DE-NOISING OF DISTANCE MAPS SENSED BY TIME-OF-FLIGHT DEVICES IN POOR SENSING ENVIRONMENT

Mihail Georgiev[1], Atanas Gotchev[1], Miska Hannuksela[2]

Tampere University of Technology, Tampere, Finland[1]; Nokia Research Center, Tampere, Finland[2]

ABSTRACT

We propose a non-local de-noising approach aimed at filtering range data sensed by Photonic Mixer Device sensors. We address specifically the case of poor sensing environment when the reflected signal amplitude is low. In our approach, signal components of phase-delay and amplitude of the sensed signal are regarded as components of a complex-valued variable and processed together in a single step. This imposes better filter adaptivity and similarity weighting. The complex-domain filtering provides additional feedback in the form of improved noise-level confidence, which can be utilized in iterative de-noising schemes. Pre-filtering of individual components is proposed to suppress structural artifacts. Our approach compares favorably with state of the art approaches.

Index Terms— PMD, Time-of-Flight, de-noising, Non-Local Means, complex, noise confidence, range device

1. INTRODUCTION

A number or imaging applications, such as 3D scene capture and reconstruction, 3D video capture and virtual view synthesis require a precise knowledge about the scene depth. Therefore, sensors measuring distances, so-called range sensors as well as the corresponding sensing techniques have emerged recently. A class of range sensing devices uses the so-called *Time-of-Flight* (*ToF*) principle, where the distance is measured by computing elapsed time between emitted and reflected signals [1]. Photonic Mixer Device (PMD) is an evolving technology which utilizes a specific ToF technique of measuring phase-delay of reflected signal emitted by a continuously-modulated infrared light source [2, 3].

While being quite precise in measuring distances in indoor conditions, PMDs exhibit generic drawbacks related with their principles of operation. Some specific properties of the sensed scene, ambient light conditions or technological limitation of the used device can lead to erroneous range measurements, generally denoted as noise added to the range image. In this work, we present an approach aimed at suppressing the noise present in range images under certain imaging conditions. Specifically, we keep the sensing principle and hardware unchanged and tackle the problem by an image denoising technique implemented as a post-processing step.

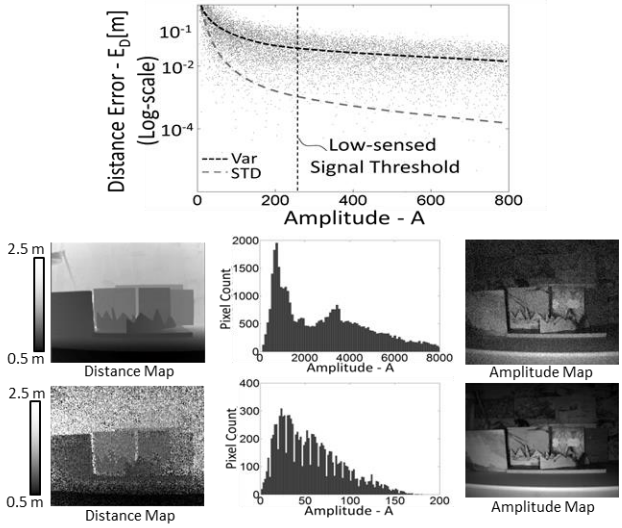

Fig. 1. Example of PMD capturing in low-sensing environment(top-to-bottom): a) relation between A and E_D, b) default capturing mode, c) low-sensing environment

2. PMD PRINCIPLE AND NOISE MODEL

A typical PMD consists of a beamer, an electronic light modulator and a sensor chip (e.g. CMOS or CCD). The beamer is made of an array of light-emitting diodes (LED) operating in near-infrared wavelengths (*e.g. 850 nm*). It radiates a point-source light of a continuously-modulated harmonic signal which illuminates the scene. The reflected light from object surfaces is sensed back by pixels of the sensor chip, which collects pixel charges for some period denoted as *integration time*. For each pixel, the range data is estimated in relation to phase-delay between sensed signal and the one of the light modulator [1, 2, 4]. The phase-delay estimation is performed as a discrete cross-correlation process of several successively captured samples taken between equal intervals during same modulation periods of fixed frequency. Denote the sample data as R_n *(n=1, 2,…N-1, N≥4*). The mixed signal components (amplitude and phase) are estimated from the sampled data as follows

