Fixed-pattern noise modeling and removal in Time-of-flight sensing

M. Georgiev, R. Bregovic, A. Gotchev

***Abstract*—In this article, we discuss the modeling and removal of fixed pattern noise (FPN) in photonic mixture devices employing the Time-of-flight principle for range measurements and scene depth estimation. We present a case which arises from low-sensing (LS) conditions caused by either external factors related with scene reflectivity or internal factors related with the power and operation mode of the sensor or both. In such case the FPN becomes especially dominating and invalidates previously adopted noise models, which have been used for removal of other noise contaminations in ToF measurements. To tackle LS cases we propose a noise model specifically addressing the presence of FPN and develop relevant FPN removal procedure. We demonstrate, by experiments with synthetic and real-world data, that the proper modeling and removing of FPN is substantial for the subsequent Gaussian denoising and yields accurate depth maps comparable with the ones obtainable in normal operating mode.**

***Index Terms*— Time-of-Flight (ToF), Photonic Mixer Device, Fixed-pattern noise (FPN), low-sensing environment, denoising**

I. Introduction

Various automotive, entertainment, and gaming applications deal with 3D visual scenes. For these, it is important to measure or estimate distances between an observer at sensor position and scene objects. Relevant technologies are termed as range or depth sensing and include approaches based on structured light, LiDAR, depth from stereo, and Time-of-flight (ToF) active sensing. The latter approach has become especially attractive as it delivers the scene geometry information simultaneously in a form of 2D (range, depth) map, which can be interpreted as a 2D gray-scale image [1], [2], [3]. In a ToF device, range data is measured by the elapsed time during which a light signal travels from the device to the scene (object) and back to the device. Contemporary ToF devices have achieved substantial progress in operating outdoors beyond a kilometer range [4], and delivering higher range resolution [5]. However, they still have a rather low spatial resolution (e.g. 204 by 204 pixels) and relatively high power consumption [6], [7]. For keeping the measurement uncertainty low, ToF devices require a strong reflection signal [8], [9], [10]. This is problematic for low-reflectivity surfaces as well as when such devices need to be integrated in mobile devices where power-consumption of electronic elements is vital for the feasible integration. In both cases, denoising of captured range data has to be applied [11], [12], [13]. For experimental purpose, such problematic sensing environment can be easily simulated by intentionally reducing the power of the system – in this paper also referred to as low-sensing (LS) environment [12], [14], [15], [16].

In this paper, we provide noise analysis for a ToF device operating in LS environment. We discuss the role of various noises present in the signal and suggest proper order of denoising and develop and compare appropriate denoising methods. We put special emphasize on removing the so-called fixed-pattern noise (FPN) [8], [9], [17], [18], [19]. FPN is a well-known phenomenon in sensors for infrared light [17], [18]. Here we develop the FPN modeling and removal taking into account the ToF sensor properties. We also develop further our linear filtering approach presented in [20] for utilizing the periodic pattern of FPN [8], [19] for its removal. After removing FPN, we employ several modern denoising algorithms [21], [22], [23] for removing the random noise. The correctness of the presented noise analysis and model and the overall denoising procedure are demonstrated by means of several experiments on synthetic and real data. For experiments, we use devices based on the Photonic Mixer Device (PMD) sensor technology [6], [9], [24], [25] – this being one of the most popular practical implementation of ToF sensing hardware.

The outline of the paper is as follows: In Section II we give a brief overview of the ToF sensing principle and define what we call LS environment. Section III introduces a noise model of a ToF system with a special emphasizes on the FPN. It is followed by description of filter design for FPN removal in Section IV. Finally, in Section V, the overall denoising procedure is demonstrated by means of examples and some concluding remarks are given in Section VI.

A. Notations

In the paper, P and Q are used to denote the observed (measured) value and true value of data, respectively. In vector form they are denoted in bold, that is, $\boldsymbol{Q} = [Q_0\ Q_1\ Q_2\ Q_3]$. The data is organized in *spatial phase frames*, which are 2D functions of spatial coordinates (pixels) $Q(x, y)$. Subscripts k and l in $Q_{k,l}$ refer to the number of phase frames within an observation period, and observation number, respectively. If not mentioned otherwise, it is assumed that $k = 0,1,2,3$. Accents hat (e. g. $\hat{Q}$) and tilde (e. g. $\tilde{Q}$) are used for denoting the estimated true value and true value with added Gaussian noise, respectively. Values $A_{\boldsymbol{Q}}$, $\varphi_{\boldsymbol{Q}}$, and $C_{\boldsymbol{Q}}$ refer to the amplitude, phase, and offset of the sinusoidal

Manuscript received June 07, 2015. The work was supported by the Academy of Finland, grant No. 137012: 'High-Resolution Digital Holography: A Modern Signal Processing Approach'.

M. Georgiev, R. Bregovic and A. Gotchev are with the Department of Signal Processing, Tampere University of Technology, P.O.Box 553, 33101 Tampere, Finland (e-mail: mihail.georgiev@tut.fi, robert.bregovic@tut.fi, atanas.gotchev@tut.fi).

function used for estimating the depth, respectively, with sub index determining the phase frame used for their evaluation. To avoid doubt, we reserve the term 'removal' for tackling the FPN, while we use the term 'denoising' for tackling the additive Gaussian noise.

II. Basic Principles of Time of Flight Operation

For the purpose of FPN analysis and modeling, in this section we give a brief overview of the ToF sensing principle [8], [9], [10] and its performance in LS environment [12], [14], [15], [16].

A. Time-of-flight sensing principle

A typical ToF device (c.f. Fig. 1) consists of an opto-electronic system that beams a continuously modulated harmonic light signal (typically in the near-infrared wavelength range) in moment t, and senses back the delayed light reflection from the objects in the scene at time $t + t_d$. A cross-correlation operation is applied for continuous signals of emitted $s(t)$ and reflected $r(t + t_d)$ signals. Assuming ideal system and conditions, the result of cross correlation is an continuous intensity output signal q for a given correlation (offset) parameter τ:

$$q(t+\tau) = r(t+t_d) \otimes s(t+\tau) \tag{1}$$

with

$$\begin{aligned} s(t,\omega) &= C_s + A_s \cos(\omega t + \varphi_s) \\ r(t+t_d,\omega) &= C_R + A_R \cos(\omega(t+t_d) + \varphi_R) \end{aligned} \tag{2}$$

for $\omega = 2\pi f$ and $f = 1/T$, where A denotes amplitude, φ denotes phase delay, C denotes modulation offset (or modulation contrast), ω denotes modulation frequency, and T is the modulation period.

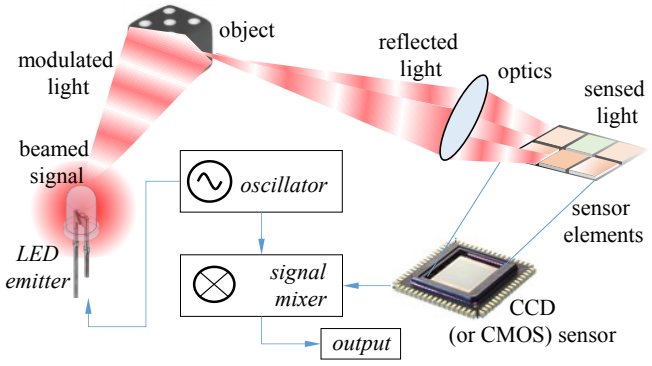

Fig. 1. A typical Time-of-flight (ToF) device and its components.

In practice, the cross-correlation in (1) is implemented by integrating the output of the correlator for a given value of τ over several signal periods depending on the integration time. For integration times that are a multiplier of the modulation period, that is $I_T = t_1 - t_0 = K\,T$, $K \in \mathbb{N}$ this output is a cosine function:

$$Q(\tau) = \int_{t_0}^{t_1} q(t+\tau)dt = C_Q + A_Q \cos(\tau + \varphi_Q) \tag{3}$$

where $\varphi_Q = \omega t_d$ is the phase from which the range can be determined. By sampling $Q(\tau)$ in four points $\boldsymbol{Q} = [Q_0, Q_1, Q_2, Q_3]$, corresponding to $\tau = 0, \pi/2, \pi, 3\pi/2$, the amplitude $A_{\boldsymbol{Q}}$, the offset $C_{\boldsymbol{Q}}$, and most importantly the phase $\varphi_{\boldsymbol{Q}}$ of the $Q(\tau)$ can be estimated as [9], [26]

$$\begin{aligned} A_{\boldsymbol{Q}} &= \frac{\sqrt{(Q_0 - Q_2)^2 + (Q_3 - Q_1)^2}}{2} \\ C_{\boldsymbol{Q}} &= (Q_0 + Q_1 + Q_2 + Q_3)/4 \\ \varphi_{\boldsymbol{Q}} &= \tan^{-1}\left(\frac{Q_0 - Q_2}{Q_3 - Q_1}\right). \end{aligned} \tag{4}$$

Finally, the measured distance D is proportional to the estimated phase φ_Q and can be evaluated as

$$D_{\boldsymbol{Q}} = k_1 \frac{\varphi_{\boldsymbol{Q}} c_L}{4\pi f} + k_2, \tag{5}$$

where c_L is speed of light in dry air (c_L~2.98×10^8[m/sec]). Parameters k_1 and k_2 are sensor-dependent constants that should be determined by calibration for a particular sensor (e.g. compensating for imperfections of the signal generator and phase offset / delays due to signal propagation through the device). Without loss of generality we can assume that k_1=1 and k_2=0. It is obvious that due to the phase-wrap, the distance can be measured only inside a range interval $\varphi_Q \in [0,2\pi] \rightarrow D \in [0, D_{MAX}]$, e.g. for $f = 20 \times 10^6[Hz]$, $D_{MAX} = c_L/2f = 7.5\,[m]$ [9]. The data samples $\boldsymbol{Q}$ obained from the ToF sensor are expressed in ToF system units (12 bit integer values). Consequently, following (4), the amplitude $A_{\boldsymbol{Q}}$ and offset $C_{\boldsymbol{Q}}$ are also in ToF system units. The phase $\varphi_{\boldsymbol{Q}}$, estimated from phase frames $\boldsymbol{Q}$ by using (4), is given in radians. Eq. (5) converts the phase to the estimated distance $D_{\boldsymbol{Q}}$ expressed in meters. If not mentioned otherwise, these units are assumed/used in the rest of this paper.

