensity into an electrical output signal $y_{\text{DL}}^{\text{PD}}(t)$. The probing signal is further encapsulated as the received DL signal by performing the convolution between the probing signal (7) and the LoS channel response given in (9). This can be mathematically represented as [46],

$$y_{\text{DL}}^{\text{PD}}(t) = p_{\text{TX}}^{\text{VLC}}(t) * h_{\text{LoS}}(t) + n(t) \quad (8)$$

where, $*$ denotes the convolution operator and $n(t)$ is the independent Additive White Gaussian Noise (AWGN) with zero-mean and unit-variance. Moreover, the LoS channel response is a time-shifted and scaled delta function that signifies the amplitude drop and delay of the transmitted data signal. As a result, optical path loss becomes a significant parameter for characterizing the LED illumination potential. Indoor VLC channels may contemplate the effect of both the LoS and the non-LoS components at the receiver end. LoS seems to be the most frequent method used for indoor optical wireless communication and illumination settings. In line with [50], [51], the LoS channel resides between the optical TX and RX. In this work, we ignored the non-LoS components of the optical link without sacrificing generality and considered a single LoS channel between TX and RX. The LoS *channel impulse response* (CIR) can be written as,

$$h_{\text{LoS}}(t) = \eta \delta(t - \Delta t) \quad (9)$$

where η is the attenuation coefficient that is introduced due to the optical channel propagation losses (path loss), Δt is the propagation time (delay-offset of the LoS path) that the optical signal undergoes in free space, and $\delta(t)$ is the Dirac delta function.

2) LiFiMAX: LiFi is a wireless networking technology that has recently been developed for commercial applications as a fully-networked device. We used the LiFiMAX system in this work. It can be easily installed on the ceilings of a conference room or workplace and it provides network access to any device equipped with a plug-and-play LiFiMAX dongle. This enables internet connectivity for 16 users within 28 m² cell size with throughputs of 60 Mbps and 40 Mbps in the DL and UL respectively. LiFiMAX uses *carrierless amplitude and phase* (CAP) modulation, one of the efficient spectral methods to overcome modulation bandwidth challenges in VLC. CAP modulation has seen extensive research interest in VLC applications as a result of its excellent spectral efficiency as well as its simplicity [52]. The transmitted signal can be

expressed by,

$$x_{\text{CAP}}(t) = \sum_{n=-\infty}^{+\infty} [a_n p(t - nT) - b_n \hat{p}(t - nT)] \quad (10)$$

where a_n and b_n are the real and imaginary parts of the n^{th} symbol, T is the time period, and n is the symbol index. The pulse-generating orthogonal filters are given as follows:

$$p(t) = g(t) \cos(\omega_c t), \quad \hat{p}(t) = g(t) \sin(\omega_c t) \quad (11)$$

where $g(t)$ is the root raised cosine filter (RRC), ω_c is the angular frequency of the filter, and T is the symbol time period. CAP is a variant modulation scheme of *quadrature amplitude modulation* (QAM) signals for LiFi communication [53]. The received signal is further demodulated and decoded to retrieve the original data stream. Note that the uplink VLC signals are not considered in this case. For the sensing purpose, the received signal at the reference PD is an analog signal that needs to be converted to the binary signal for the synchronization of the ToF camera.

C. Passive ToF Sensing Model

In the bistatic passive ToF paradigm, opportunity illumination sources (i.e., VLC sources) illuminate the scene, and the ToF camera acquires the reflected signal. We studied a PB-ToF method based on VLC-modulated light signals. The VLC signal interacts with the scene, the observed ToF signal is affected by the *scene response function* (SRF), defined in (12), denoted by $h_s(t)$, where s represents the scene. Single reflection $K = 1$ is often considered in the general setting of the ToF camera. In the most general case, ($K \geq 1$), a single ToF pixel may receive multiple reflections from the scene. In this context, the SRF is given by the expression [54],

$$h_s(t) = \sum_{k=0}^{K-1} \Gamma_s[k] \delta(t - \tau_s[k]), \tau_s[k] = 2d_s[k]/c \quad (12)$$

where $\delta(t)$ denotes the Dirac delta function and $\{\Gamma_s[k], \tau_s[k]\}_{k=0}^{K-1}$ are the reflective components of the targets and corresponding delays which are introduced by K back-scattered light paths [14]. The reflected signal is obtained by performing the convolution between the probing signal and the SRF, i.e., $r(t) = (p_{\text{TX}}^{\text{VLC}} * h_s)(t)$ and this follows (7). The reflected signal is given by [55],

$$r(t) = (p_{\text{TX}}^{\text{VLC}} * h_s)(t) \quad (13)$$

where the SRF is the shift-invariant response function. The conventional CW-ToF method uses the internal signal generated by the ToF system as DCS [14], [56]. Differently, in this work, we exploited the thresholded version of the *Reference Photodiode* (RPD) signal, $y_{\text{RPD}}(t)$, as the DCS. This is a binary signal used to synchronize the ToF camera.

The received signal is sampled at every bit duration T_b . Therein, we employed a thresholding scheme that compares the received signal amplitude to the threshold voltage V_{th} ; this allows us to make a decision and generate logical levels *Low* and *High*. The signal prior to thresholding is denoted as the

RPD signal, $y_{\text{RPD}}(t)$, and the thresholding operation can be written as follows,

$$y_{\text{RPD}}^{\text{th}}(t) = \begin{cases} A, & y_{\text{RPD}}(T_b) \geq V_{\text{th}} \\ 0, & y_{\text{RPD}}(T_b) \leq V_{\text{th}} \end{cases} \quad (14)$$

where A is the required amplitude level of the ToF reference signal. Let us first explain the measurement gathering process to gain insights into how our proposed method processes the acquired data. During the acquisition time interval, the VLC emitter transmits signals continuously, which is long enough to capture the reflected light signals. We can work with a fixed time window of size τ , which is facilitated by the main lobe of the autocorrelation of VLC signals. This allows us to use a given range of samples in simulation and hardware experiments and accounts for the finite exposure time of the ToF camera. The cross-correlation between (13) and (14) provides continuous-time measurements $m(t)$ as given in (15). The measurements are obtained by shifting the control signal within a given range of delays, related to the depth range to cover. This yields samples of the cross-correlation function.

$$m(t) = \int r(\tau)y_{\text{RPD}}^{\text{th}}(t, \tau)d\tau \quad (15)$$

where τ is the shift component. The discrete measurements can be written as $m[i] = m(t)|_{t=iT, i \in \mathbb{Z}}$, where $0 \leq t < T$ is the sampling range.

IV. DEPTH RECOVERY APPROACH

The focus of this section is now shifted to depth reconstruction using matched filtering. Now, our discussion is linked to our numerical experiments, how we built our simulations from real optical signals and attempted to invert the model. Later, we will use real measurements from PMD pixels. The starting point is the generation of measurements in order to emulate our numerical experiments. We briefly discuss the measurement generation process to perform matched filtering for depth recovery. The GT cross-correlation function is shifted and downsampled by a shift operator, $\mathcal{S}_q : \mathbb{R}^N \rightarrow \mathbb{R}^N$ and a downsampling operator, $\mathcal{D}_q : \mathbb{R}^N \rightarrow \mathbb{R}^Q$, defined by $[\mathcal{S}_q(\vec{Y})]_p := Y_{p-q}$ and $[\mathcal{D}_q(\vec{Y})]_p := Y_{1+(p-1)q}$, respectively. This yields a measurement vector that can be written as follows,

$$\vec{Y}_{\text{Meas}} = \mathcal{D}_R(\mathcal{S}_{u_\tau}(\vec{Y}_{\text{GT}})) \quad (16)$$

where $u_\tau = N\tau/T$ and the spacing between the samples is defined as, $R = N/Q$. For the number of measurements acquired according to the considered cases, we adhered to the standard number of measurements, namely, $Q \in \{4, 8, 16\}$, while $N \gg Q$ and Q is different representation of M as measurements. The MF correlates the test vector obtained from the reference cross-correlation vector and the measurement vector [22], [57]. To demonstrate the correct operation of MF, AWGN is added to the measurements for our simulation, which is a common channel model for noise. The SNR ranges from -30 dB to 100 dB for attaining measurement vectors. The reference cross-correlation is generated in simulations and calibrated in experiments. We obtained the test vector

by applying the same sequence of shifting and downsampling operations, but for a candidate delay to test, $\Delta\tau$, so that a vector of the same dimension as the measurement vector is obtained that is a function of $\Delta\tau$. The test vector can be generated as follows,

$$\vec{Y}_{\text{Test}}(\Delta\tau) = \mathcal{D}_R(\mathcal{S}_{u_{\Delta\tau}}(\vec{Y}_{\text{GT}})) \quad (17)$$

