Vision, Modeling, and Visualization (2011)
Peter Eisert, Konrad Polthier, and Joachim Hornegger (Eds.)

Assessment of Time-of-Flight Cameras for Short Camera-Object Distances

Michael Stürmer[1,2,3], Guido Becker[3] and Joachim Hornegger[1,2]

[1]Pattern Recognition Lab, Department of Computer Science, Friedrich-Alexander-University Erlangen-Nürnberg, Germany
[2]Erlangen Graduate School of Advanced Optical Technologies (SAOT), Germany
[3]Lion Systems S.A., Luxembourg

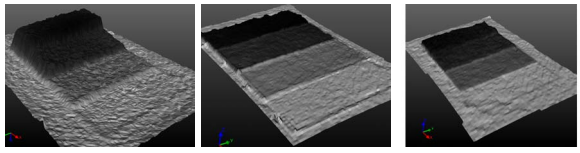
Abstract

In this paper we have compared Time-of-Flight cameras of different vendors at object-camera distances of 500 mm*,* 1500 mm *and* 2500 mm*. The aim was to find the highest possible precision at the distance of* 500 mm*, to estimate the change of the accuracy depending on scene-reflectivity and working distance and to investigate the possibility to use the cameras as per-pixel sub-centimeter accurate measuring devices. To this end, we have evaluated the variation of the distance measurement noise over several distances as well as the minimum noise we could achieve with each camera. As the amplitude-dependent distance error may become significantly large, we also tried to quantify it in order to estimate if it can be reduced to fulfill given accuracy requirements. We compared a Camcube3 from PMD Technologies, a Swissranger4000 from MESA Imaging and a C70E from Fotonic. All cameras showed different behaviors in terms of temporal noise, variation of noise and amplitude dependency. The Camcube showed the strongest amplitude dependent effects. The minimum standard deviations at* 500 mm *distance were at* 4.8 mm *for the Camcube,* 1.6 mm *for the Swissranger and* 0.9 mm *for the C70E.*

1. Introduction

Time-of-Flight (ToF) cameras offer a convenient way to acquire 3-D-data for realtime processing purposes [KBKL10]. Commercial products are available from several manufacturers, each implementing the range sensing technology in different ways [Lan00, OLK*04]. While there are different realtime-capable range sensing modalities available at comparable and lower prices, ToF offers the advantage of easily combining several sensors in one setup using different modulation frequencies of the reference signal.

For all ToF-sensors available on the market, the cornerstones of the cameras are the same: all cameras work with an active illumination unit which emits light at wavelengths of ca. 800 nm. The optical signal is modulated with frequencies of about $20 - 40$ MHz. The reflected signal is sampled on the ToF-chip and by demodulation the per-pixel phase shift with respect to the reference signal is detected. This principle is well known and published in literature [OLK*04]. As ToF cameras are often claimed to be very accurate range imagers, a high quality of the distance data is expected from these sensors in general. The better the accuracy and precision of the sensor is expected, the more hardware related effects, which impair the theoretical optimal measurement, become obvious and cannot be neglected. The most significant effects that affect the quality of the measured range data are



(a) Camcube3 (b) Swissranger4000 (c) C70E

Figure 1: Amplitude-dependent measurement deviations for the tested cameras at 500 mm. The mean error in the darkest regions are from left to right: 41.8 mm (a), 4.4 mm (b), 9.2 mm (c). The four intensity levels in each image correspond to 0%, 30%, 60% and 100% white-level (see text for explanation). For the visual representation outliers where reduced by a 3×3 median filter.

DOI: 10.2312/PE/VMV/VMV11/231-238

- a per-pixel fixed pattern noise offset of the chip
- the amplitude-dependency of the distances [LK07]
- the so-called 'wiggling' effect (which is due to imperfect shapes of the reference light signal)
- the temperature-dependency and other drift-effects of the sensor
- the dependency of the precision of the sensor on the amplitude of the received signal.

In addition, due to the high dynamic range of the received amplitude data and the limited resolution of the sensor, oversaturation and high noise due to weak signal strength may occur even within the same recorded frame. This property makes operating ToF-cameras even more complex, as the devices must be adjusted constantly in terms of integration time in order to produce reliable data. Calibration and correction procedures for most of these effects have been published in literature [LK07], yet the provided results mostly cover the whole measurement range of the cameras (5 m for 30 MHz, 7.5 m for 20 MHz modulation frequency) and are neglecting close-range applications with distances < 1 m.