The above procedure is calculated pixel-wise for spatial phase frames $Q_k(x,y)$, in order to obtain the offset map $C_{\boldsymbol{Q}}(x,y)$, amplitude map $A_{\boldsymbol{Q}}(x,y)$, and phase delay map $\varphi_{\boldsymbol{Q}}(x,y)$. In practice, one is not able to directly measure the true values $\boldsymbol{Q}$. Instead, the observed (measured) values $\boldsymbol{P}$ are mixtures of $\boldsymbol{Q}$ values and measurement errors of various kinds. The amplitude and phase parameters are then calculated from the potentially noisy measurements $\boldsymbol{P}$. As an example, sensed data captured in normal operating mode is shown in Fig. 2 (image of the used scene is given in Fig. 3b).

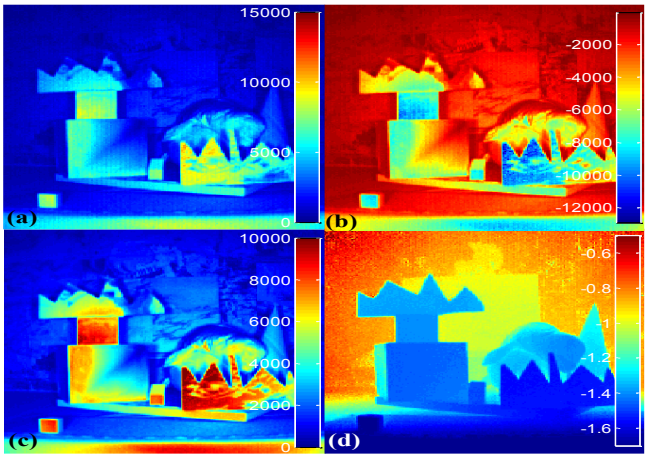

Fig. 2. Sensed data maps by a ToF device in normal operating mode. a) Phase frame $P_0(x,y)$. b) Phase frame $P_1(x,y)$. c) Amplitudes $A_P(x,y)$. d) Phase delays $\varphi_P(x,y)$.

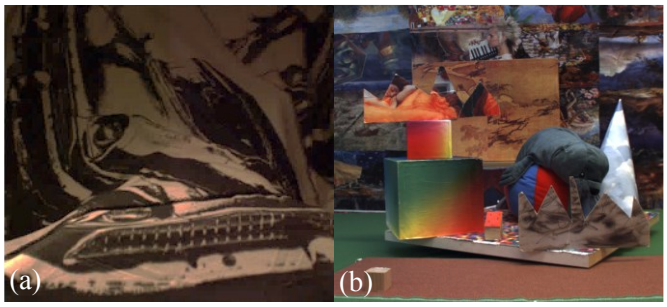

Fig. 3. Experimental scenes. a) Scene with smooth depth for FPN analysis. b) Scene with varying reflectance and depths for denoising experiments.

B. *Low sensing environment and operating mode*

We define low-sensing (LS) environment [12], [15] as an operating mode of a particular ToF device that works in suboptimal conditions and therefore generates higher measurement errors than the ones defined for normal operating mode. There are a number of external (e.g. materials of low-reflectivity, low-incident angles of reflected light, strong ambient light, multi-reflectivity path) and internal (low-powered beamers, short integration times, high sensor resolution) factors that can cause LS operating mode. External factors cannot be influenced by a user or system designer, but internal ones can be intentionally introduced in the design, e.g. reducing the power of the light beamers for power efficiency reasons when using ToF device on / with mobile devices.

For a typical ToF device, we consider an LS environment, when the measurement error E_D exceeds the maximum error E_{ToF} observed in normal conditions:

$$E_D = |D_P - D| > E_{ToF} \tag{6}$$

with D_P and D being the measured (observed) and real distance, respectively. For example, in the case of the ToF device under consideration, *CamCube* 2.0 by PMDTec™, the maximum error for the normal operational mode is given as $E_{ToF} = 0.05[m]$ [27]. It has been observed that for cases when the observed amplitude becomes $A_P < 300$ [*ToF system units*] then the error becomes larger then E_{ToF}. Consequently, this value of A_P is considered as an upper limit defining the LS operating mode for this particular device.

An example for a device operating in LS environment is given in Fig. 4. Here, the varying integration time is the factor causing an LS environment. As seen in the figure, the lower the integration time is, the lower the computed amplitude is. As a consequence more noise is present in the phase map. Although the noise is quite extreme, as it will be shown later, it can be effectively filtered by specific denoising techniques thus making it feasible for the ToF device to operate robustly in LS conditions. In next subsection we will discuss the noise in ToF devices in more detail as basis for developing suitable denoising approaches.

III. Noise Model for Time of Flight Devices Operating in Low-Sensing Environment

The sensing mechanism in ToF devices is based mainly on Complimentary Metal-oxide Semiconductors (CMOS) or Charge-coupled Device (CCD) sensors [28]. There are many internal causes, due to the sensor or associated electronic circuits, that influence the output of the sensors as discussed in [8], [24]. All those manifest themselves as noise added to the true signal. Among various sources of noise, for practical reasons we identify and model two dominant noise types that contaminate the data sensed by a ToF device. Those can be modeled as additive Gaussian noise and fixed-pattern noise. They are discussed in the following sections.

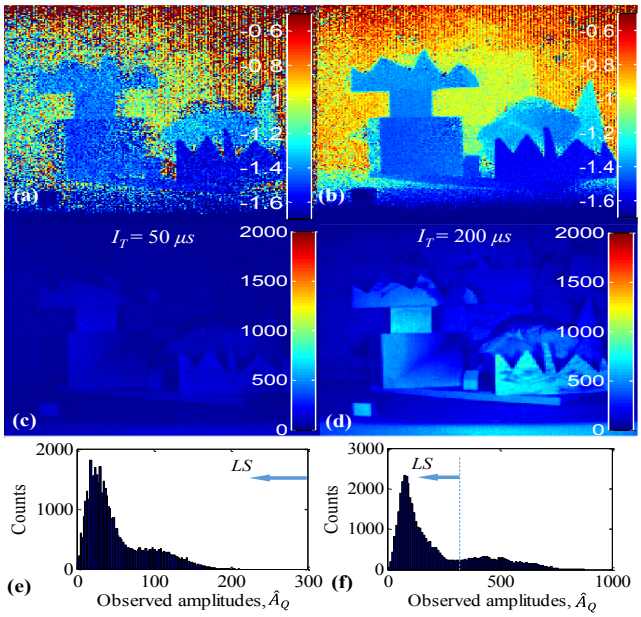

Fig. 4. Time-of-flight device in low sensing environment for integration times $I_T = 50\mu s$ (left column) and $I_T = 200\mu s$ (right column). Upper row: Phase delay maps $\varphi_P(x,y)$; middle row: Amplitude maps $A_P(x,y)$; lower row: Histogram plots of amplitude maps.

A. *Additive Gaussian noise*

First type of noise present in a ToF system is introduced by the photonic nature of emitted and reflected light and its dissipation in the medium and is related to the light conditions of the emitted and sensed signals. This noise is common in all imaging sensors (e.g. thermal cameras [17]) and is typically considered as a random process with Gaussian distribution.

Specifically for the case of LS environment, it dominates over the photonic-shot noise, as discussed in [8]. Thus, it is a satisfactory assumption for denoising of ToF data. Mathematically, the influence of this noise can be expressed as

$$\tilde{Q}_k(x,y) = Q_k(x,y) + \eta(x,y) \tag{7}$$

where $Q_k(x,y)$ is the value of the true signal (phase frames), $\eta(x,y)$ is Gaussian noise with zero mean and variance σ^2, and $\tilde{Q}_k(x,y)$ is the true signal contaminated with Gaussian noise.

Due to the assumption of Gaussian noise model, it is expected that in an ideal ToF device, the noise variance of each pixel is close to constant, that is,

$$Var(\tilde{Q}_0(x,y)) \cong \cdots \cong Var\left(\tilde{Q}_3(x,y)\right) \cong \tilde{\sigma}^2_{\tilde{Q}_k} \cong \sigma^2. \tag{8}$$

This assumption has been verified by empirical estimation of covariance matrices from multiple observations of phase frames [9].

Furthermore, assuming that $\tilde{\sigma}^2_{\tilde{Q}_k}$ is the variance of each phase frame, it turns out that the amplitude variance $\sigma^2_{\tilde{A}}$ and phase variance $\sigma^2_{\tilde{\varphi}}$ of the estimated correlation signal relate as [9]:

$$\begin{aligned} \tilde{\sigma}^2_{\tilde{A}} &= \tilde{\sigma}^2_{\tilde{Q}_k}/2 \\ \tilde{\sigma}^2_{\tilde{\varphi}} &= \tilde{\sigma}^2_{\tilde{Q}_k}/(2\tilde{A}^2_{\tilde{\boldsymbol{Q}}}) \,. \end{aligned} \tag{9}$$

The second relation is especially important as it shows that amplitude can be utilized as a confidence measure of expected measurement uncertainty of phase and consequently range. As such, it provides a confidence map, which can drive ToF denoising techniques. The fact that measurements and calculated parameters exhibit close-enough Gaussian behavior enables one to utilize any modern denoising algorithm for its removal. Examples are demonstrated and discussed in Section V.