$$A = \frac{2}{N}\sum_{n=0}^{N-1}\left|R_n e^{-j2\pi\frac{n}{N}}\right|, \varphi = \arg\left(\sum_{n=0}^{N-1} R_n e^{-j\frac{2\pi n}{N}}\right) \quad , \qquad (1)$$

where A is the modulation amplitude, and j is the imaginary unit. The sensed distance D is proportional to the phase

$$D \propto \frac{\varphi}{4\pi f} c_L, \quad (2)$$

where f is the frequency of the emitted signal and c_L is speed of light through dry air (~298.10^9 *km/h*). The exact value of D is calculated after precise calibration of the sensor [5]. It has been proved that the measured signal amplitude is an optimal estimator for the reliability of range measurements [4, 6]. That is expressed by a relation between the variance of measured error of the sensed D, denoted by σ^2 and the modulation amplitude

$$\sigma^2 \propto \frac{1}{A^2}, \quad (3)$$

An empirical estimation of the distance error E_D proving Eq.(3) is plotted in Fig. 1a. As seen in the figure, high measurement error is related with very low-amplitude signals which thus indicate a poor sensing environment. An amplitude threshold can delineate the operating range. Below the threshold, the distance measurement error goes above the specifications given in PMD operating manuals (cf. Fig. 1a). Erroneous sensing environment can be caused by both sensory-internal and external causes. Internal factors include low power of the emitted signal or short integration time for forming the reflected signal samples. External factors include reflection from surfaces of small incident angles, sensing objects having low-reflectivity materials or colors, objects situated close to the range sensing limit, signals formed by multiple reflections. The effect of such factors is an erroneous range image where errors are manifested as noise. Fig. 1b illustrates the case where the sensor works in normal mode. This mode is characterized by histogram of measured amplitudes ranging from 200 to 8000 units thus well above the empirical threshold of 200 units. Fig. 1c illustrates the case where all amplitudes are below the threshold. While the amplitude map looks structurally intact, the distance map is grainy (noisy). Furthermore, there might be erroneous distance measurements even when the amplitude is relatively high. Such cases are caused by e.g. objects with very high reflectivity, which become additional light sources for other objects thus causing multiple reflections. Strong ambient light containing the same wavelength as of the emitted signal (e.g. sensing outdoor scenes under sun light) is also a source of measurement errors. Such errors are manifested in the distance map as structural artifacts rather than as a noise. We are particularly interested in the cases when the PMD sensor is forced, by technological limitation, to work in low power or short integration time mode (e.g. requirements for miniaturization leading to limited beamer size, decreased number of LED elements; embedding into portable low-power devices; cost of hardware). A remedy for such cases would be to introduce a denoising procedure applied to the computed distance map in a post-measurement stage.

3. RELATION TO PRIOR WORK

Image de-noising is an active area of research. Among the myriad of denoising methods, edge-preserving bilateral filtering [7] and non-local (patch based) methods [8, 9] have shown superior results. Correspondingly, those methods have been modified for the case of PMD data denoising as well. The work [10] has addressed a case of combing PMD data with aligned 2D color camera. The de-nosing benefits from the additional color texture mode. Both sets of data are projected on a mutual grid and a cross-modality Non-Local Means filter is applied that search for filter weights in color image to de-noise distance data. Several articles have proposed de-noising methods which use the modulation amplitude A as a given prior of noise presence and adapt parameters of the de-noising filters [6, 11, 12, 13]. Techniques utilize bilateral or multi-lateral filters [11, 12] or wavelet transforms with amplitude-adaptive thresholding [13]. De-noising of PMD data in complex-valued domain is first proposed in [6]. There, the de-noising filters are based on median smoothing and normalized convolution with Gaussian smoothing kernel.

We propose an approach which contains two novel ideas. First, we employ non-local means filtering in complex-valued signal domain, which to the best of our knowledge has not been studied yet. The patch similarity search for complex-valued image patches effectively adapts the prior knowledge of noise influence to the de-noising process and simplifies the block search stage leading to better weighting for complex-valued signals with reduced computational complexity. The non-local de-noising in complex domain implies simultaneous filtering of all signal components in a single step thus it additionally improves the noise confidence parameter given by A, which can be further used effectively for an iterative de-noising scheme. Second, we propose also a preliminary filtering step of individual signal components that provides additional enhancement effect by suppressing system artifacts to the also improved de-noising output.