The estimated time-delay (18) is provided by determining the argument that maximizes the MF function, and from this, the depth estimation can be carried out as given in (19).

$$\hat{\tau} = \arg \max_{\Delta\tau} \langle \vec{Y}_{\text{Meas}}, \vec{Y}_{\text{Test}}(\Delta\tau) \rangle \quad (18)$$

where $\langle \cdot, \cdot \rangle$ denotes the standard inner product of vectors. The proposed methodology allows depth reconstruction, depending on various measurement SNR values. The translation of time-shift into the distance is a simple linear computation.

$$\hat{d} = \frac{c \times \hat{\tau}}{2} \quad (19)$$

Different sampling schemes are used in the time-shift domain. The depth range is given by the width of the main lobe of the cross-correlation function. We sampled the cross-correlation function according to uniform, random, and sparse ruler [58], [59] sampling schemes.

A. Bistatic Geometry

Bistatic ToF imaging is a new line of research; bistatic configurations of emitters and receivers have not yet been explored in 3D ToF imaging. In the passive modality setting, the estimated depth defines a 3D ellipsoid where the target point may lie due to the bistatic geometry. The foci of the ellipsoid are the VLC emitter and the ToF receiver. Provided that we know the observation direction of each pixel of the ToF array thanks to lens calibration, we can compute the intersection between the ellipsoid and the direction vector. This defines the 3D location of the target point, thus retrieving the accurate depth between the camera and the target. The 3D location of the target can be written as follows:

$$\mathbf{T} = \mathbf{R} + \vec{u}_{\text{R}}d_{\text{RT}} \quad (20)$$

where \mathbf{T} is the target position, \mathbf{R} is the position of the receiver, \vec{u}_{R} is the unit vector, and d_{RT} is the distance between the receiver and the target. The 3D ellipsoid is defined by,

$$d - d_o = d_{\text{ET}} + d_{\text{RT}} \quad (21)$$

where d_{ET} is the Euclidean distance between the emitter and the target, and d_{RT} defines the Euclidean distance between the receiver and the target. The depth offset, d_o , is associated with the used cables and electronics. It is not related with the scene geometry. The complete proof takes profit from the lens calibration and uses the observation vector \vec{u}_{R} to attain the true depth. The proof of the final depth for the bistatic geometry is provided in Appendix I and can be expressed as,

$$d_{\text{RT}} = \frac{d_{\text{ER}} - d^2}{2[(C_x u_{\text{Rx}} + C_y u_{\text{Ry}} + C_z u_{\text{Rz}}) - d]} \quad (22)$$

where $d_{\text{ER}} = \|\mathbf{E} - \mathbf{R}\|_2^2$, C_x , C_y , and C_z are defined in the Appendix I. d is the total distance that is obtained by (19).

u_x , u_{Ry} , and u_{Rz} are the observation vector components in x , y , and z directions. The demonstrated results in Section VI-E are carried out by applying (19) first and this estimated depth is further used in (22) to obtain the true depth between the receiver and the target.

B. Generalized Sampling in Time-Shift Domain

Classical sampling theory, based on the Shannon-Nyquist theorem, assumes uniform sampling, but related work has shown that non-uniform sampling schemes may reduce the required number of samples without compromising the signal reconstruction quality [60]. We analyze three different sampling schemes in the time-shift domain: uniform sampling, pseudo-random sampling, and sparse-ruler-based sampling.

1) *Uniform Sampling*: Reconstruction of a signal from a discrete number of sampling points, where the signal is sampled kT uniformly at fixed-time intervals is known as *uniform sampling*, where T is the time interval and is probably the most widely-spread sampling technique. In an arbitrarily-fine discrete domain, the samples would be located at the positions $k \in \{0, s, 2s, 3s, \dots\}$, where s is a discrete version of T . Fig. 3 (a) demonstrates uniform sampling with fixed time intervals. Provided that our signals of interest are ultimately K -sparse, with $K = 1$ in the single-bounce case. Shannon sampling yields redundant and unnecessary samples, raising the computation complexity of subsequent processing. To overcome this limitation, one can make use of non-uniform sampling [61].

2) *Non-Uniform Sampling*: Non-uniform sampling deals with sets of sampling points that are not uniformly distributed over the sampling domain. Different methods can be used to define the location of the sampling points, which may be random or deterministic. In this work, we analyze the depth reconstruction and performance obtained by employing *random* [58], [62], and *sparse ruler* ([59], and references therein) sampling on the time-shift domain. The current upsurge of interest in sampling signals at rates lower than their Nyquist rate can be credited to compressive sampling (also known as compressed sensing) [63], [64], which has fueled a substantial amount of research over the last several years, including in the ToF 3D imaging community [1].

Conclusively, we consider sparse measurements in our simulation settings. Theoretically, a sparse ruler is one that misses some marks, but it can still measure all integer distances between 0 and the ruler's length L . For example, a sparse ruler with the length $L = 23$ and $M = 8$ marks would be $\{0, 1, 2, 11, 15, 18, 21, 23\}$. The cardinality of the set M determines the number of measurements, and the value of the marks denotes the distance between each measurement sample and the reference sample. We have exploited an optimal sparse ruler of type *Wichmann W(2, 5)*. Note that, since all markers are set on a discrete grid (like a sequence), the distances between them are always expressed in terms of integers rather than time-shift units.

3) *On-grid and Off-grid sampling*: The sampling schemes considered in this work rely on an underlying regular grid. The regular or fine grid sampling is governed by the hardware (i.e., oscilloscope). It is assumed that the sample location fits within

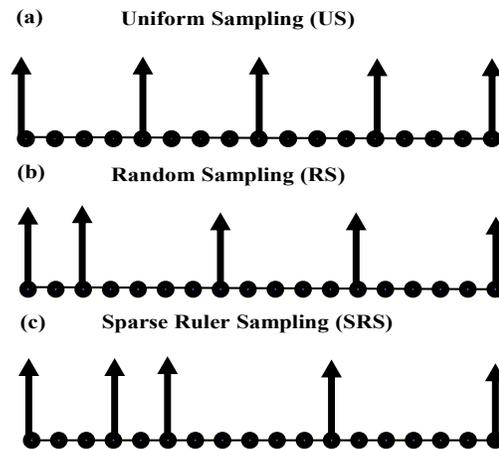


Fig. 3. Sampling methods. (a) A common sampling approach is uniform sampling. (b) Pseudo-random approach samples the signal in an irregular fashion, and (c) sparse-ruler-based sampling.

the fine grid's resolution. For uniform and random sampling methods, on-grid samples can always be generated, regardless of the step size of the fine grid. In the context of the sparse-ruler sampling method, samples may not coincide with the on-grid locations. In this case, the signal is further down-sampled to match the fine grid resolution required by the sparse ruler. This results in a grid mismatch or locations which are off-grid [65]. One has to deal with the grid mismatch between the fine grid and the sparse-ruler-based sampling grid (off-grid). To this end, the interpolation method can be used to predict the samples located between the fine-grid GT samples and form a smooth curve between on-grid and off-grid sample locations.