B. Fixed-pattern noise

Second type of noise is related more to the hardware implementation of the overall ToF device, including the sensor, signal generator, LED, correlator, amplifiers, etc. This noise is present in every digital sensor and is referred to as fixed pattern noise (FPN) [17], [18], [24]. The source of this noise can be explained by the structural non-uniformities of sensing elements, that is, capabilities of sensing elements to integrate different signal intensities. Such variations have fixed behavior, but are randomly distributed among the pixels. Therefore, this effect has been named FPN. Furthermore the sensor elements are read column-wise or row-wise. Therefore, FPN typically appears as "column-" or "row-wise" stripes of slightly brighter and darker intensity across the acquired data.

The FPN is clearly visible both for raw data provided by the measured phase frames $\boldsymbol{P}(x,y)$ as well as amplitude $A_{\boldsymbol{P}}(x,y)$ and phase $\varphi_{\boldsymbol{P}}(x,y)$ evaluated from those measurements (c. f. Fig. 4 a,b and Fig. 6a). Furthermore, FPN appears visually stronger in LS environment and is negligible in normal operating mode (c. f. Fig. 2). In normal operating mode, the influence of FPN appears usually below the resolution of the device and can be excluded from consideration. In LS environment, the noise influence becomes important, since its levels start to be comparable to the sensed charges – FPN usually appears dominant over resulted sensed signal that holds true, but noisy information.

FPN is modeled as a signal-dependent noise for each phase frame

$$\begin{aligned} P_k(x,y) &= \alpha(x,y)\tilde{Q}_k(x,y) + \beta_k(x,y) \\ \alpha(x,y) &= f_{x,y}\big(\tilde{Q}_k(x,y)\big), \end{aligned} \tag{10}$$

with $\alpha(x,y)$ and $\beta_k(x,y)$ being the FPN gain and offset component, respectively, $P_k(x,y)$ the observed (measured) signal and $\tilde{Q}_k$ the true signal plus Gaussian noise mixture. The offset component is also sometimes referred to as the "pedestal maps".

Two points regarding the FPN model have to be emphasized. First, the gain component $\alpha(x,y)$ is signal dependent. As it will be seen later, it takes values close to one for normal operational mode and gets values significantly deviating from one in LS mode. Second, the same gain component is used for all phase frames ($k = 0,1,2,3$), however each phase frame k has a different offset $\beta_k(x,y)$. This is significantly different from, for example, thermal camera imaging where only one offset per sensor element is used – here it was necessary to have different offsets due to the fact that different hardware might be involved for capturing different phase frames.

Removing the FPN is a necessity for enhancing the operating performance of a ToF device in LS environment. While FPN could be effectively tackled by various hardware solutions [24], those are expensive and difficult to implement – especially for high-frequency sampling devices such as ToF. As an alternative, FPN can be corrected digitally as post-capturing operation thereby improving the signal-to-noise ratio of the sensed signal [8].

The FPN is cast as 'fixed', since the offset $\beta_k(x,y)$ is spatially fixed to sensed intensities and always appears as initial parasitic value of the sensor element added to the phase frame even when no data is sensed. The gain component map $\alpha(x,y)$ also remains fixed in terms that it always amplifies the captured intensity of certain magnitude for a certain pixel in the same functional manner. However, although termed as 'fixed', the effect of FPN is unique for each sensor device, and tends to change during exploitation time due to sensor wearing out. This means that certain easy-to-implement FPN estimation and filtering mechanism should be developed that adapts to possible changes in the nature of FPN over time.

In this paper we propose a two-stage technique for removing the FPN. First stage is responsible for removing the offset $\beta_k(x,y)$, and second one tackles the gain component $\alpha(x,y)$. They are discussed in the following two sections.

C. Removing the FPN offset component

The offset components $\beta_k(x,y)$ are generally considered as fixed masks. For estimating the offset, we occlude the optical system (if possible then the camera beamers should be also switched off in order to avoid thermal noise) and gather observation data for a large number of frames (e. g. $L = 400$).

We denote those observations $P_{k,l}(x,y)$ for $l = 0,1,\dots L-1$ and $k = 0,1,2,3$. Since we are effectively sensing 'nothing', in an ideal case each pixel would record only random noise. Since this is not the case due to the FPN offset, the idea is to determine, for each pixel, such parameters $\beta_k(x,y)$ that when those are extracted for corresponding measured phase frame, $\hat{Q}_{k,l}(x,y) = P_{k,l}(x,y) - \beta_k^l(x,y)$, for each pixel the estimated amplitude and phase would have close-to-zero mean and minimum variance for amplitude ($\hat{\mu}_{\hat{A}} \to 0$, $\hat{\sigma}_{\hat{A}}^2 \to 0$) and at the same time minimum mean and maximum variance for phase ($\hat{\mu}_{\hat{\varphi}} \to 0, \hat{\sigma}_{\hat{\varphi}}^2 \to \pi$) with mean and variance calculated per pixel over all observations. Those means and variances before and after optimization are shown in Fig. 5.

After estimating the offset values $\beta_k(x,y)$, they are stored in a table and can be subtracted in real time. They should be periodically recalculated to compensate for changes due to aging of the sensor (device). However, it is expected that those changes are not significant over time as long as the sensor does not malfunction and small changes can be mitigated by the denoising step in the proposed method. Therefore, although the offset changes with time, there is no need to measure the offset too often. This has been experimentally observed on the sensor under consideration. The offset has been measured twice in the beginning and the end of a yearlong period. While the two measurement differed, the difference was in the form of random pattern (white noise) and was compensated by the applied denoising.

D. Removing the FPN gain component

For estimating and removing the FPN gain $\alpha(x,y)$, we exploit the fact that visual stripes that are the result of the FPN gain can be considered as a texture [8], [24]. Such texture pattern can be isolated by utilizing frequency-domain analysis. When identifying the texture pattern and in order to minimize the influence of the scene content, we considered a smooth natural scene from a high reflectivity material and a continuous change in distance D (e.g. $1m \le D \le 5m$). Such scene provides varying values of phase maps $P_k(x,y)$ and at the same time is free of edges between scene objects situated at different distances. This avoids the appearance of high-frequency components in the frequency domain representation that are not related to FPN. The special scene used for this analysis is shown in Fig. 3a. This scene is used only for determining the FPN dominant frequency. Experiments quantifying the performance of the FPN removal filters are done on real-life scenes, which typically exhibit edges in the corresponding depth maps (c.f. Fig. 3b).

For a captured phase frame map $P_{k,l}(x,y)$ we remove the offset components $\beta_k(x,y)$, to get $P_{k,l}(x,y) - \beta_k(x,y)$, and calculate the Fourier transform of the resulting phase frame. The phase frame is shown in Fig. 6a and the Fourier domain representation in Fig. 7a. As seen in Fig. 7a, for the sensor under consideration, two frequency components related to the FPN gain component can be easily identified, namely, $(\omega_X, \omega_Y) = (2/3\pi, 0)$ and $(\omega_X, \omega_Y) = (\pi, 0)$. For other sensors, these frequencies might be different [8]. However, they can be determined by the same frequency domain analysis as described above.

A finite number of dominant frequency components can be easily removed by a properly designed linear-phase FIR filter. The filter design procedure takes the estimated dominant frequencies as input parameters. In the case under consideration, the undesired components are along the x-axis and one can use 1D filters that are applied horizontally (row by row) on the signal $P_{k,l}(x,y) - \beta_k(x,y)$. A good filter must satisfy the following requirements. First, is should remove the undesired components. Second, it should preserve the signal (spectra of the signal) as well as possible, that is, the passband ripple should be as small as possible. Third, it should not introduce ringing artifacts – in real scenes this will be taken care by adaptive filtering. For now it is important to point out that for our test scene this is not an issue since the scene does not have edges, while ringing is related to processing edges.

A method for designing FIR filters satisfying aforementioned requirements is discussed in Section IV. Here, we applied one of the designed filters ($N = 20$ and $\rho = 0.14$) on our test data (row by row). We assume that after filtering, we obtain $\tilde{Q}_{k,l}(x,y)$ which is the true value of the measured phase frame contaminated with Gaussian noise (c.f. Fig. 6d). This is confirmed by the Fourier domain representation of the filtered phase frame that is shown in Fig. 7b. As seen, the undesired frequency components responsible for FPN pattern are effectively removed without any noticeable modifications in the rest of the spectra.

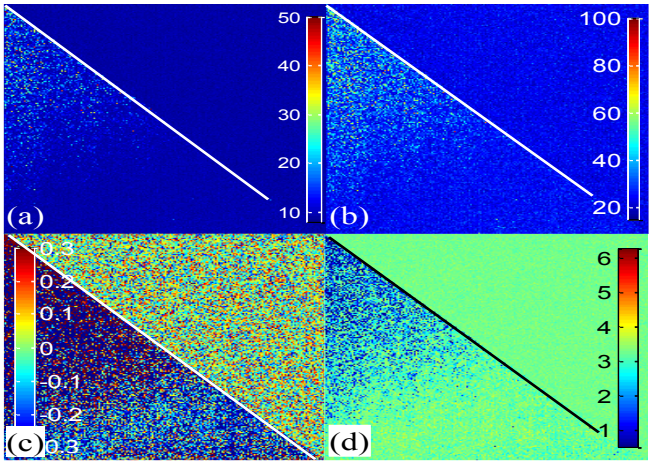

Fig. 5. Removing FPN offset $\beta_k(x,y)$ for an occluded ToF system – input data (lower-left parts) and masked data (upper right part): a) $\hat{\mu}_{\hat{A}}$, b) $\hat{\sigma}_{\hat{A}}^2$, c) $\hat{\mu}_{\hat{\varphi}}$, and d) $\hat{\sigma}_{\hat{\varphi}}^2$.