4. PROPOSED DE-NOISING APPROACH

In our approach, we adopt the non-local de-noising paradigm. The general idea is to find and stack similar blocks (patches) of pixel data together and utilize their similarity measures as weights in a filtering process based on averaging. The reference for similarity is chosen as a block or patch of pixels that surrounds the filtered pixel. We based our approach on the original NLM [8], however other non-local transform-domain filtering schemes are also possible [9]. The general NLM for a pixel with coordinate x is defined as follows [8]:

$$NLM[x] = \frac{1}{C_N(x)} \int_{\Omega} exp\left(-\frac{G \times |U(x+.) - U(y+.)|^2}{h^2}(0) \right) U(y) dy \quad , \quad (4)$$

where C_N –normalization factor, G – Gaussian kernel, Ω – search range, U – pixel map, x – index of filtered pixel, y –

index of center pixel in similarity patch, h – filter parameter tuned in relation with σ^2, $\times\cdot(0)$ – centered convolution operator, $(+\cdot)$ – pixel indices of spatial neighborhood. The map for similarity search, denoted by U in our approach is chosen to be the pre-computed maps of (A, φ) – (A_U, φ_U) given in Eq.(1) and combined together pixel-wisely into a complex-valued map, denoted by Z:

$$(\varphi_U, A_U) \rightarrow Z = A_U\left(e^{j\varphi_U}\right),\ Z \rightarrow A_U = |Z|,\ Z \rightarrow \varphi_U = \arg(Z)\cdot \quad (5)$$

The modified NLM filter is given by:

$$NLM_{CMPLX}[x] = \frac{1}{C_N(x)}\int_{\Omega} exp\left(-\frac{G\times|Z(x+.)-Z(y+.)|}{h^2}(0)\right)Z(y)dy\ . \quad (6)$$

Furthermore, we propose a preliminary filtering of individual components, i.e. A, φ by some mild smoothing filtering (e.g. Gaussian) before their coupling into a complex number thus leading to their filtered values A_D, φ_D. The reason for this preliminary smoothing is to tackle possible structural artifacts. The effect is illustrated in Fig. 2, where the source exhibits areas of multi reflectivity, low reflectivity and small angle of incidence. For such cases, the preliminary smoothing improves search confidence of block patches, and tackles possible wrapping effects of the phase-delay φ. The latter effect can be explained as follows: when the noise influences values of φ being close to wrapping boundaries, the phase wraps-over and provide measures of extreme errors among closely situated pixels which in fact are sensing similar distances. At the same time, the confidence given by A indicates small error. A wrapping noise effect could appear for sensed objects on far distances, or in local areas of multi-reflected sensed signals. A pre-filtering is capable of mitigating this effect. Pre-filtering of both components gives usually best results, however, for low-complexity reasons it can be simplified to either filtering the amplitude or the phase. Filtering of either components leads to different enhancement effects summarized in the following options:

Option 1. Filter A to get A_D. De-noising and edge preservation in distance map is improved, but structural artifacts remain

Option 2. Filter φ to get φ_D only. Edges are slightly preserved, but structural artifacts are better suppressed

The effects of the above options are explained as follows: when amplitudes $A(y)$ of a filtered patch are low, then similarity weights for patches of the same condition will have stronger impact and noise suppression will be weak. Some enhancement will be obtained only when the phase-delay data for those patches is modified (pre-filter φ). Preliminary smoothing of φ might blur some edges. In the other hand, when the patch being filtered has high amplitudes (indicating small measurement errors), it is preferred that the range content is to be kept rather than smoothed. Thus, smoothing only the amplitudes is expected to improve the similarity block search. An adaptively-switching pre-filtering can be devised combining benefits of both options. Pixels with very low values of A are treated by Option 2, all other cases are treated by Option 1(see the results in subsection 5.3).