V. EXPERIMENTAL SETUP

In this section, the focus is on designing a bistatic setup for passive ToF imaging. We provide the details of the hardware components which are used in the setup. The bistatic geometry of the VLC-enabled passive ToF setup is demonstrated in Fig. 4(a). Our proposed experimental setup is depicted in Fig. 4(b). The simulations and experiments are developed based on our proposed pipeline (cf. Fig. 1). We used an evaluation board containing a PMD camera module endowed with built-in external reference capabilities to control the PMD pixels. This module needs binary signals with 0 V as low and 1.8 V as high voltage levels of the external reference signal.

In this case, we employed a thresholding operation that emulates the thresholding circuit to digitize the analog signal for external referencing. This circuit should come before the evaluation board. For the thresholding operation in our experimental settings, we used a pulse generator (HP 8082A) that carries out the thresholding operation. The key parameters of ToF setup and evaluation board are given in Table. I and Table. II. VLC communication experiments were carried out in dark and bright room conditions. ToF sensing tests were conducted in the darkroom settings. In the experimental setup, we used an optical rail to move the emitter and receiver easily. The rail width is our x , the height of the emitter/receiver is y , and the depth is denoted by z . The length of the rail is 1 m.

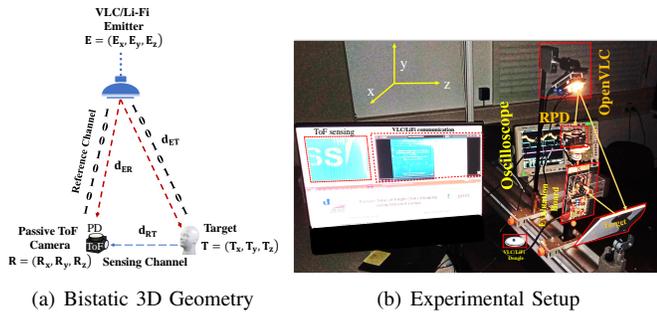


Fig. 4. (a) Bistatic geometry of passive ToF sensing. (b) Experimental setup for simultaneous communication and sensing.

TABLE I

EXPERIMENTAL KEY PARAMETERS OF VLC-ENABLED PASSIVE TOF SETUP

Modules	OpenVLC1.3	LiFiMAX
Modulation	MC-OOK	CAP
Modulation Frequency	1 MHz	50 MHz
Throughput	400 kbps	65 Mbps
LED spectrum	VL LED	IR LED
Transmit Power	3 mW	<7 W
Oscilloscope Bandwidth (MDO4104-6)	up to 1 GHz	
PD Bandwidth (PDA10A-EC)	up to 150 MHz	
PD responsivity	0.5 A/W	
PD active area and rise time	0.8 mm and 2.3 ns	
Threshold circuit (HP-8082A)	-1.5 to 1.5 V	
Delayer (PSD-MOD)	Delay range 50 ns	

TABLE II

PMD MODULE PARAMETERS

Model No	IRS 2877C
Resolution	153k pixels
Frame rate	5 FPS
Exposure time	5 ms
Channels	A and B
TV Lens	25 mm

Here we have provided a procedure for the proposed method and the process for acquiring data as follows:

- Communication signals are transmitted continuously via OpenVLC/LiFiMAX. We placed the emitter at the position of $E = (5, 47.3, 56.3)$ cm.
- The reference photodiode (RPD) is placed at the top of the ToF sensor. It is responsible for obtaining the optical signal for synchronizing the ToF camera. This establishes a direct link between the emitter and the RPD.
- The RPD signal is inserted into the thresholding circuit, converting the analog signal to the digital signal. Furthermore, the thresholded signals are fed to the picosecond delayer (PSD), enabling custom delays. In our case, we selected our range of scenarios based on the range offered by the PSD.
- The output signal of the PSD is launched into the evaluation board after the necessary voltage-level adaptation.
- On the flip side, the reflections are acquired by the PMD camera. The ToF camera is mounted on the optical rail at the position of $R = (7, 21.5, 60)$ cm. The evaluation board has a reference mixer, which takes the reflected

signal and mixes it with the external reference signal. Assuming that the signal which is fed into the sensor board is a thresholded representation of the VLC signal. This yields a cross-correlation operation and by delaying the signal, we can obtain measurements.

- The target is placed at a distance of 25 cm from the ToF evaluation board.

In our case, we used a *MiraVera PMD version* that features a high-resolution of 640×240 pixels. The Bohler star is used as a target in 3D. This is useful for evaluating the camera's real lateral resolution.

VI. RESULTS AND DISCUSSION

A. Evaluation of OpenVLC and LiFiMAX modules

For evaluating the performance of both exploited modules, we obtained optical signals using a Thorlabs photodiode (PDA 10A-EC) via an Agilent Oscilloscope (MDO4104-6). The optical signals can be seen in Fig. 5 (a) and (b). The random data signals are generated from both modules (OpenVLC and LiFiMAX). These signals are used to emulate the passive modality. The acquired data signals are thresholded by following (14) with respect to their local average. Fig. 5 (c) and (d) show the thresholded signals of both modules. The output signal of the thresholded signal is kept 1.8 V maximum to adhere to the PMD evaluation board requirements. The cross-correlation functions, between the optical signals and their respective thresholded versions, are demonstrated in Fig. 5 (e) and (f). The magnified parts of both signals demonstrate the expected central peaks and the provided time-span which can be used for shift-domain sampling.

B. Depth Reconstruction error in presence of noise

A number of simulations are carried out using real optical signals of both modules. We exploited the two different ranges of each of the modules. For the OpenVLC module, which has the largest time period, the range can be covered up to 150 m, but we restricted simulations to a range of 6.72 m because of our PSD, which can provide shifts only up to the considered range.

In this section, we group our results based on different sampling approaches, according to section IV-B and the number of measurements. Synthetic ToF data measurements were generated and modeled as pulse-shaped cross-correlation samples, as reported in the previous section. The shape is approximately Gaussian with standard deviation of 44.8 ns for OpenVLC and 10 ns for LiFiMAX. The measurements are generated via (15) by following the considered sampling methods in the time-shift domain. We took into account 40 realizations for each approach used.

Four samples of the correlation function are often used for phase computation in consumer devices, such as PMD cameras. Without sacrificing generality, samples are studied using a single-echo case in which the cross-correlation function range has been partitioned into several measurement samples. The depth reconstruction results of this study are depicted in box plots as shown in Fig. 6 for the OpenVLC module. We

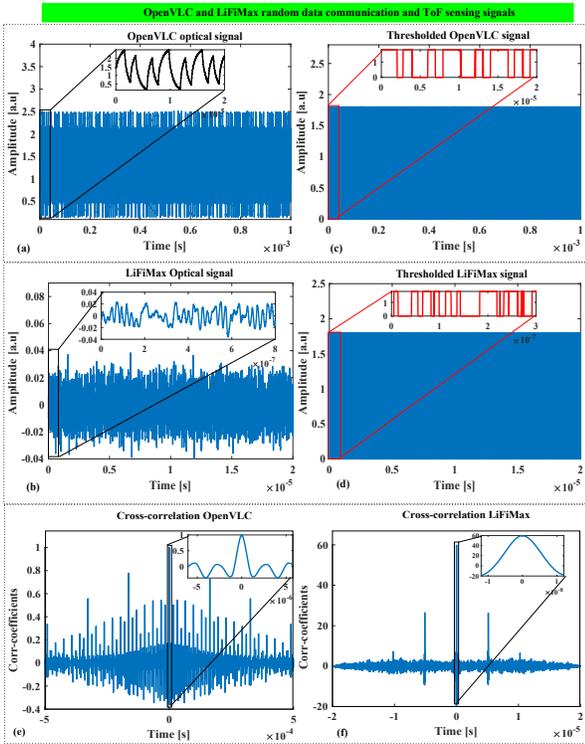


Fig. 5. OpenVLC and LiFiMAX module signal characterization.

utilized RMSE as our performance metric to evaluate the depth reconstruction performance of our method, that is,

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{n=1}^N (\hat{d}_n - d_{\text{GT},n})^2} \quad (23)$$

where N is the number of independent depth scenarios, \hat{d} is the estimated depth, and d_{GT} is the ground truth depth. The RMSE is considered in logarithmic scale $\text{RMSE}[\text{dB}] = 10 \log_{10}(\text{RMSE})$. The measurements SNR is controlled ranging from -30 dB to 100 dB. The matched filtering results are restricted up to 50 dB since a plateau at a negligible error was attained, as presented in Fig. 6.