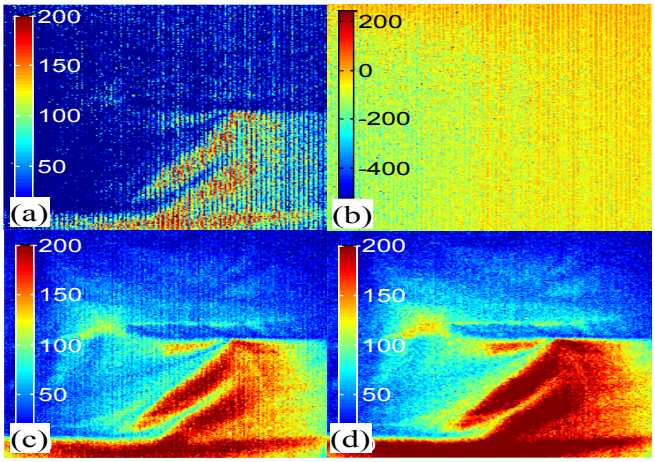

Fig. 6. Example of FPN removal. a) Measured phase frame $P_0(x,y)$. b) FPN offset $\beta_0(x,y)$. c) $P_0(x,y)-\beta_0(x,y)$. d) Filtered output $\tilde{Q}_0(x,y)$.

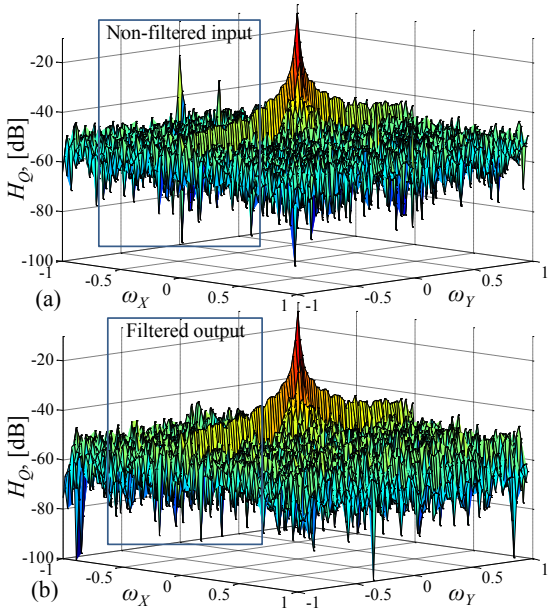

Fig. 7. Filtering of FPN amplitude gain component $\alpha^{U}_{\tilde{Q}(x,y)}$: a) Input frequency response H_Q. b) Filtered output for FIR filter of order 20 with ρ =0.28.

Assuming that filtering successfully removed the gain component of FPN, we can use the filtered values of the phase map $\tilde{Q}_{k,l}(x,y)$ to identify the behavior of the gain component. According to (10), the estimated FPN gain component is:

$$\hat{\alpha}_{k,l}(x,y) = \big(P_{k,l}(x,y) - \beta_k(x,y)\big)/\tilde{Q}_{k,l}(x,y) \qquad (11)$$

In order to cope with the Gaussian noise still present in $\tilde{Q}_{k,l}(x,y)$, we take multiple observations of $P_{k,l}(x,y)$ for several integration times in the range $50\mu s \le I_T \le 2000\mu s$ and estimate corresponding $\hat{\alpha}_{k,l}(x,y)$ values. We plot all obtained $\hat{\alpha}_{k,l}(x,y)$ against $\tilde{Q}_{k,l}(x,y)$ for several pixels (x,y) and fit curves for the selected pixels as illustrated in Fig. 8(a). As seen from the figure, the gain values converge to one for higher-magnitude measurements and deviate considerably of that value for lower values of $\boldsymbol{Q}$.

We can extend the comparison and investigate the relation between the FPN gain $\alpha_A(x,y)$ and amplitude $A_{\tilde{\boldsymbol{Q}}}(x,y)$. Applying similar procedure as explained above for $\hat{\alpha}_{k,l}(x,y)$, we can show a different FPN gain behavior at different sensing conditions quantified by different amplitudes. This is illustrated for several pixels in Fig. 8(b). As seen in the figure, the estimated gain deviates from one for low amplitudes. This can be interpreted as a high nonlinearity between the observed amplitude and the true one for the case we do not remove the FPN gain for LS environment $A_{\tilde{\boldsymbol{Q}}}(x,y) < 300$.

We aim at modeling the FPN gain component for the sensor under consideration as texture with two dominant frequencies. As seen in Fig 7a, the pattern is very well localized in frequency domain with two clear peaks and, correspondingly, has strong sinusoidal pattern in spatial domain. Given that the relation between $\alpha(x,y)$ and $\tilde{Q}_{k,l}(x,y)$ is tabulated by the curves $\hat{\alpha}_{k,l}(x,y)$ in Fig. 8(a), the model is rewritten as contribution of two sinusoidal components (patterns) in horizontal direction

$$f_{x,y}\big(\tilde{Q}(x,y)\big) = 1 + \gamma(\tilde{Q}(x,y))\,(\sin(\omega_1 x) + \sin(\omega_2 x)) \qquad (12)$$

with $\gamma\big(\tilde{Q}(x,y)\big) = \hat{\alpha}(x,y) - 1$ and $\hat{\alpha}(x,y)$ being an approximation of one of the steepest curves $\hat{\alpha}_{k,l}(x,y)$ in Fig. 8 (we use the steepest curve to simulate the worst-case scenario). Following this model, during the exploitation time of the sensor, two things can change. The frequency can drift and/or the alpha factors can vary. Since all alpha curves are similar, changing alpha factors causes change in the amplification only and not in the shape of the curve. In both cases this would result in deviating the dominant frequencies causing FPN from their estimated values. As long as the deviations are reasonably small (large changes are expected to occur only in the case the sensor is malfunctioning), they will not influence the performance for FPN gain removal with the designed filters since the filters have relatively wide stopbands and therefore will still properly remove FPN. This makes the proposed FPN removal method robust to aging of the sensor.

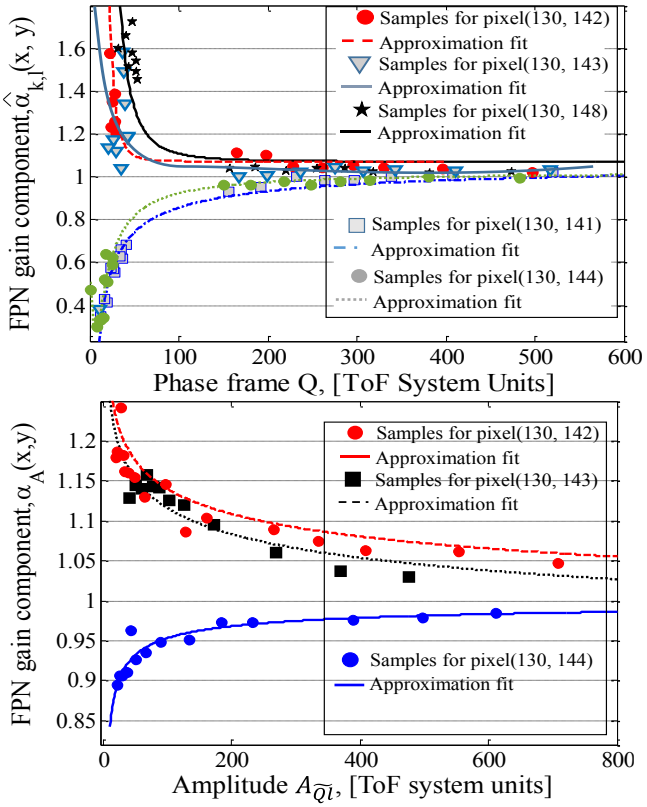

Fig. 8. Behavior of the FPN gain component $\alpha(x, y)$ for several pixels with respect to: (a) Phase frame $\tilde{Q}_{k,l}(x, y)$. (b) Amplitude $\hat{A}_{\tilde{Q}_l}(x, y)$.

E. Effect of FPN removal on ToF noise model

Based on our ToF noise model, the FPN noise removal yields a signal, whose noise component should exhibit mostly Gaussian distribution in which case the phase variance $\tilde{\sigma}_{\tilde{\varphi}}^2$ can be directly linked to variance of phase frames $\tilde{\sigma}_{\tilde{Q}_k}^2$ and the amplitude estimate can be used as a measure for that variance (c.f. (9)). Furthermore, it is expected that the phase variance (and in turn the error of the estimated range) will reduce, that is, the LS mode should perform closer to the normal operational mode. We illustrate this in Fig. 9 that shows the relation between amplitude A and phase variance σ_{φ}^2 before and after FPN removal. The data related to measurements $P_{k,l}(x, y)$ is shown in green and data after FPN removal (filtering) $\tilde{Q}_{k,l}(x, y)$ in blue. For easier comparison, we fitted curves according to Eq.(9) over observed data (please note that the x-axis is in log-scale). As expected, after removing the FPN, the phase delay variance becomes considerably smaller. Extremely small amplitudes and corresponding high phase variances are gone, the remaining variances are tolerable even for low amplitudes $A_Q < 20$ [ToF system units]. For high-enough amplitudes $A_Q > 100$ the two curves coincide. This suggests that in normal operating mode the FPN is not an issue, however, it should be specifically tackled in LS mode in order to minimize the expected phase measurement uncertainty.