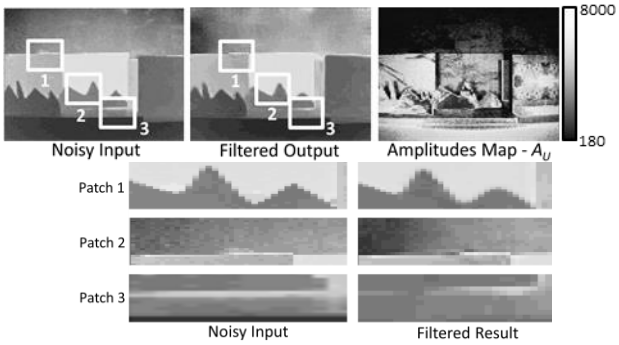

Fig. 2. Artifact Suppression by pre-filtering step: *Patch 1*) area of multi-reflectivity, *Patch 2*) area of low-reflectivity, *Patch 3*) area of surface reflection in small angles of incidence

As the de-noising of complex signal modified both signal components, the modified amplitude can be further utilized for iterative de-noising procedure. The modified amplitude serves as an improved (*de-noised*) confidence as illustrated by Fig. 3. An optimal iterative solution will require adapting the filter parameter – h in every iteration due to the changed noise variance in the phase map. The proposed iterative de-noising approach is summarized follows:

0. Initialize. Iteration counter $N=0$
   - 0.1. Capture PMD data. Compute A, φ (Eq.(1))
1. IF pre-filtering?
   - 1.1. YES – Pre-filter optionally A, $\varphi \rightarrow A_{[D]}$ or $\varphi_{[D]}$
   - 1.2. NO – Skip step
2. Calculate Z_N from $A_{[D]}$ or $\varphi_{[D]}$ (Eq.(5))
3. Apply de-noising filter on Z_N (*e.g.* Eq.(6))
   - 3.1. SAVE output – Z_N
   - 3.2. Calculate A_N, φ_N from Z_N (Eq.(5))
4. IF Iterate?
   - 4.1. YES – N=N+1. Adapt filter parameter – h.
     - 4.1.0. REPEAT Steps 1÷4
   - 4.2. NO – Calculate D from φ_N. (Eq.(2))
5. SAVE D. Finish

5. EXPERIMENTS AND RESULTS

5.1. Test Setup

Our experimental setup includes a captured scene where the relative object position to camera are kept unchanged, but the sensing conditions are varied by changing sensor operational parameters. A PMD Vision CamCube 2.0 has been used in the experiments [14]. Default camera capturing settings and integration time of *2000 μs* has been set to model normal sensing conditions (i.e. amplitude confidence ensuring measuring error smaller than specified for the device). We prepared Ground Truth (*GT*) images by averaging *200* consecutively captured frames. The testing scene is designed from planar objects facing frontally the camera, but placed in arbitrary positions. The objects are made from materials of different reflection or painted in different color textures. We model the varying sensing environment by decreasing the integration times for

capturing raw samples. The following times were experimented with *200 μs, 100 μs, 50 μs* (cf. Fig.4, Fig. 5). Table 1 gives the percentage of pixels which amplitudes are below the threshold.

5.2. Experiments

We test our de-noising approach denoted by NLM_{CMPLX} for varying pre-filtering schemes and number of iterations. We compare it with the classical NLM approach (denoted by NLM) applied to distance data only. The pre-filtering scenarios are tested for four cases of NLM smoothing filter: 1) φ_D only 2) A_D only *3*) Both A_D, φ_D, *4*) Adaptive switching between A_D, φ_D. We compare also with prior work. We have ported a Matlab code of Frank et al. [4] (denoted by (FRANK)) and use their comments how to initialize properly filter parameters. We reproduced also the algorithm of Chan et al. in [12] (denoted by CHAN) in an implementation of bi-lateral filter [7]. We implement one more solution with the adaptive approach by CHAN, but utilizing Non-Local Means filter instead, denoted by $NLM_{ADAPT.}$ The de-noising performance is measured by *PSNR*:

$$MSE = \frac{1}{N_P}\sum_{x=0}^{x<N_P}\left(D_C(x)-D_{GT}(x)\right)^2, PSNR[dB] = 20\log_{10}\left(\frac{D_{MAX}}{MSE}\right) \quad , \quad (7)$$

where N_P – number of pixels, D_C – evaluated range map, D_{GT} – *GT* range map, D_{MAX} – maximum expected distance value in PMD data (Eq.(2), for φ=2π). For fair comparison, we tune filter parameters of tested algorithms until best result is obtained for each method.