Several depth tests are carried out for a number of considered ranges. When the SNR is greater than 0 dB, matched filtering method attained a negligible error. More comprehensive analysis showed few outliers in all sampling schemes. One may observe a large error for sparse ruler sampling due to the combined effect of noise and the off-grid phenomenon due to a mismatch between the fine grid and the grid required by the sparse ruler. The matching filtering approach performs well, with uniform and random sampling.

The matched filter performs better, since the transition from failure to successful reconstruction happens at a much lower measurement SNR. Fig. 7 presents the depth reconstruction error for matched filtering using the LiFiMAX module. The same procedure is conducted for the LiFi module. We attained a -90 dB RMSE value for uniform, -95 dB for random, and almost -80 dB for sparse ruler. It is observed here that sparse

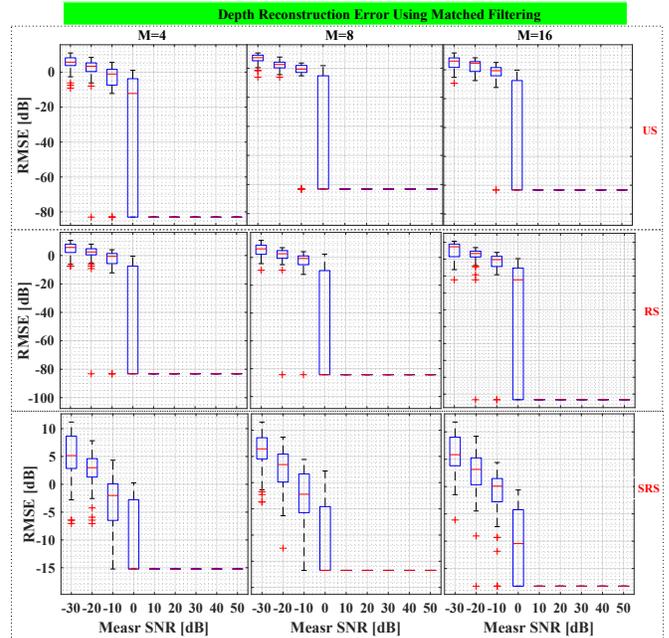


Fig. 6. Depth reconstruction RMSE obtained for the OpenVLC module at different measurements ($M = 4, 8, 16$) and SNR ranges using Matched Filtering. Uniform Sampling (US) (1st row), Random Sampling (RS) (2nd row), and Sparse Ruler Sampling (SRS) (3rd row).

ruler depth reconstruction error has a minimal difference in results with respect to the other two techniques.

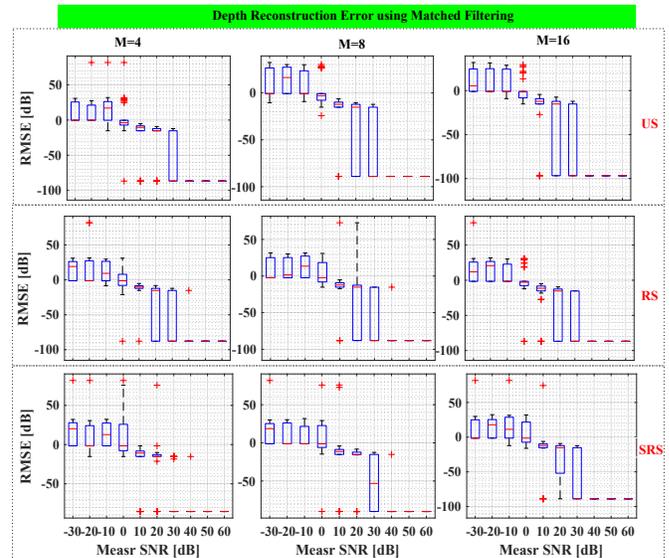


Fig. 7. Depth reconstruction RMSE obtained for the LiFiMAX module at different measurements ($M = 4, 8, 16$) and SNR ranges using Matched Filtering. Uniform Sampling (US) (1st row), Random Sampling (RS) (2nd row), and Sparse Ruler Sampling (SRS) (3rd row).

Our results demonstrate how depth retrieval from samples gained according to different sampling schemes performed in noisy environments. Assume we have noisy measurements, the noise might be from the ToF sensors or the surroundings. We attempted to provide the results of the measurement SNR and the estimated depth SNR. The effects of measurement

SNR on the estimated depth SNR were studied for both the OpenVLC and LiFiMAX modules. We provide the OpenVLC SNR results here. The estimated SNR becomes constant after 0dB SNR in matched filtering, as shown in Fig. 8.

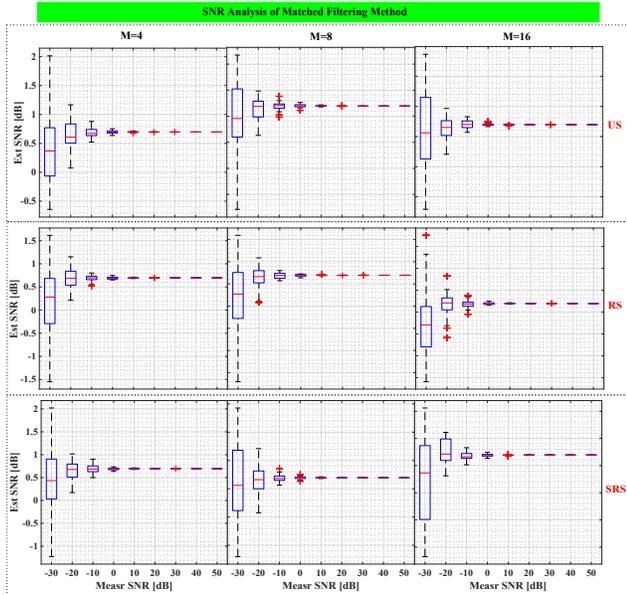


Fig. 8. SNR analysis of OpenVLC module using Matched Filtering for different measurements ($M = 4, 8, 16$) and SNR ranges. Uniform Sampling (US) (1st row). Random Sampling (RS) (2nd row) and Sparse Ruler Sampling (SRS) (3rd row).