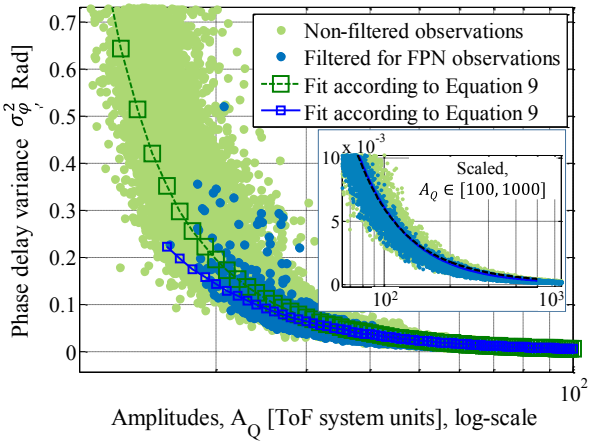

Fig. 9. Experimental observation of phase delay variance relation to signal amplitudes A_Q for FPN filtered data

IV. Design of Filters for FPN Removal

As shown in the previous section, for removing the FPN gain, we need custom-made notch-type filters. In this section we discuss methods for efficient design of such filters. Their cut-off frequencies are sensor dependent and have to be estimated before the actual design procedure takes place. However, the same general filter design principle described here can, be applied for other sensors since the frequencies relates to FPN are input parameters to the filter design algorithm.

A. Filter specifications

Based on the discussion in Section III.D, a 1D filter for removing the FPN gain for the sensor under consideration should remove (or sufficiently suppress) two dominant frequencies, namely, $\omega_1 = 2/3\pi$ and $\omega_2 = \pi$ (c.f. Fig. 7a), and at the same time preserve all other information in the signal. Ideal frequency domain representation of a filter satisfying these requirements is given in Fig. 10a and can be expressed as:

$$\left|D(e^{j\omega})\right| = \begin{cases} 1 & \text{for} \quad \omega \in X_1 \\ 0 & \text{for} \quad \omega \in X_2 \end{cases} \tag{13}$$

With regions X_1 and X_2 being defined as

$$\begin{aligned} X_1 &= [0, \omega_{p1}] \cup [\omega_{p2}, \omega_{p3}] \\ X_2 &= [2/3\,\pi; \quad \pi]. \end{aligned} \tag{14}$$

In the design, the filter transition bands (around $2/3\,\pi$ and π) should be kept as short as possible, but at the same time they have to be wide enough in order to be able to design filters with small passband ripples and reasonable orders. A method for designing linear-phase FIR filters that can achieve a good approximation of the desired response is given in the next section.

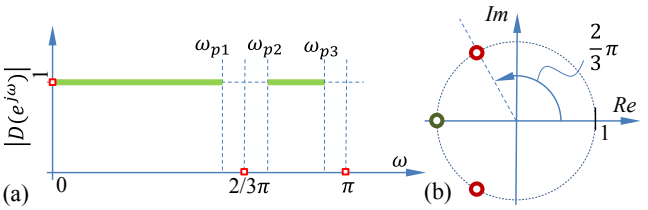

Fig. 10. Filter requirements. (a) Frequency domain specifications. (b) Desired position of fixed zeros.

B. Filter design approach

Based on the requirements listed in the previous section, the design problem can be specified as: Determine the unknown filter coefficients $h[n]$ of a filter of order N with transfer function $H(z) = \sum_{n=0}^{N} h[n]z^{-n}$ to

$$\text{minimize} \quad \delta \tag{15}$$

such that

$$\begin{aligned} 1 - \delta \le \left|H(e^{j\omega})\right| \le 1 + \delta \ \text{ for } \ \omega \in X_1 \\ \left|H(e^{j\omega})\right| = 0 \ \text{ for } \ \omega \in X_2. \end{aligned} \tag{16}$$

To simplify the design, the desired filter transfer function $H(z)$ can be split into fixed and variable part as:

$$H(z) = F(z)G(z) = F(z)\sum_{n=0}^{N_G} g[n]z^{-n} \tag{17}$$

with

$$F(z) = 1 + 2z^{-1} + 2z^{-2} + z^{-3}, N_G = N - 3. \tag{18}$$

The fixed part represented through transfer function $F(z)$ corresponds to the fixed zeros (c.f. Fig. 10b) and $g[n]$ for $n = 0,1,\dots N_G$ are the unknown filter coefficients. Those filter coefficients can be calculated by utilizing the Remez multiple exchange algorithm [29]. The procedure for applying the Remez algorithm on the above design problem is discussed in detail in [20], [30], [31]. After obtaining coefficients $g[n]$ for $n = 0,1,\dots N_G$, the filter values $h[n]$ for $n = 0,1,\dots,N$ can be easily obtained by convolving $G(z)$ with $F(z)$ as given by (17).

C. Designed filters

We used the design method proposed in the previous section and designed a family of filters having various orders and passband edges. We limited our self to filters of orders less than 60 (since we are processing finite length signals, longer filters are not to useful) and selected all passband edges at equal distances from fixed zeros in the filter frequency response, that is,

$$\omega_{p1} = \left(\frac{2}{3}\right)\pi - \rho\pi; \omega_{p2} = \left(\frac{2}{3}\right)\pi + \rho\pi; \omega_{p3} = \pi - \rho\pi \tag{19}$$

with $0.01 < \rho < 1/6$. Larger ρ values do not make practical sense, since for $\rho > 1/6 \rightarrow \omega_{p2} > \omega_{p3}$.

The obtained passband behavior (ripples) for filters of various orders and transition bandwidths are shown in Fig. 11(a,b) and the impulse response and the frequency response of one of the designed filter are shown in Fig. 11(c,d). Those filters will be considered as good candidates for FPN removal. Several comments regarding the designed filters: First, as expected, filters of higher order and wider transition band have smaller passband ripple and vice versa. Second, for, for the same performance, filters with even-order must be of higher order than the ones with odd-order. This is due to the fact that an even order filter will have, due to its properties [30], two zeros at $\omega = \pi$ instead of one, thereby making the passband optimization more challenging. Third, filters of odd order introduce a half sample delay that has to be taken care of during filtering, e.g. adjusting filtered data with original one for the purpose of comparison.

Since it is not clear which of those filters will give the best results for FPN removal, in the next section we will test all of them on synthetic and measured data in order to see which one performs best.

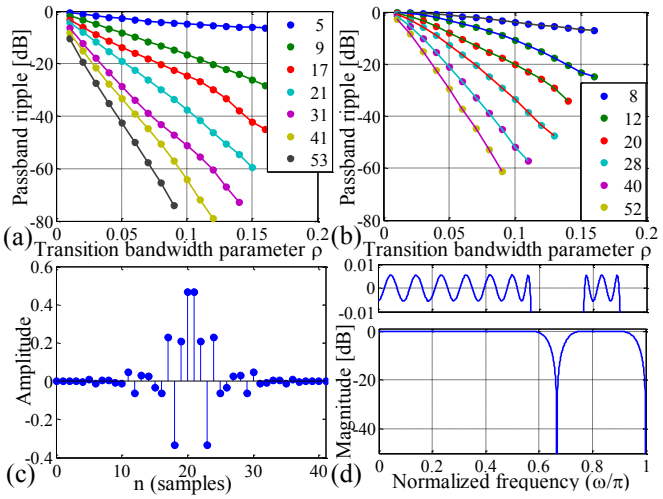

Fig. 11. Performance of designed filters. (a) Odd filter orders. (b) Even filter orders, c) & d) Impulse response and magnitude response of filter with $N = 41$ and $\rho = 0.1$.

V. Experiments

We performed two sets of experiments. In the first set we used synthetic data for evaluating the performance of designed filters for FPN removal in order to select the best filter(s). The performance of the overall FPN removal approach for estimating the depth map in LS mode is then illustrated on real data in the second set of experiments.

A. Testing performance of FIR filters for FPN removal on synthetic data

We apply the proposed filtering approach by reproducing the FPN influence in a simulated LS environment, where the focus is on the gain component $\alpha^{U}_{Q(x,y)}$ removal. We create a synthetic scene (c.f. Fig. 12), which is rich of objects of fine details, edges, and contours in different orientation and thus influencing the whole frequency response spectrum. We choose a distance sensing range of typical PMD device (e.g. PMDTec™ CamCube 2.0) $D \in [0, 7.5]m$. We denote the depth map of such scene as $D_{GT}(x, y)$. We simulate the influence of the sensor for $A_Q \in [7, 300]$ units, by applying the observed noise model as given in Section II (c.f. Fig. 9), which corresponds to the LS case. Thus, we provide a challenging testing scenario for the proposed FPN suppression approach. We apply a periodic FPN pattern that follows our observation on the same device, with modulation frequencies $\omega_{FPN} = [1, 2/3] \times \pi$ and intensity corresponding to the sensed

values of phase frames Q_k as discussed in Section III.We ensure that the added FPN is at least 25% of the signal value for very low-sensed signals and less than 5% for signals of highest values of the sensing range. Since the FPN is more dominant for signals of low intensity, we propose also an adaptation technique of the filtering process. We adapt in a manner that filtered output $\hat{D}_{\hat{Q}}$ and original data input D_P are fused in a certain normalized fraction $(\delta_A, 1-\delta_A)$ according to the sensed signal amplitudes $\hat{A}_{\hat{Q}}$:

$$\hat{D}_{adapt}(x,y) = \delta_A \hat{D}_{\hat{Q}}(x,y) + (1-\delta_A) D_p(x,y). \quad (20)$$

The adaptation helps in preserving data edges by minimizing the effect of filter-caused ringing. The adaptation parameter δ_A is select such that $\delta_A = 1$ for very low amplitudes $(A < 70)$, while for the range of amplitudes $[70, 350]$, δ_A decreases linearly between 1 and 0.2. In Fig. 14, we plot the results for proposed filters along with the effect of FPN removal on the subsequent Gaussian noise suppression applied using the Non-local Means (NLM) [21]. In Fig. 13 we plot visual output of filtered content. Table I shows the performance for some FPN removal filters quantified in terms of mean absolute error (MAE), mean square error (MSE) and peak signal-to-noise ratio (PSNR) between the ground true signal and the NLM-denoised signal. The MAE, MSE and PSNR are calculated by using following expressions:

$$MAE = \frac{1}{XY}\sum_{x=0}^{X-1}\sum_{y=0}^{Y-1}\left|\hat{D}(x,y) - D_{GT}(x,y)\right|$$
$$MSE = \frac{1}{XY}\sum_{x=0}^{X-1}\sum_{y=0}^{Y-1}\left(\hat{D}(x,y) - D_{GT}(x,y)\right)^2 \quad (21)$$
$$PSNR = 20\log_{10}\left(\frac{D_{MAX}}{\sqrt{MSE}}\right)$$

with Y and X being the number of rows and columns, respectively and $D_{MAX} = 7.5$[m]. The above three measures are used to quantify the difference between the processed depth map and the ground true one over the whole image. The results demonstrate that for a range of filters, the FPN is effectively suppressed and a 6-10 dB improvement is observed for the denoising performance for the filtered data content. Similar improvements can be noticed in terms of MAE and MSE.