As long as the sensors are only used for rough estimation of 3-D positions in space, most of the stated effects can be neglected. For detection of hand gestures or body pose estimations as input modality for interaction with a computer program a perfectly accurate surface model of the hand or the body will most likely be not of high importance. Fitting of skeleton models and estimating the motion of these is not a trivial task either, but the precision of the data does not need to be as high. If the sensors are meant to be used as 3-D acquisition devices which should deliver accurate 3-D surface models of the scene that is imaged, all the above mentioned problems must be addressed. Close-range high-accuracy applications of ToF-cameras have been proposed in the past and are still subject to research [PHS*09, SSBH11].

The details how the manufacturers solve particular problems in reference signal generation and ToF-chip based hardware related problems are undisclosed to the user of a ToF-camera. From experience one can say that different approaches are implemented by each manufacturer which exhibit different influence and side-effects on the data. Many of these effects have been addressed by the community both by theoretic explanation as well as experiments [SJ09]. Approaches for reduction of the amplitude-related distance error and the 'wiggling' have been proposed [LK07, KRI06]. Studies were mostly made using one specific camera and investigating the errors of that device in much detail. A comprehensive overview that compares a set of contemporary ToF-cameras and their properties regarding high accuracy (i.e. sub-centimeter) measurements is not available so far.

A thorough evaluation of the long-term-stability of the investigated cameras is out of the scope of this work. In the past, ToF cameras tended to show global shifts in their distance measurements of up to several centimeters. For a rigidly mounted set of pre-calibrated cameras with known extrinsic positions this requires the system to be re-calibrated very often. Nevertheless, this effect is negligible if a single camera is used and the environment can be controlled well enough to reduce the drift to a minimum. Even though ToF-cameras can measure distances of up to 7 m, we restrict ourselves to distances of maximal 2.5 m, as for larger distances the accuracy we want to achieve will most likely be not achievable anymore. The mentioned "wiggling" effect has not been investigated as well. The wiggling will have most influence for images with a large range of distance values. It will not affect the accuracy of objects significantly, if they are measured at a common distance with no large variations within the distances of the dataset. To come up with an accurate and stable ToF 3-D measuring device the neglected effects of course must be taken into account as well. Yet both of them influence the measurement more on a macro-level that can be ignored when determining the best precision and relative accuracy of the acquired data. As a last restriction, moving objects where not investigated. One of the main strengths of the modality of course is the ability to measure dynamic scenes. In order to find the absolute limits of accuracy the cameras can deliver, dynamic properties where not evaluated.

In this work we provide a quantitative evaluation of the precision and relative accuracy achievable with three ToF-cameras from different vendors and quantify those effects that have a strong influence on the measurement accuracy for objects within common distance ranges. A similar comparison was conducted by [AHK05], yet in that case different cameras where evaluated and the main interest was not on close-range applications but an overall characterization of the acquired data. For accurate close-range measurement devices, precisions of 1 mm are preferable. We want to evaluate the capability of the cameras to reach this accuracy without further post processing. Additional filtering would certainly improve the result, but the quality of the filtered data will still be guided by the raw input. Thus an evaluation of the sensor output will provide a good starting point when estimating what accuracy can be reached after processing.

2. Methods

The accuracy of ToF-sensors is mainly guided by the noise of the distance measurements which depends on the amplitude of the received signal at the ToF-chip. The relation between amplitude and standard deviation of distances of a ToF-sensor is well known and has been investigated earlier [FPR*09]. When trying to find the upper limit for the reachable accuracy of a ToF-sensor, the optimal camera parameters must be determined first. For a given distance and reflectivity of an object, the amplitude can be steered by changing the time used to sample the reference signal on the ToF-chip. This so-called integration- or shutter-time can be changed individually for any ToF-camera. Common values for integration times lie within 0.1 ms and 25 ms, the dis-

cretization of the parameter is between 0.001 ms and 0.1 ms. Adjusting the integration time of a ToF-camera means finding a trade-off between the lowest possible noise and the least saturation of the sensor. While lower noise will provide more precise and accurate results, saturation will make reconstruction of the phase of the received signal impossible and render distance computations for these pixels invalid.

For our study, we identified three working distances for which high accuracy measurements might be preferable: 500 mm, 1500 mm and 2500 mm. At the maximum distance of 2500 mm the complete body of a standing human is visible with most ToF-cameras and thus applications like full-body tracking can be implemented. The minimum working distance we identified at 500 mm, where a reasonable field of view can be covered and objects like hands or feet can be captured and tracked. This distance might be attractive both for interactive applications with a ToF-camera as input modality as well as for accurate scanning of small objects. When incorporating several cameras, even moving objects could be imaged and reconstructed. As a third distance for evaluation we used a working distance of 1500 mm, where the torso of a human will be well visible. This distance can be used for medical purposes like patient positioning and respiratory motion gating and management [PSHK08, SPH08, BBHRss].