C. Parametric Modeling

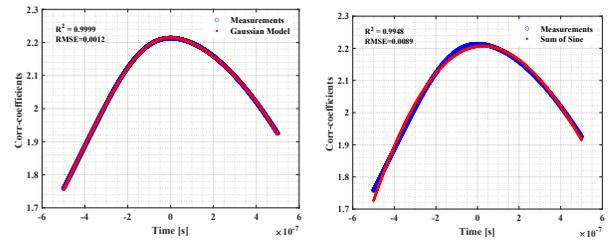
The measurements are samples of the cross-correlation model in (15). The data points were constrained to a range that meets the range of our PSD. In the time-shift domain, a window size of about 45 ns was taken into consideration. The cross-correlation was obtained from the full range. The cross-correlation is approximated as a sum of Gaussian functions in (24) and a sum of sines model in (25) for the OpenVLC. The sum of sines and Fourier models in (26), were found to be best in describing the LiFiMAX cross-correlation function.

$$\vec{Y}_{GT} = \sum_{i=1}^n a_i e^{-\left(\frac{t-b_i}{c_i}\right)^2} \quad (24)$$

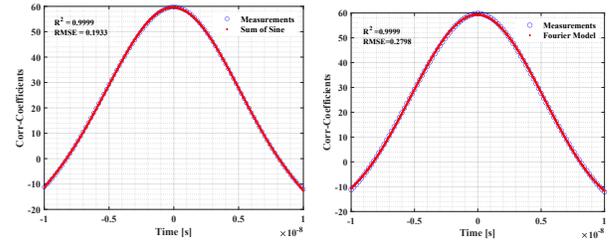
$$\vec{Y}_{GT} = \sum_{i=1}^n a_i \sin(b_i t + c_i) \quad (25)$$

$$\vec{Y}_{GT} = a_o + \sum_{i=1}^n a_i \cos(i\omega t) + b_i \sin(i\omega t) \quad (26)$$

where t is the time scale of the cross-correlation function, a_i , b_i , c_i , and ω are the model parameters and $n = 2$. The fitted models for the considered range cases are demonstrated in Fig. 9 against the points used for fitting. Plots of Fig. 9 demonstrate an excellent match with the cross-correlation functions, with $R^2 = 0.9999$ in almost all cases considered. Fig. 10 shows the depth reconstruction using the Gaussian model for the OpenVLC module and the sum of sines model for the LiFiMAX cross-correlation function. We attained successful depth reconstruction up to constant for both models.



(a) OpenVLC depth range 6.72 m (b) OpenVLC depth range 6.72 m



(c) LiFiMAX depth range 1.47 m (d) LiFiMAX depth range 1.47 m

Fig. 9. Parametric modeling results of cross-correlation for different target ranges.

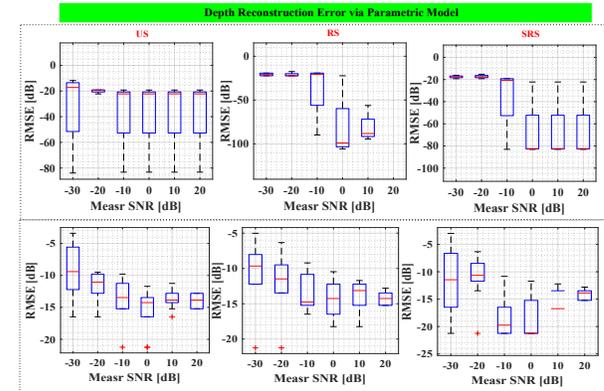


Fig. 10. Depth Reconstruction RMSE by means of parametric models using Matched Filtering for $M = 4$ measurements. OpenVLC module (1st row) and LiFiMAX module (2nd row).

D. Performance evaluation of VLC Communication

This part focuses on the communication-oriented system performance, considering both modules. Here, we concentrate our analysis on the OpenVLC module due to the fact that it is a research-based module, while the LiFiMAX module is a commercial product. We evaluated the OpenVLC module performance in terms of SNR and theoretical bit error rate (BER) as a function of distance. These metrics are used to examine the MC-OOK VLC link for indoor communication. The BER is given by $BER = Q_f(\sqrt{SNR})$, where Q_f is a quality factor that can be defined as, $Q_f = \text{erfc}\left(\frac{SNR}{\sqrt{2}}\right)$. Fig. 11 (a) represents the throughput of the OpenVLC link as a function of distance. It can be seen that this can provide illumination without affecting the communication data rate. Fig. 11 (b) shows the measured SNR and theoretical BER of an indoor VLC channel using the ThorLabs photodiode and the Agilent oscilloscope. The eye diagrams are measured after down-sampling to one sample per bit. The MC-OOK modu-

lation is used and an eye diagram is generated by making use of OpenVLC random data signals. The eye-diagram settings are enabled in the oscilloscope to see the eye patterns. It is possible to define the quality of a transmission network using eye diagrams. The SNR can be computed as follows:

$$\text{SNR}_{\text{dB}} = 20 \log_{10} \left(\frac{V_{\text{signal}}}{V_{\text{noise}}} \right) \quad (27)$$

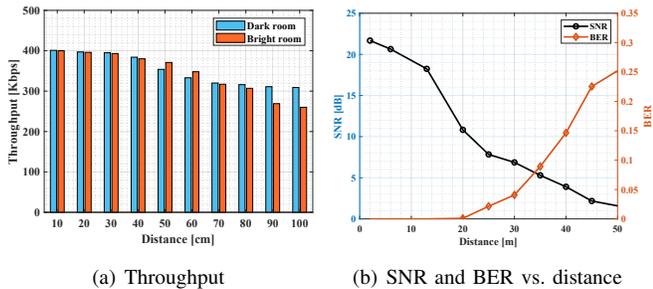


Fig. 11. Communication performance of the VLC module. (a) Measured Throughput. (b) SNR and BER curves as a function of distance.

First insights can be gained by analysis of the eye diagram's vertical and horizontal characteristics. We choose to measure the vertical amplitude and horizontal eye openings, as illustrated in Fig. 12.

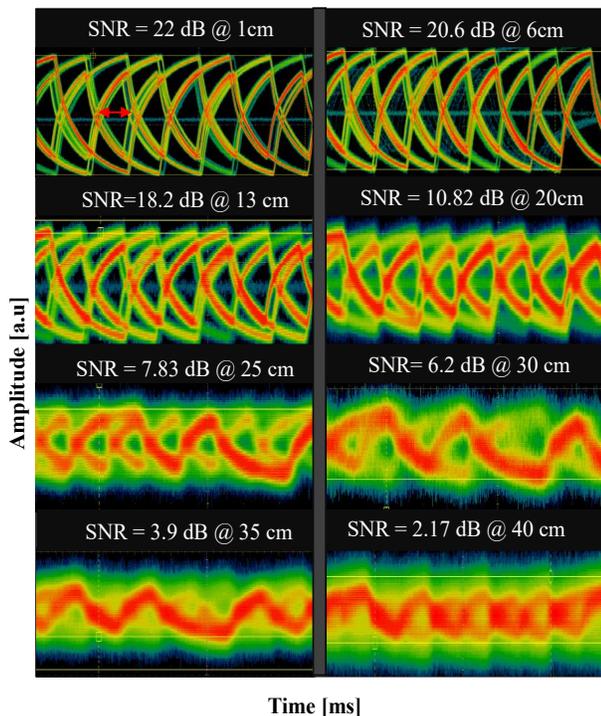


Fig. 12. Eye diagrams obtained for the optical signals of the OpenVLC module at different distance ranges.

E. Preliminary 3D Imaging Results

This experiment aims to demonstrate that the concept of passive depth recovery works in practice. To this end, we used only four uniformly-distributed samples for depth reconstruction. The results are shown in Fig. 13. Fig. 13 (a)

depicts the so-called Boehler star [66] used as a target, which is a three-dimensional representation of the Siemens star used to determine the spatial resolution of depth sensors. The VLC module is used to acquire raw images with custom delays, and the exposure time is maintained at 5 ms. The depth can be obtained via (22) using the lens normals. In the experiments, we recorded the measurements for the reference function, \vec{Y}_{GT} , in the range provided by our PSD. One can cover the complete range of the main lobe of the cross-correlation function. A plane is used to acquire the reference function data. Then, MF is used to retrieve the depth from the measurements by leveraging the recorded reference function. The 3D reconstruction in Fig. 13 is the first one ever obtained from a ToF camera without a dedicated illumination system and validates the proposed passive ToF imaging concept.