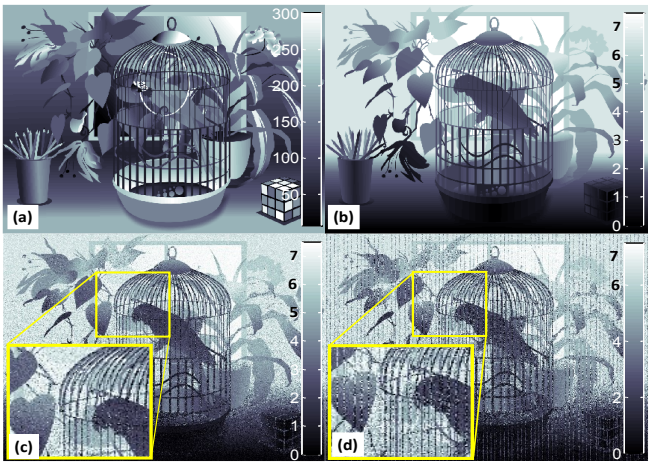

Fig. 12. Synthetic scene for low-sensing ToF sensing environment: (a) GT amplitude map, (b) GT range map, (c) noise contamination (d) FPN influence on measured depth; (zoomed region in yellow-bordered squares)

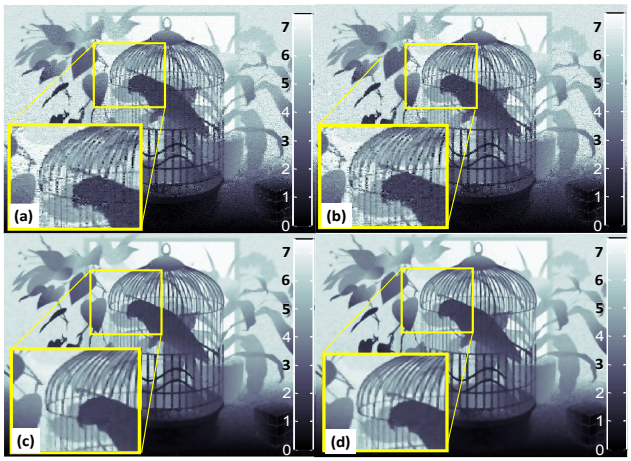

Fig. 13. FPN suppression using filter {20, 0.28} for the synthetic scene experiment: (a) filtered and (b) adaptively filtered output, and their corresponding subsequently denoised versions (c), (d).

TABLE I
DENOISING RESULTS FOR SYNTHETIC DATA

Metrics – MAE[m], *MSE*[m], **PSNR**[db]				
	With added noise [dB]		With FPN [dB]	
Input data	.40, *.35*, **27.9**		.54, *.70*, **24.9**	
NLM, *no filter*	.13, *.05*, **35.4**		.23, *.12*, **31.6**	
	FPN Filtering		FPN filtering & denoising	
{order, ρ}	F	AF	F	AF
{20, 0.28}	.45, *.44*, **27.0**	.42, *.38*, **27.8**	.16, *.06*, **34.6**	.17, *.07*, **34.8**
{28, 0.26}	.46, *.47*, **26.8**	.43, *.41*, **27.4**	.17, *.08*, **34.4**	.17, *.08*, **34.7**
{40, 0.18}	.47, *.49*, **26.6**	.44, *.42*, **27.2**	.17, *.08*, **34.2**	.18, *.08*, **34.5**
{17, 0.28}	*.47*, *.48*, **26.7**	.44, *.42*, **27.3**	.18, *.09*, **33.7**	.18, *.08*, **34.3**
{31, 0.28}	.48, *.50*, **26.5**	.45, *.43*, **27.1**	.19, *.10*, **33.4**	.18, *.08*, **34.0**
{53, 0.24}	.50, *.53*, **26.2**	.46, *.45*, **27.0**	.19, *.11*, **33.1**	.18, *.09*, **33.9**

* F and AF denote denoising results for filtered and adaptively filtereddepth map estimation, respectively.

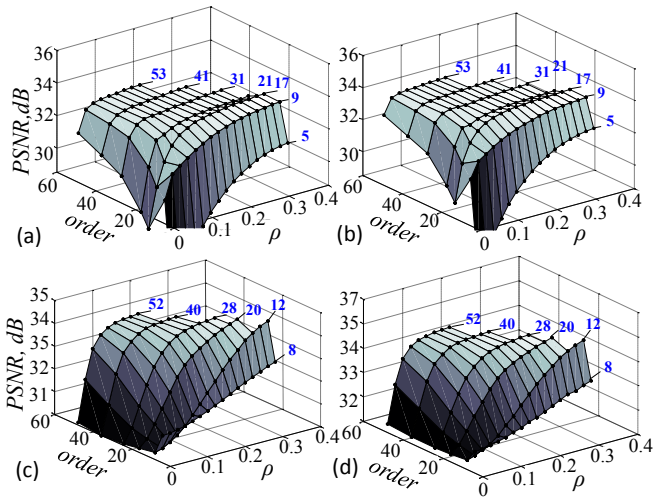

Fig. 14. Performance (PSNR) of NLM denoising method after suppression of FPN with filters of odd order (first row) and even order (second row) with (second column) and without adaptation (first column).

B. Denoising performance of FPN removal filters for real-case data

We apply the designed filters for FPN removal for a real-case experimental setup. We use a scene (c.f. Fig. 3b), composed of planar objects at different depths with respect to the sensing device. The objects are made of materials with different reflectivity and in different colors and textures. The position of sensed objects remains unchanged, while the sensing conditions are changed by varying the sensor operational parameters (c.f. Fig. 4). The default sensor parameters (e.g. integration time $I_T = 2000[\mu s]$ or higher) have been used to produce data in normal sensing conditions. In such conditions, the measured data yields high-enough amplitudes, which suggest that the uncertainty in phase estimate is lower than (or equal to) the one specified for the device. We have prepared a reference depth map, which, for consistency with the previous case, is denoted as Ground Truth (GT) depth map $D_{GT}(x, y)$. It is of size 204 x 204 pixels and obtained by averaging 100 consecutive frames, captured in normal sensing conditions. The GT depth map estimated in such way is shown in Fig. 15(a). The corresponding uncertainty map, expressed through the standard deviation $\sigma_{GT}(x, y)$ evaluated pixel-wise over the 100 observed frames, is shown in Fig. 15(b). As seen in the figure, 95% of $\sigma_{GT}(x, y)$ are below 2.11 cm. This value is in line with the expected performance of this particular sensor (c.f. Fig. 6 in [27]) confirming that our measurement setup is correct. It should be pointed out that this depth map is just a reference and it might differ from the 'true' depth map. However, it is the best estimate that can be achieved by this device while working in normal operational mode. Our goal is to quantify the performance of the proposed FPN removal method for the LS mode, against normal mode, therefore it is justifiable to use the latter for creating the GT depth map.

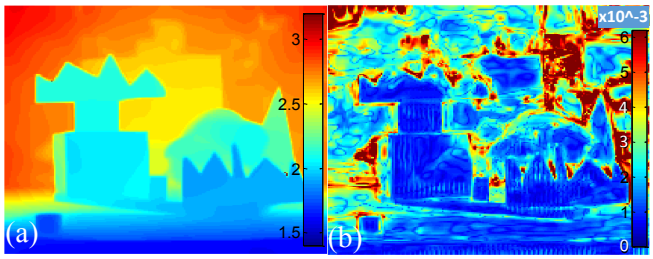

Fig. 15. Estimated data for real scene: (a) Ground truth depth map $D_{GT}(x, y)$ and (b) corresponding uncertainty map expressed through standard deviation $\sigma_{GT}(x, y)$.

The LS mode is enforced by decreasing the integration intervals to be in the range of $[50 \div 200]$ $[\mu s]$ (c.f. Fig. 4). Table II lists the percentage of pixels, where the sensed signals are with amplitudes $\widehat{\mathrm{A}}_{\boldsymbol{Q}} < 300$ ToF units, i.e. below the low-sensing threshold (denoted by LOW in Table II). As seen from the table, as well as Fig. 4(a), the scene has been assembled in such a way that for the 50 $[\mu s]$ integration time, all pixels have amplitude smaller than 300 ToF units with most of them having amplitude smaller than 200 ToF units. In this way we can test the performance of the algorithm in an extreme LS environment.