2.1. Variation of Noise

For the proposed working distances and a set of different reflectivities, a large range of amplitude values can expected to be obtained when setting the integration time to a fixed value. Determination of the smallest standard deviation can only be done using a reference object with one certain reflectivity. For lower amplitude values acquired from objects with less reflectivity at the same distance a higher standard deviation of the distances will result. Yet this does not imply that the standard deviation of the distances is merely dependent on the measured amplitude value. We first investigate the variation of the distance noise depending on the reflectivity of the scene and the distances towards the objects. Evaluation was done by recording a set of different reference reflectivities where the integration time of the camera was adapted such that amplitude values within the same range where generated for all sequences. These recordings where done for 10 reflectivities at the three defined working distances.

2.2. Minimum Noise of Cameras

After investigation of the variation of the noise over the working distances and different reflectivities, the minimum reachable noise level for each camera was determined at the closest distance of 500 mm. The determined minimum noise is only valid for amplitude values which fall in the range of the amplitude values which where acquired when evaluating the noise. In real-world applications of ToF-devices, homogeneous amplitude images with values which are within the optimal range cannot be expected. Thus it is not only important to detect the upper accuracy limit but also to estimate the trend of the noise for decreasing amplitudes. For applications which acquire images with a wide range of amplitude values the stability of the noise over amplitude values becomes more and more important. If the captured objects can be steered to always show similar reflectivities, most accurate results can be expected.

(a) Camcube3 (b) Swissranger4000 (c) C70E

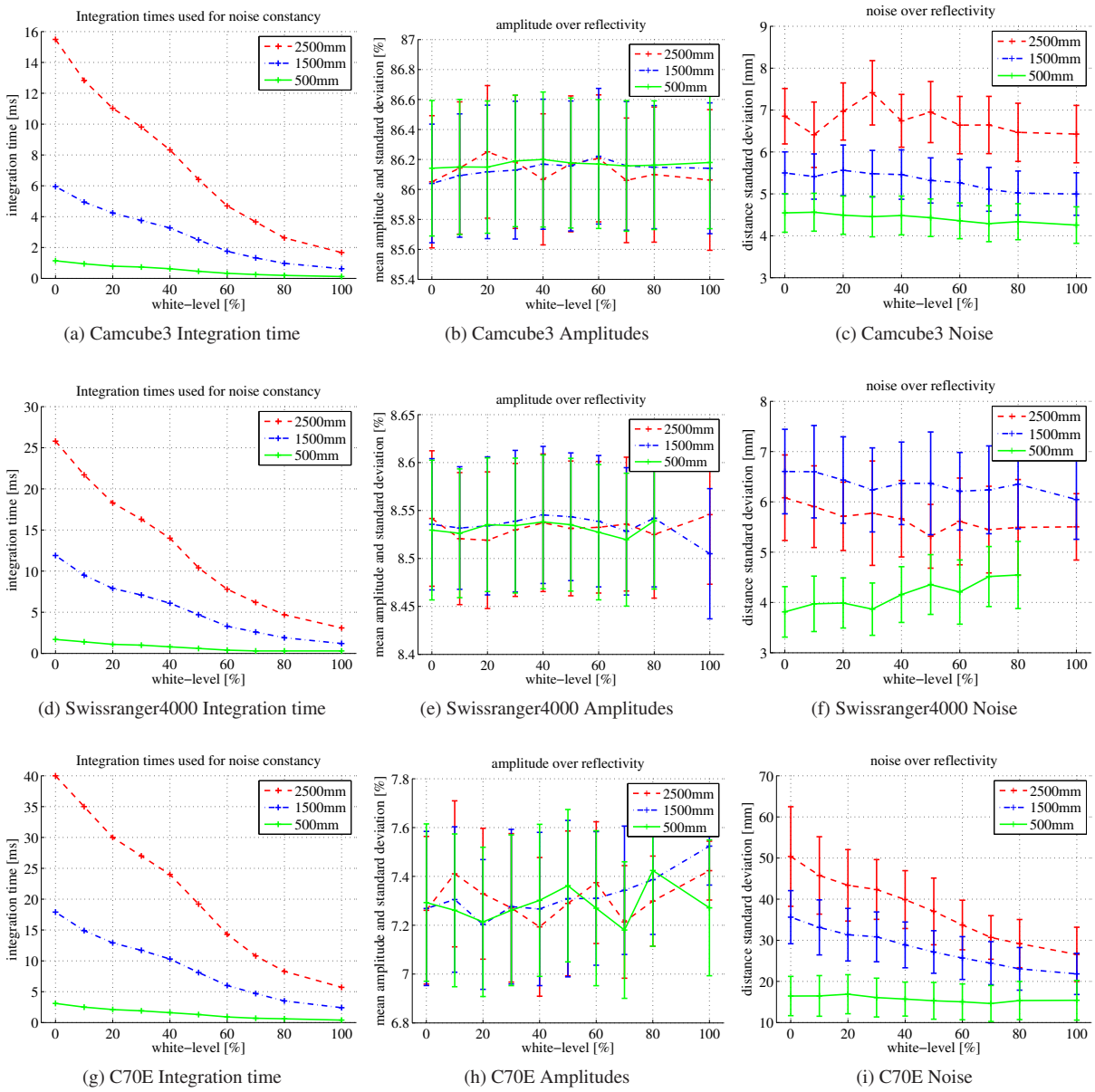
Figure 2: Time-of-Flight cameras evaluated in our study.