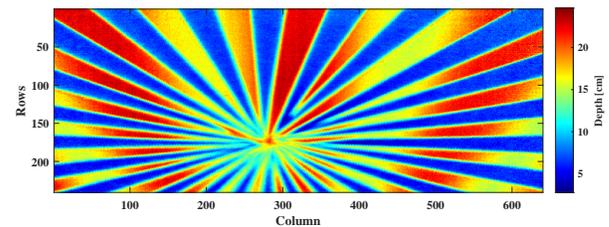
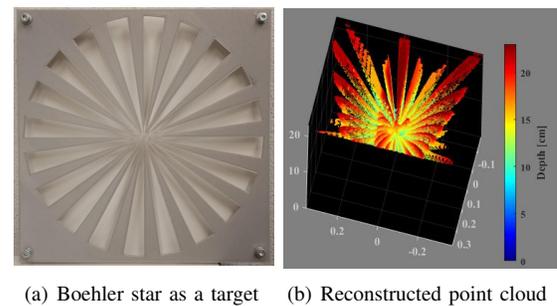


Fig. 13. First 3D reconstruction results of a real target obtained via the proposed passive ToF approach.

VII. CONCLUSION AND FUTURE OUTLOOK

In this work, we proposed a novel VLC-enabled passive imaging pipeline that allows depth estimation up to machine precision in simulations. This study opens up new possibilities for simultaneous communication and 3D sensing systems by exploiting different sampling schemes in the time-shift domain. The validity of the proposed method has been demonstrated using matched filtering, resulting in a depth reconstruction accuracy of 95% at suitable SNR levels for both modules. The key advantage of our work is a drop-in replacement of the classical ToF illumination units by existing, uncontrolled VLC sources. This method drastically reduces the power consumption of ToF cameras and eliminates temperature drift effects from measurements which are generated by the light source [67]. We have also shown that this method performs well with noisy measurement data in simulations. Low-complexity parametric models, including Gaussian and the sum of sines, have been proposed to characterize the

cross-correlation functions of both modules. Extremely low fitting errors confirmed the validity of the proposed models. Hardware experiments validate our proposed methodology by attaining the worst-case depth error 20 mm at 25 cm target distance. In addition, we acknowledged that the depth accuracy depends on the VLC source bandwidth. The communication signals are not optimized for sensing purposes. The design of modulation waveforms that are jointly optimized for both purposes is a subject of future work. On the theoretical front, we have provided a complete formulation of the direct sensing model and the method we use for solving the inverse problem. Our method raises interesting considerations regarding the hardware. We intend to explore our hardware implementation further in the future.

APPENDIX I PROOF OF EQUATION (22)

Remember that (22) is derived from (20) and (21), since the emitter and the receiver locations are known. In this regard, (21) is further re-arranged and represented in Euclidean distance between the emitter and the target. By omitting the d_o , this can be expressed as follows,

$$d - d_{RT} = \|E - T\|_2 \quad (28)$$

$$d - d_{RT} = \sqrt{(E_x - T_x)^2 + (E_y - T_y)^2 + (E_z - T_z)^2} \quad (29)$$

where $E = (E_x, E_y, E_z)$ and $T = (T_x, T_y, T_z)$ are 3D locations of the emitter and the target, respectively.

Equation (21) is broken down into 3D coordinate components and it can be defined as,

$$\begin{aligned} T_x &= R_x + u_{R_x} d_{RT} \\ T_y &= R_y + u_{R_y} d_{RT} \\ T_z &= R_z + u_{R_z} d_{RT} \end{aligned} \quad (30)$$

where $R = (R_x, R_y, R_z)$ are the 3D coordinates of the receiver and $\vec{u}_R = (u_{R_x}, u_{R_y}, u_{R_z})$ are the components of observation vector. We substitute (30) in (29) and, after carrying out some re-arrangements and manipulation, we get,

$$\begin{aligned} d^2 - 2dd_{RT} + d_{RT}^2 &= (C_x^2 + C_y^2 + C_z^2) \\ &+ d_{RT}^2 (u_{R_x}^2 + u_{R_y}^2 + u_{R_z}^2) \\ &- 2d_{RT} (C_x u_{R_x} + C_y u_{R_y} + C_z u_{R_z}) \end{aligned} \quad (31)$$

where $d_{ER}^2 = C_x^2 + C_y^2 + C_z^2$ is the Euclidean distance between emitter and receiver, where $C_x = E_x - R_x$, $C_y = E_y - R_y$, and $C_z = E_z - R_z$ are the differences of x , y , and z coordinates of emitter and receiver, respectively. At this point, it is noted that the sum of normal vector components is equal to unity, ($u_{R_x}^2 + u_{R_y}^2 + u_{R_z}^2 = 1$). By substituting this in (31) and making some re-arrangements, we obtain the bistatic depth recovery formulation as given in (22).

REFERENCES

[1] M. Heredia Conde, "Compressive sensing for the photonic mixer device," in *Compressive Sensing for the Photonic Mixer Device*, pp. 207–352, Springer, 2017.

[2] A. Tontini, L. Gasparini, and M. Perenzoni, "Numerical model of spad-based direct time-of-flight flash lidar cmos image sensors," *Sensors*, vol. 20, no. 18, p. 5203, 2020.

[3] A. D. Griffiths, H. Chen, D. D.-U. Li, R. K. Henderson, J. Hermsdorf, M. D. Dawson, and M. J. Strain, "Multispectral time-of-flight imaging using light-emitting diodes," *Optics express*, vol. 27, no. 24, pp. 35485–35498, 2019.

[4] H. Elgala, R. Mesleh, and H. Haas, "Indoor broadcasting via white leds and ofdm," *IEEE Transactions on consumer electronics*, vol. 55, no. 3, pp. 1127–1134, 2009.

[5] L. Li, P. Hu, C. Peng, G. Shen, and F. Zhao, "Epsilon: A visible light based positioning system," in *11th USENIX Symposium on Networked Systems Design and Implementation (NSDI 14)*, pp. 331–343, 2014.

[6] M. Kavehrad, "Sustainable energy-efficient wireless applications using light," *IEEE Communications Magazine*, vol. 48, no. 12, pp. 66–73, 2010.

[7] D. Karunatilaka, F. Zafar, V. Kalavally, and R. Parthiban, "Led based indoor visible light communications: State of the art," *IEEE Communications Surveys & Tutorials*, vol. 17, no. 3, pp. 1649–1678, 2015.

[8] G. B. Nair and S. Dhoble, "Current trends and innovations," *The Fundamentals and Applications of Light-Emitting Diodes*, p. 253, 2021.

[9] L. E. M. Matheus, A. B. Vieira, L. F. Vieira, M. A. Vieira, and O. Gnawali, "Visible light communication: concepts, applications and challenges," *IEEE Communications Surveys & Tutorials*, vol. 21, no. 4, pp. 3204–3237, 2019.

[10] H. Abumarshoud, L. Mohjazi, O. A. Dobre, M. Di Renzo, M. A. Imran, and H. Haas, "LiFi through reconfigurable intelligent surfaces: A new frontier for 6G?," *arXiv preprint arXiv:2104.02390*, 2021.

[11] S. Soderi and R. De Nicola, "6G networks physical layer security using rgb visible light communications," *IEEE Access*, 2021.

[12] A. H. Azhar, T.-A. Tran, and D. O'Brien, "A gigabit/s indoor wireless transmission using mimo-ofdm visible-light communications," *IEEE photonics technology letters*, vol. 25, no. 2, pp. 171–174, 2012.

[13] S. Zhang, "High-speed 3D shape measurement with structured light methods: A review," *Optics and Lasers in Engineering*, vol. 106, pp. 119–131, 2018.

[14] A. Bhandari, M. Heredia Conde, and O. Loffeld, "One-bit time-resolved imaging," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 42, no. 7, pp. 1630–1641, 2020.

[15] Y. He and S. Chen, "Recent advances in 3d data acquisition and processing by time-of-flight camera," *IEEE Access*, vol. 7, pp. 12495–12510, 2019.