A number of techniques for suppression of additive Gaussian noise relevant to ToF data have been applied on non-filtered and filtered data. The following techniques have been applied: NLM [12], [21], Block matching over 3D (BM3D) [22], cross-bilateral filtering [32] with depth hypothesis regularizer (HypBL) [33], and Local polynomial approximation with intersection of confidence intervals rule (LPA-ICI) [23], [34]. The denoising performance is quantified in terms of MAE, MSE, and PSNR calculated according to $D_{MAX} = 7.5[m]$ following (21). The experimental results demonstrate that the denoising of ToF data measured in LS environment benefits from the proposed filtering for preliminary FPN removal. As seen in Table II, for an extreme measurement case ($I_T = 50\mu s$), the denoising performance is 5-7 dB better, in terms of PSNR, with respect to the case when FPN is not handled. Similar performance can be observed in the table also in the terms of MSE and MAE. It is important to emphasize that we utilize those three metrics since we are interested in the overall depth map and not in individual pixels. An interesting observation shows that the periodicity of FPN contributes to an artifact produced especially by non-local denoising techniques (NLM, BM3D). In these cases, the FPN is even amplified after the denoising (c.f. Fig. 17a). Fig. 17(c)-(d) illustrate the performance of denoising algorithms for a different integration time ($I_T = 200\mu s$) while FPN is present or removed. Obviously, the role of FPN is less prominent when the integration time is increased, that in turn, results with less pixels having amplitude in the LS range. This is illustrated in Fig. 16. The same three metrics are evaluated for different integration times. The figures show two things. First they demonstrate the effectiveness of the proposed FPN denoising approach – the estimated depth map benefits considerably from the proposed approach for most of the cases shown – exception are long integration times ($I_T > 1000\mu s$) for which the amplitudes are high and as such the influence of

FPN is negligible (effectively FPN is not an important factor). Second, it illustrates a compromise between the integration time (that is how much power we are ready to spend) and the achieved quality of the depth map – longer time means better depth map. This enables a user to make a proper choice based on the requirements of the application at hand.

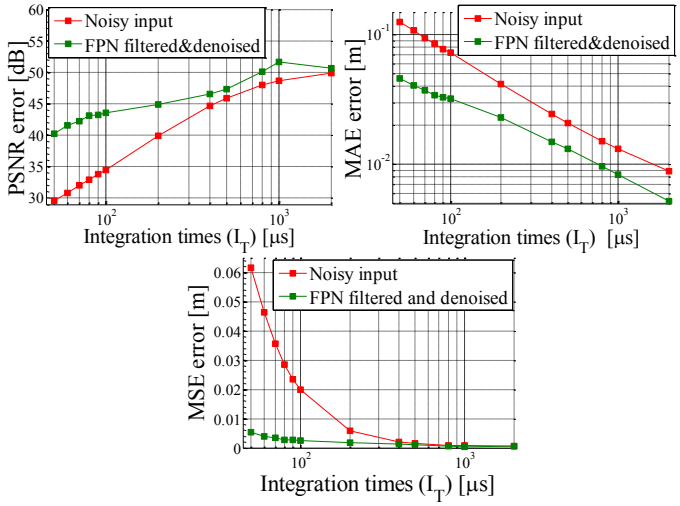

Fig. 16. Performance of proposed approach for different integration times: (a) PSNR, (b) MAE, and (c) MSE.

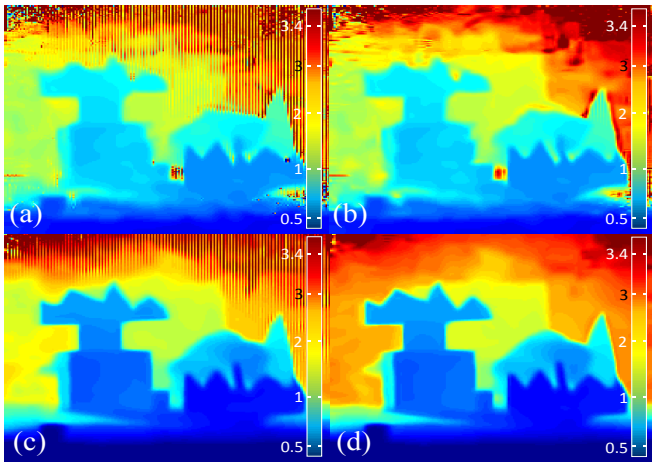

Fig. 17. Estimated depth map for a noisy input for $I_T = 50\mu s$ (upper row) and $I_T = 200\mu s$ (lower row) using BM3D. Denoising with no preliminary FPN removal (left column) and (b) denoising with preliminary FPN removal (right column).

TABLE II
DENOISING RESULTS FOR REAL-CASE DATA

Metrics – MAE[m], *MSE*[*m*], **PSNR**[db]				
Integration time I_T	$50\mu s$		$200\mu s$	
LOW	100%		81%	
Method	NF	AF	NF	AF
No denoising	.41, *.49*, **20.5**	.38, *.44*, **21.1**	.12, *.05*, **30.3**	.09, .02, **33.5**
NLM[21]	.25, *.22*, **24.1**	.16, *.08*, **28.3**	.07, *.01*, **35.8**	.04, *.00*, **40.2**
HypBL [33]	.24, *.10*, **27.1**	.16, *.05*, **30.0**	.07, *.01*, **35.7**	.05, *.00*, **38.6**
LPA-ICI [23]	.20, *.12*, **26.5**	.16, *.06*, **29.2**	.07, *.01*, **36.6**	.05, *.00*, **39.9**
BM3D [22]	.25, *.19*, **24.5**	.16, *.07*, **28.9**	.07, *.02*, **35.1**	.04, *.00*, **40.6**

* NF and AF denote denoising results for not filtered and FPN adaptively filtered depth map estimation, respectively.

In order to quantify further the quality of the estimated depth map, we repeated the measurement, for a selected integration time, 100 times thereby obtaining 100 observations of the depth maps captured under identical conditions. On those observations, we applied the proposed FPN removal approach obtaining 100 depth map estimates, $\widehat{D}_k$ for $k = 1,2,\dots 100$. For each depth map estimate, we calculated the difference between the map and the GT depth map. Finally we computed pixel-wise average and standard deviation of those differences as

$$\mu(x,y) = \frac{1}{N}\sum_{k=1}^{N} \widehat{D}_k(x,y) - D_{GT}(x,y)$$
$$\sigma(x,y) = \sqrt{\frac{1}{N}\sum_{k=1}^{N}\left(\widehat{D}_k(x,y) - D_{GT}(x,y) - \mu_k(x,y)\right)^2} \quad (22)$$

with N 100. In this way, the map $\mu(x,y)$ represents the accuracy of the depth map estimation in the LS mode in comparison to the normal operational mode of the device and the map $\sigma(x,y)$ represents the uncertainty of the estimated depth map. $\mu(x,y)$ and $\sigma(x,y)$ maps for $I_T = 50\mu s$ and $I_T = 200\mu s$ are shown in Fig. 18. In order to emphasize regions with high accuracy and uncertainty deviation from the GT, we limited the colorbars in the figures to encompass 95% of the values. The figure shows that the high deviations, as expected, are mostly on the edges of the object and some background regions but overall the accuracy and uncertainty of the depth maps are reasonable. We repeated the experiment for various integration times and compared the estimated deviations from GT and uncertainty with and without proposed FPN removal (c.f. Fig. 19). We can see that the accuracy increases (with or without FPN removal) as the integration time increases. Nevertheless, it is always beneficial to use the proposed approach since it always gives better results than unprocessed data.

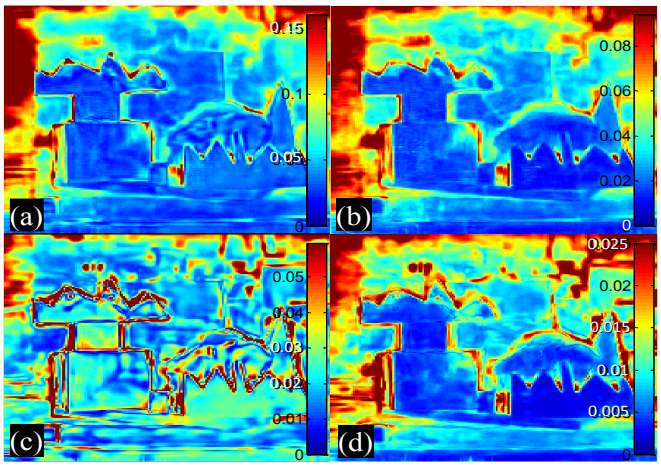

Fig. 18. Estimates of depth map accuracy μ (left column) and uncertainty σ (right column) for integration times $I_T = 50\mu s$ (upper row) and $I_T = 200\mu s$ (lower row).

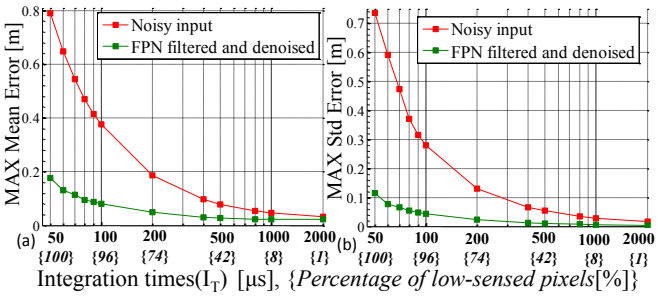

Fig. 19. Performance of the proposed approach for different integration times – c.f. Eq. (22): (a) accuracy expressed through μ and (b) uncertainty expressed through σ.

Finally, we analyzed the repeatability of the proposed method for estimating the depth map. Since, as already mentioned, we are interested in the quality of the depth map, we use the MAE as the evaluation criteria. We calculate MAE for each of the 100 observations of the scene (for $I_T = 50\mu s$). The results are shown in Fig. 20. As seen in the figure, there are very small variations between independent estimation of the depth map – standard deviation of the observation being $\sigma = 0.0066$ with $\mu = 0.0648$.

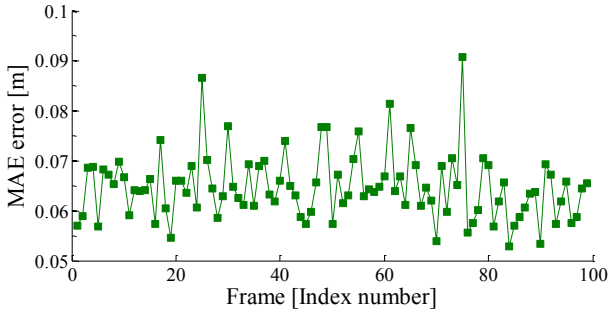

Fig. 20. Variation of MAE for the real scene ($I_T = 50\mu s$) for 100 successive observations.