Table 1

PMD Sensor Integration Times[μs]	*200*	*100*	*50*
Low-sensed pixels[%]	*57*	*93*	*100*
Noisy Input[*dB*]	26.58	22.41	18.10
Methods	De-noised Results[*dB*]		
FRANK	30.69	28.34	26.09
CHAN	32.24	30.56	28.50
NLM(Buades)	33.11	30.95	26.98
NLM_{ADAPT}	33.95	31.06	27.16
Proposed Solutions			
NLM_{CMPLX}	35.02	31.39	27.55
NLM_{CMPLX}(*1 iteration*)	35.15	31.62	27.94
NLM_{CMPLX} (φ_D)	35.05	32.58	30.74
NLM_{CMPLX} (A_D)	35.15	31.60	28.03
NLM_{CMPLX} (φ_D, A_D)	**35.29**	32.48	**30.76**
NLM_{CMPLX} (*Selective* φ_D, A_D)	35.11	**32.60**	30.74

Bold Numbers – *Best Metric Results*

5.3. Results and Conclusions

The obtained results are given in Table 1 and some are depicted in Fig. 4, Fig. 5. For the given experimental setup, the proposed solution NLM_{CMPLX} for all tested scenarios show better performance of about 0÷2 dB in terms of PSNR, when compared to results from the other tested algorithms. The visual observation of the results suggests that obtained map bears also closer resemblance to GT map. Results of algorithms including pre-filtering show further performance boost and appear visually better (edges better preserved). The iterative approach brings additional improvement. It is worth to mention, that the selective pre-filtering approach

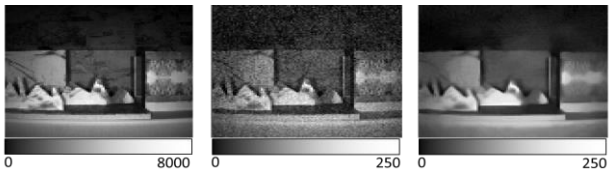

Fig. 3. De-noising example of amplitude map: a) Ground Truth data, b) noisy input(A_U), c) de-noised result(A_D)

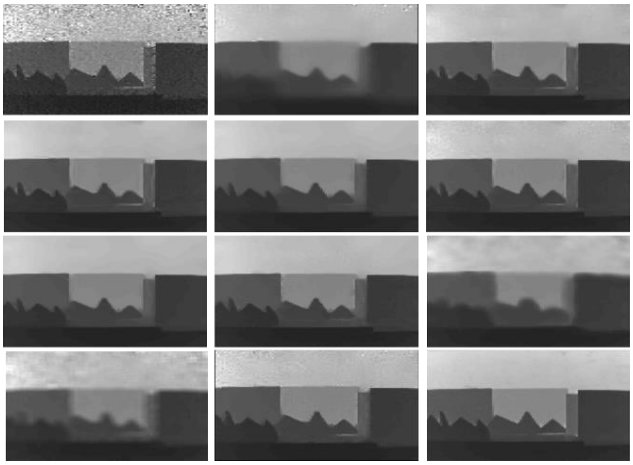

Fig. 4. De-noising Results(200 μs) (left-to-right, top-to-bottom) a)Noisy Input, b)NLM, c)NLM_{CMPLX}, d)NLM_{CMPLX}*(1iter.)*, e)$NLM_{CMPLX}(\varphi_D)$, f)$NLM_{CMPLX}(A_D)$, g)$NLM_{CMPLX}(\varphi_D, A_D)$, h)$NLM_{CMPLX}$(Selective φ_D, A_D), i)FRANK, j)CHAN, k)$NLM_{ADAPT.}$ i)GT

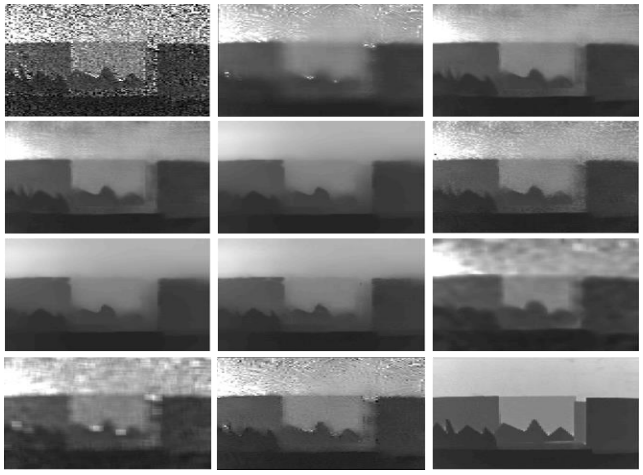

Fig. 5. De-noising Results(50 μs) (left-to-right, top-to-bottom) a)Noisy Input, b)NLM, c)NLM_{CMPLX}, d)NLM_{CMPLX}*(1iter.)*, e)$NLM_{CMPLX}(\varphi_D)$, f)$NLM_{CMPLX}(A_D)$, g)$NLM_{CMPLX}(\varphi_D, A_D)$, h)$NLM_{CMPLX}$(Selective φ_D, A_D), i)FRANK, j)CHAN, k)$NLM_{ADAPT.}$ i)GT