[16] B. Langmann, *Wide Area 2D/3D Imaging: Development, Analysis and Applications*. SpringerLink : Bücher, Springer Fachmedien Wiesbaden, 2014.

[17] F. Ahmed, M. Heredia Conde, and O. Loffeld, "Pseudo-passive indoor tof sensing exploiting visible light communication sources," in *2021 IEEE Sensors*, pp. 1–4, IEEE, 2021.

[18] A. Tsiatmas, C. P. M. J. Baggen, F. M. J. Willems, J. M. G. Linnartz, and J. W. M. Bergmans, "An illumination perspective on visible light communications," *IEEE Communications Magazine*, vol. 52, no. 7, pp. 64–71, 2014.

[19] A. Galisteo, D. Juara, and D. Giustiniano, "Research in visible light communication systems with OpenVLC1.3," in *2019 IEEE World Forum on Internet of Things (WF-IoT)*, pp. 539–544, IEEE, 2019.

[20] B. Béchadargue and B. Azoulay, "An industrial view on LiFi challenges and future," in *2020 12th International Symposium on Communication Systems, Networks and Digital Signal Processing (CSNDSP)*, pp. 1–6, IEEE, 2020.

[21] E. J. Candès *et al.*, "Compressive sampling," in *Proceedings of the international congress of mathematicians*, vol. 3, pp. 1433–1452, Citeseer, 2006.

[22] G. Turin, "An introduction to matched filters," *IRE transactions on Information theory*, vol. 6, no. 3, pp. 311–329, 1960.

[23] A. Eftekhari, J. Romberg, and M. B. Wakin, "Matched filtering from limited frequency samples," *IEEE Transactions on Information Theory*, vol. 59, no. 6, pp. 3475–3496, 2013.

[24] X. Jiang, W.-J. Zeng, H. C. So, S. Rajan, and T. Kirubarajan, "Robust matched filtering in l_p -space," *IEEE Transactions on Signal Processing*, vol. 63, no. 23, pp. 6184–6199, 2015.

[25] R. Schwarte, "Method and apparatus for determining the phase and/or amplitude information of an electromagnetic wave for photomixing," May 30 2006. US Patent 7,053,357.

[26] T. Möller, H. Kraft, J. Frey, M. Albrecht, and R. Lange, "Robust 3D measurement with pmd sensors," *Range Imaging Day, Zürich*, vol. 7, no. 8, 2005.

- [27] R. Lange, "3D time-of-flight distance measurement with custom solid-state image sensors in cmos/ccd-technology," *Ph.D. dissertation, Univ. Siegen, Dept. Elect. Eng. Comput. Sci., Germany*, 2000.
- [28] R. Lange and P. Seitz, "Solid-state time-of-flight range camera," *IEEE Journal of quantum electronics*, vol. 37, no. 3, pp. 390–397, 2001.
- [29] S. Foix, G. Alenya, and C. Torras, "Lock-in time-of-flight (tof) cameras: A survey," *IEEE Sensors Journal*, vol. 11, no. 9, pp. 1917–1926, 2011.
- [30] M. Lehmann, "Smart pixels for future 3D-tof sensors," in *Proc. IEEE Workshop CCD and Advanced Image Sensors*, 2005.
- [31] H. Rubinsztein-Dunlop, A. Forbes, M. V. Berry, M. R. Dennis, D. L. Andrews, M. Mansuripur, C. Denz, C. Alpmann, P. Banzer, T. Bauer, *et al.*, "Roadmap on structured light," *Journal of Optics*, vol. 19, no. 1, p. 013001, 2016.
- [32] Y. Zhang, L. Wang, K. Wang, K. S. Wong, and K. Wu, "Recent advances in the hardware of visible light communication," *IEEE Access*, vol. 7, pp. 91093–91104, 2019.
- [33] B. Béchadergue, L. Chassagne, and H. Guan, "Simultaneous visible light communication and distance measurement based on the automotive lighting," *IEEE Transactions on Intelligent Vehicles*, vol. 4, no. 4, pp. 532–547, 2019.
- [34] H. Sarbolandi, M. Plack, and A. Kolb, "Pulse based time-of-flight range sensing," *Sensors*, vol. 18, no. 6, p. 1679, 2018.
- [35] S. Lang, J. Zhang, Y. Cai, X. Zhu, and Q. Wu, "Classification of materials using a pulsed time-of-flight camera," *Machine Vision and Applications*, vol. 32, no. 1, pp. 1–19, 2021.
- [36] F. Wagner, F. Schiffers, F. Willomitzer, O. Cossairt, and A. Velten, "Intensity interferometry-based 3D imaging," *Optics Express*, vol. 29, no. 4, pp. 4733–4745, 2021.
- [37] R. H. Brown and R. Twiss, "A test of a new type of stellar interferometer on sirius," *Nature*, vol. 178, no. 4541, pp. 1046–1048, 1956.
- [38] R. Brown and R. Q. Twiss, "Correlation between photons in two coherent beams of light," *Nature*, vol. 177, no. 4497, pp. 27–29, 1956.
- [39] J. Boger-Lombard and O. Katz, "Passive optical time-of-flight for non line-of-sight localization," *Nature communications*, vol. 10, no. 1, pp. 1–9, 2019.
- [40] J. Herrnsdorf, J. McKendry, M. Stonehouse, L. Broadbent, G. C. Wright, M. D. Dawson, and M. J. Strain, "Led-based photometric stereo-imaging employing frequency-division multiple access," in *2018 IEEE Photonics Conference (IPC)*, pp. 1–2, IEEE, 2018.
- [41] E. Le Francois, J. Herrnsdorf, L. Broadbent, M. D. Dawson, and M. J. Strain, "Top-down illumination photometric stereo imaging using light-emitting diodes and a mobile device," in *Laser Science*, pp. JTu3A–106, Optical Society of America, 2019.
- [42] J. Herrnsdorf, J. McKendry, M. Stonehouse, L. Broadbent, G. C. Wright, M. D. Dawson, and M. J. Strain, "Lighting as a service that provides simultaneous 3D imaging and optical wireless connectivity," in *2018 IEEE Photonics Conference (IPC)*, pp. 1–2, IEEE, 2018.
- [43] Y. Almadani, D. Plets, S. Bastiaens, W. Joseph, M. Ijaz, Z. Ghassemlooy, and S. Rajbhandari, "Visible light communications for industrial applications—challenges and potentials," *Electronics*, vol. 9, no. 12, p. 2157, 2020.
- [44] A. Al-Kinani, C.-X. Wang, L. Zhou, and W. Zhang, "Optical wireless communication channel measurements and models," *IEEE Communications Surveys & Tutorials*, vol. 20, no. 3, pp. 1939–1962, 2018.
- [45] J. M. Kahn and J. R. Barry, "Wireless infrared communications," *Proceedings of the IEEE*, vol. 85, no. 2, pp. 265–298, 1997.
- [46] Z. Ghassemlooy, L. N. Alves, S. Zvanovec, and M.-A. Khalighi, *Visible light communications: theory and applications*. CRC press, 2017.
- [47] H. Henniger and O. Wilfert, "An introduction to free-space optical communications," *Radioengineering*, vol. 19, no. 2, 2010.
- [48] H. Haas, L. Yin, C. Chen, S. Videv, D. Parol, E. Poves, H. Alshaer, and M. S. Islam, "Introduction to indoor networking concepts and challenges in LiFi," *Journal of Optical Communications and Networking*, vol. 12, no. 2, pp. A190–A203, 2020.
- [49] M. S. Amjad and F. Dressler, "Integrated communications and non-invasive vibrations sensing using strobing light," in *ICC 2020-2020 IEEE International Conference on Communications (ICC)*, pp. 1–6, IEEE, 2020.
- [50] S. Long, M.-A. Khalighi, M. Wolf, S. Bourennane, and Z. Ghassemlooy, "Channel characterization for indoor visible light communications," in *2014 3rd International Workshop in Optical Wireless Communications (IWOW)*, pp. 75–79, IEEE, 2014.
- [51] S. Long, M.-A. Khalighi, M. Wolf, Z. Ghassemlooy, and S. Bourennane, "Performance of carrier-less amplitude and phase modulation with frequency domain equalization for indoor visible light communications," in *2015 4th International Workshop on Optical Wireless Communications (IWOW)*, pp. 16–20, IEEE, 2015.
- [52] R. Le Priol, M. Héliard, S. Haese, and S. Roy, "Experimental comparison of pam and cap modulation for visible light communication under illumination constraints," *IEEE Photonics Journal*, vol. 14, no. 2, pp. 1–11, 2022.
- [53] K. O. Akande, P. A. Haigh, and W. O. Popoola, "On the implementation of carrierless amplitude and phase modulation in visible light communication," *IEEE Access*, vol. 6, pp. 60532–60546, 2018.
- [54] A. Bhandari and R. Raskar, "Signal processing for time-of-flight imaging sensors: An introduction to inverse problems in computational 3-d imaging," *IEEE Signal Processing Magazine*, vol. 33, no. 5, pp. 45–58, 2016.
- [55] M. Heredia Conde, K. Kagawa, T. Kokado, S. Kawahito, and O. Loffeld, "Single-shot real-time multiple-path time-of-flight depth imaging for multi-aperture and macro-pixel sensors," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 1469–1473, IEEE, 2020.
- [56] M. Heredia Conde, T. Kerstein, B. Buxbaum, and O. Loffeld, "Fast multipath estimation for pmd sensors," in *5th International Workshop on Compressed Sensing Theory and its Applications to Radar, Sonar, and Remote Sensing (CoSeRa 2018)*, 2018.
- [57] F. Boßmann, S. Krause-Solberg, J. Maly, and N. Sissouno, "Structural sparsity in multiple measurements," *IEEE Transactions on Signal Processing*, vol. 70, pp. 280–291, 2021.
- [58] P. Maechler, N. Felber, and A. Burg, "Random sampling ADC for sparse spectrum sensing," in *2011 19th European Signal Processing Conference*, pp. 1200–1204, IEEE, 2011.
- [59] Y. Zhang, Y. Wang, Z. Tian, G. Leus, and G. Zhang, "Efficient super-resolution two-dimensional harmonic retrieval with multiple measurement vectors," *IEEE Transactions on Signal Processing*, 2022.
- [60] M. Heredia Conde, A. Bhandari, and O. Loffeld, "Nonuniform sampling of echoes of light," in *2019 13th International conference on Sampling Theory and Applications (SampTA)*, pp. 1–4, IEEE, 2019.
- [61] T. Beyrouthy, L. Fesquet, and R. Rolland, "Data sampling and processing: Uniform vs. non-uniform schemes," in *2015 International Conference on Event-based Control, Communication, and Signal Processing (EBCCSP)*, pp. 1–6, IEEE, 2015.
- [62] J. Laska, S. Kirolos, Y. Massoud, R. Baraniuk, A. Gilbert, M. Iwen, and M. Strauss, "Random sampling for analog-to-information conversion of wideband signals," in *2006 IEEE Dallas/CAS Workshop on Design, Applications, Integration and Software*, pp. 119–122, IEEE, 2006.
- [63] D. Donoho, "Compressed sensing," *IEEE Transactions on Information Theory*, vol. 52, no. 4, pp. 1289–1306, 2006.
- [64] E. J. Candes and M. B. Wakin, "An introduction to compressive sampling," *IEEE Signal Processing Magazine*, vol. 25, no. 2, pp. 21–30, 2008.
- [65] A. Abtahi, S. Gazor, and F. Marvasti, "Off-grid localization in mimo radars using sparsity," *IEEE Signal Processing Letters*, vol. 25, no. 2, pp. 313–317, 2018.
- [66] W. Boehler, M. B. Vicent, A. Marbs, *et al.*, "Investigating laser scanner accuracy," *The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. 34, no. Part 5, pp. 696–701, 2003.
- [67] M. Heredia Conde, T. Kerstein, B. Buxbaum, and O. Loffeld, "Near-infrared, depth, material: Towards a trimodal time-of-flight camera," *IEEE Sensors Journal*, vol. 22, no. 12, pp. 11271–11279, 2022.