VI. Conclusions

We have addressed the operation of ToF sensors in what we call LS mode. In such mode, the phase measurement uncertainty exceeds the limits specified for the particular device. Factors causing erroneous measurements can be external (scene of low-reflectivity, low-incident light-reflected angles or multi-reflectivity paths) as well as internal (weak emitted signal, short integration times) and they impose the mandatory use of denoising techniques at a post-processing stage. For proper use of such techniques, one needs an adequate noise model. We have proposed a noise model composed of two components: an additive Gaussian noise component and a signal-dependent component cast as FPN. The latter is subsequently modeled as signal-dependent gain component and additive (offset) component. FPN becomes especially visible in the LS mode and has to be handled before any additive Gaussian noise removal takes place. Based on the proposed FPN model, we developed a corresponding two-stage noise removal procedure involving 1) estimation and compensation of the offset component and 2) estimation and filtering of periodic pattern (gain) component.

Purposely-designed linear notch filters tackle the latter. We have validated our noise model by empirical measurements showing dependence between the gain component and the correlation signal amplitude, where the latter is considered a measure of varying sensing conditions. The empirical relation illustrates the importance of tackling FPN first as well as the adequacy of our noise model and the corresponding filtering method. Furthermore, we demonstrated the influence of the FPN and its removal on the widely-adopted relation between the phase variance and amplitude. While this dependence is highly violated in the LS case and in the strong presence of FPN, it gets close to the model when FPN is removed. This observation is further strengthened by experiments involving techniques for denoising signals contaminated by additive Gaussian noise applied either on non-filtered or filtered signals. The denoising results are much better while FPN has been removed before denoising.

VII. References

[1] R. Lange and P. Seitz, "Solid State ToF Range Camera," *IEEE Trans. of Quantum Electronics,* vol. 37, no. 3, pp. 297-390, March 2001.

[2] M. Hansard, S. Lee, O. Choi and R. Horaud, "Time-of-fligth Cameras: Principles, Methods and Applications," *Springer Press, ISBN-10: 1447146573,* p. 96, December 2012.

[3] "Kinect for windows," Microsoft Corporation, [Online]. Available: https://www.microsoft.com/en-us/kinectforwindows/default.aspx. [Accessed 2014].

[4] A. McCarthy, N. Krichel, N. Gemmell, X. Ren, M. Tanner, S. Dorenbos, V. Zwiller, R. Hadfield and G. Buller, "Kilometer-range, High Resolution Depth Imaging via 1560 nm Wavelength Single-photon Detection," *Optics Express,* vol. 21, no. 7, pp. 8904-15, July 2013.

[5] B. Langmann, K. Hartmann and O. Loffeld, "Depth Camera Technology Comparison and Performance Evaluation," *IEEE Proc. of ICPRAM,* vol. 2, pp. 438-444, Algarve, Portugal, 2012.

[6] "PMD[Vision] CamCube 2.0," pmd Tech, [Online]. Available: http://www.pmdtec.com/news_media/video/camcube.php. [Accessed 2014].

[7] "3D Smart Cameras," Fotonic(TM), [Online]. Available: http://www.fotonic.com/content/Products/fotonic-products-e-series.aspx. [Accessed 2014].

[8] X. Luan, Experimental Investigation of Photonic Mixer Device and Development of TOF 3D Ranging Systems Based on PMD Technology, Doctoral Thesis: University of Siegen, Germany, 2001, p. 136.

[9] H. Rapp, Experimental and Theoretical Investigation of Correlating ToF-camera Systems, Doctoral Thesis: University of Heidelberg, Germany, 2007, p. 71.

[10] M. Frank, M. Plaue, H. Rapp, U. Koethe and B. Jaehne, "Theoretical and Experimental Error Analysis of

Continuous-wave Time-of-Flight Range Cameras," *Optical Engineering,* vol. 48, no. 1, p. 16, January, 2009.

[11] D. Chan, H. Buisman, C. Theobalt and S. Thrun, "A Noise-aware Filter for Real-time Depth Upsampling," *Proc. of Workshop on Multimedia and Multimodal Sensor Fusion Algorithms and Applications,* p. 12, Marseille, France, 2008.

[12] M. Georgiev, A. Gotchev and M. Hannuksela, "De-noising of Distance Maps Sensed by Time-of-flight Devices in Poor Sensing Environment," *Proc. of IEEE ICASSP,* vol. 1, pp. 1533-1537, Vancouver, Canada, May, 2013.

[13] M. Frank, M. Plaue and F. Hamprecht, "Denoising of Continuous-wave Time-of-flight Depth Images Using Confidence Measures," *Optical Engineering,* vol. 48, no. 7, p. 13, July 2009.

[14] M. Georgiev, A. Gotchev and M. Hannuksela, "Real-time Denoising of ToF Measurements by Spatio-temporal Non-local Mean Filtering," *Proc. of IEEE ICMEW,* p. 6, San Jose, CA, July, 2013.

[15] M. Georgiev, A. Gotchev and M. Hannuksela, "Joint De-noising and Fusion of 2D Video and Depth Map Sequences," *Proc. of IEEE ICME,* p. 4, San Jose, CA, July, 2013.

[16] A. Chuchvara, M. Georgiev and A. Gotchev, "A Speed-optimized RGB-Z Capture System with Improved Denoising Capabilities," *Proc. SPIE 9019,* p. 12, March, 2014.

[17] M. Maggioni, E. Monge and A. Foi, "Joint Removal of Random and Fixed-pattern Noise Through Spatiotemporal Video Filtering," *IEEE Trans. on Image Processing,* vol. 23, no. 10, pp. 4282-4296, October 2014.

[18] J. Pezoa and O. Medina, "Spectral Model for Fixed-pattern-noise in Infrared Focal-plane Arrays," *LNCS,* vol. 7042, pp. 55-63, November 2011.

[19] B. Langmann, "PMD Imaging," in *Wide area 2D/3D imaging*, Hannover, Germany, Vieweg+Teubner Verlag, 2013, pp. 21-47.

[20] M. Georgiev, R. Bregovic and A. Gotchev, "Fixed-pattern Noise Suppression in Low-sensing Environment of Time-of-flight Devices," *Proc of Sensors Applications Symposium,* p. 6, Zadar, Croatia, April, 2015.

[21] A. Buades and J. Morel, "A Non-local Algorithm for Image Denoising," *Proc. of CVPR,* vol. 2, pp. 60-65, San Diego, CA, June, 2005.

[22] K. Dabov, A. Foi, V. Katkovnik and K. Egiazarian, "Image Denoising by Sparse 3-D Transform-Domain Collaborative Filtering," *IEEE Trans. on Image Processing,* vol. 16, no. 8, pp. 2080-2095, August 2007.

[23] V. Katkovnik, A. Foi, K. Egiazarian and J. Astola, "Directional Varying Scale Approximations for Anisotropic Signal Processing," *Proc. of EUSIPCO,* pp. 101-104, Vienna, Austria, September, 2004.

[24] M. Loose, A Self-calibrating CMOS Image Sensor with Logarithmic Response, Doctoral Thesis: University of Heidelberg, Germany, October, 1999, p. 208.

[25] A. P. Jongenelen, D. G. Bailey, A. D. Payne, A. A. Dorringtom and D. A. Carnegie, "Analysis of errors in ToF range imaging with dual-frequency modulation," *IEEE Transactions on Instrumentation and Measurement,* vol. 60, no. 5, pp. 1861-1868, 2011.

[26] S. Hussmann, F. Knoll and T. Edeler, "Modulation method including noise model fio minimizing the wiggling error in TOF cameras," *IEEE Transactions on Instrumentation and Measurement,* vol. 63, no. 5, pp. 1127-1136, 2014.

[27] H. Rapp, M. Frank, F. Hamprecht and B. Jahne, "A theoretical and experimental investigation of the systematic errors and statistical uncertainties of Time-Of-Flight-cameras," *International Journal of Intelligent Systems Technologies and Applications,* vol. 5, no. 3/4, pp. 402-413, 2008.

[28] B. Buttgen, T. Oggier, M. Lehmann, R. Kauffmann and F. Lustenberger, "CCD/CMOS Lock-in Pixel for Range Imaging: Challenges, Limitations, and State-of-the-art," *Proc. 1st. Range Imaging Research Day,* pp. 22-32, Zurich, Switzerland, 2005.

[29] J. McClellan, T. Parks and L. Rabiner, "A Computer Program for Designing Optimum FIR Linear Phase Digital Filters," *IEEE Trans. on Audio Electroacustics,* vol. AU, no. 21, pp. 506-526, December 1973.

[30] S. Mitra, Digital Signal Processing, New York: McGraw-Hill, 2006.

[31] T. Saramäki, "Finite Impulse Response Filter Design," in *Handbook for Digital Signal Processing*, New York, John Willey & Sons, 1993.

[32] C. Tomasi and R. Manduchi, "Bilateral Filtering for Gray and Color Images," *Proc. of ICCV,* pp. 839-846, Bombay, India, January 1998.

[33] S. Smirnov, A. Gotchev and K. Egiazarian, "Methods for Depth-map Filtering in View-Plus-Depth 3D Video Representation," *EURASIP Journal on Advances in Signal Processing,* vol. 25, p. 21, February 2012.

[34] V. Katkovnik, J. Astola and K. Egiazarian, "Phase Local Approximation(phasela) Technique for Phase Unwrap from Noisy Data," *IEEE Trans. on Image Processing,* vol. 17, no. 6, pp. 833-846, June 2008.