NLM_{CMPLX}(Selective A_D, φ_D) shows best result in terms of PSNR in one case and very competitive in the other two while being visually the best (artifacts are better suppressed). Another important conclusion is that, while PSNR metrics shows similar or very close results among some solutions, the visual result of de-noised content could vary in great extent in terms of edge preservation (from sharp to quite blurry).

REFERENCES

[1] R. Lange, P. Seitz, "Solid State Time-of-Flight Range Camera", *Journal of Quantum Electronics*, Vol. 37, No. 3, pp. 390-297, 2001.

[2] T. Spirig, P. Seitz, O. Vietze, and F. Heitger, "The Lock-in CCD – two dimensional synchronous detection of light", *Journal of Quantum Electronics*, Vol. 31, pp. 1705-1708, 1995.

[3] T. Ringbeck, PMDTechnologies GmbH, "A 3D Time of Flight Camera for Object Detection", *Conference of Optical 3D Measurement Techniques*, Zurich, Switzerland, 2007.

[4] M. Frank, M. Plaue, H. Rapp, U. Koethe, B. Jaehne, F. Hamprecht, "Theoretical and Experimental Error Analysis of Continuous-Wave Time-Of-Flight Range Cameras", *Journal of Optical Engineering*, Vol. 48, No. 1, pp. 1-24, 2009.

[5] M. Lindner, A. Kolb, "Lateral an Depth Calibration of PMD-Distance Sensors", *Proceedings of ISVC 2006*, Vol.2, pp. I: 524–533, Lake Tahoe, USA, 2006

[6] M. Frank, M. Plaue, F. Hamprecht, "Denoising of Continuous-wave Time-of-flight Depth Images Using Confidence Measures", *Journal of Optical Engineering*, Vol. 48, No. 7, SPIE, 2009

[7] C. Tomasi, R. Manduchi, "Bilateral Filtering for Gray and Color Images", *Proceedings of the IEEE International Conference on Computer Vision 1998*, Vol.1, pp.839-847, Bombay, India, 1998

[8] A. Buades, J. Morel, "A Non-local Algorithm for Image Denoising", *Proceedings of Conference of* Computer Vision and Pattern Recognition 2005, Vol. 2, pp. 60-65, San Diego, USA, 2005

[9] K. Dabov, A. Foi, V. Katkovnik, K, Egiazarian, "BM3D Image Denoising with Shape-Adaptive Principal Component Analysis", *Proceedings of SPARS 2009*,Vol. 16, No.8, pp. 2080-2095, Saint-Malo, France, 2009

[10] B. Huhle, T. Schraer, P. Jenkke, W. Strasser, "Fusion of Range and Color Images for Denoising and Resolution Enhancement with a Non-Local Filter", *Journal of Computer Vision and Image Understanding*, Vol. 114, No. 12, pp. 1336-1345, 2010.

[11] G. Jin, J. Baozhu, Z. Xudong, H. Liangmei, "PMD Camera Calibration Based on Adaptive Bilateral Filter", *Proceedings of SOPO'11,* Vol.1, pp. 1-4, Wuhan, China, 2011.

[12] D. Chan, H. Buisman, C. Theobalt, S. Thrun "A Noise-Aware Filter for Real-Time Depth Upsampling", *Workshop on Multicamera and Multimodal Sensor Fusion Algorithms and Applications (ECCV),* Vol. 1, pp. 1-12, Marseille, France, 2008.

[13] T. Edeler, K. Ohliger, S. Hussmann, A. Mertins, "Time-of-Flight Depth Image Denoising Using Prior Noise Information", *Proceedings of ICSP 2010,* pp. 119-121, Beijing, China, 2012

[14] PMDTechnologies GmbH, "PMD[Vision] CamCube 2.0", *PMD Camera Vendor*, info@pmdtec.com., Siegen, Germany, 2010