Faisal Ahmed received his B.E. degree in Telecommunication Engineering and MEng in Telecommunication Engineering and Management from Mehran UET Jamshoro in 2016 and 2019 respectively. He had an opportunity to complete his Master thesis as an Erasmus exchange student. In collaboration under the Erasmus exchange scholarships between Mehran UET and the University of Malaga, Spain. Besides, he served as a visiting faculty member at IICT, University of Sindh, Pakistan. In 2020, he joined the Marie Skłodowska-Curie Innovative Training Network Project MENELAOS^{NT} as an Early-Stage Researcher at the Center for Sensor Systems (ZESS), University of Siegen, Germany, where he is currently pursuing his Ph.D. degree in the field of optical sensing and communication from the same university. His research interests include Visible Light Communication, Time-of-Flight imaging, compressive sensing, and indoor 3D sensing.



Miguel Heredia Conde (M'14) received the Dr. Eng. degree in the field of sensor signal processing from the University of Siegen, Siegen, Germany, in 2016. In 2013, he joined the Center for Sensorsystems (ZESS), at the University of Siegen. Since then he has also been a member of the Research Training Group GRK 1564 "Imaging New Modalities". Since 2016 he is the Leader of the research group "Compressive Sensing for the Photonic Mixer Device" and since 2020 also the General Manager of

the H2020-MSCA-ITN "MENELAOS^{NT}". His current research interests include Time-of-Flight Imaging systems, such as those based on the Photonic Mixer Device (PMD), Compressive Sensing, Computational Imaging, and unconventional sensing. Miguel Heredia was one of the recipients of the 2006 Academic Excellence Prizes, awarded by the Government of Galicia, Spain. In 2017 he was awarded the University of Siegen Prize for International Young Academics, for the excellent performance in his doctoral studies. He is a member of the ITG/VDE and the IEEE/SPS.



Paula López Martínez received the Ph.D. degree in physics from the University of Santiago de Compostela, Spain, in 2003. Later, she held a two years post-doctoral position with the Fraunhofer Institute for Integrated Circuits, Erlangen, Germany. Her main research focus have been the design of mixed signal integrated circuits particularly CMOS imagers and the physical modeling of electronic devices. She is currently an Associate Professor with the Centro Singular de Investigación en Tecnoloxías Intelixentes

(CiTIUS), University of Santiago de Compostela. Her current research interests are the design of smart image sensors and low power circuits for energy harvesting and the IoT applications.



Thomas Kerstein studied Electrical Engineering with focus on automation technologies at the University Siegen, Germany and received the Dipl.-Eng. degree in 2008 and the Dr.-Eng. degree in Computer-Assisted Surgery at the Center for Sensor Systems (ZESS) at University Siegen, in 2014. Dr. Kerstein is System Engineer at the Research and Development Department of pmstechnologies ag in Siegen, Germany.



Bernd Buxbaum received the Dipl.-Eng. degree in Electrical Engineering from University of Technology Darmstadt, Germany in 1997 and the Dr.-Eng. degree in Optoelectronics at the Center of Sensor Systems (ZESS) at University of Siegen, in 2002. From 1997 to 2001 he worked at the ZESS on the research of 3D Time-of-Flight technology. Dr. Buxbaum is CEO and Co-Founder of pmstechnologies ag in Siegen, Germany and Executive Board Member of ifm electronic group of companies in Essen, Germany.

He is the author of several books and numerous published articles and gives lectures on 2D/3D image sensing at the University of Siegen, Germany.