2.3. Amplitude Dependency of Measurements

Besides the distance noise of the ToF camera, the second main effect that is limiting the sensor as an accurate 3-D acquisition device is the amplitude dependency of the distance measurements. In a final step, the relative change of the distance measurements due to changes in the reflectivity has been quantified. Correction approaches to this effect have been published in more detail by [LSKK10]. In this work we focus on the quantization of the raw effect of the different cameras. The actual correction of the error is not addressed here.

3. Experiments and Results

For this study a set of three ToF cameras was evaluated: A Camcube3 from PMD Technologies, a Swissranger4000 from MESA Imaging and a C70E from Fotonic (see Fig. 2). All three cameras have unique characteristics. The Camcube3 is the only device that allows an individual placement of the illumination units, which makes it more flexible when integrating it into special setups. The Swissranger4000 features an optical feedback of the reference signal to reduce the signal drift to a minimum and the C70E uses a laser and an optical grating for illumination of the scene. The C70E lies within laser class 1 with and within laser class 3B without the grating, where the grating is necessary to ensure a homogeneous illumination of the complete scene. In addition, every camera uses a different ToF-chip for acquisition and demodulation of the signal. The one common parameter is the illumination wavelength which is in the near-infrared range for every camera (800 nm). A list of selected camera features is given in Table 2 at the end of the paper.



(a) Camcube3 Integration time (b) Camcube3 Amplitudes (c) Camcube3 Noise

(d) Swissranger4000 Integration time (e) Swissranger4000 Amplitudes (f) Swissranger4000 Noise

(g) C70E Integration time (h) C70E Amplitudes (i) C70E Noise

Figure 3: Constancy of noise of the evaluated cameras for fixed amplitude values. On the left, the actually used integration times are shown. In the middle, the mean amplitude value as well as the amplitude standard deviation is shown, on the right the standard deviation for changing reflectivities and distances can be seen. The amplitude values in the middle are given in % of a saturation value that was determined for each camera individually. The evaluation was done for a subset of central pixels which showed approximately the same amplitudes. The error bars on the right denote the standard deviation of the noise for the evaluated pixels.

3.1. Variation of Noise

For evaluation of the variation of noise for a given amplitude value a set of 10 different reflectivities was tested. On a laser-printer 10 different levels of gray where printed. The levels are described in percent of brightness, 0% denotes a completely black print, 50% gray and 100% a completely white paper. The aim was to get a set of different reflectivities which can be distinguished by the ToF-camera. In the remainder of this work we will refer to the different printout-intensities as "white-levels", where 100% is the raw white paper. For each camera, a reference amplitude value was defined by acquiring the 0% white-level (black) from the farthest distance and setting integration time to the maximum possible before saturation occurred within the black region. For the Swissranger4000 and the C70E this resulted in the maximum possible integration time (25.8 ms and 40 ms respectively), for the Camcube3 a integration time of 15.5 ms was chosen. For every camera thus a reference amplitude value was determined which was used throughout the evaluation of the variation of the noise (see Fig. 3, middle column). As the numerical amplitude values of the individual ToF cameras do not correspond to a common unit and the optical power of the different cameras is not the same as well, a direct comparison of the determined amplitude reference values is not possible. To come up with a comparability of the amplitudes, we determined for each camera a numerical saturation value empirically and scaled the values in percentage to this value. By looking at the levels of saturation it also becomes obvious that the noise levels from this evaluation cannot be compared between the single cameras. The results of the evaluation are shown in Figure 3. For the evaluation of the variation of noise it was necessary to retain the reference amplitude value also for highest amplitude values (100% white-level at the closest distance). For the Swissranger4000 the lowest possible integration time of 0.3 ms did not suffice to show amplitudes at the reference level, that is why the samples are missing in Figures 3(e) and 3(f). For the C70E, the integration time at 500 mm for 100% was reduced to 0.4 ms, for the Camcube3 a integration time of 0.117 ms was used. While the noise is quite constant for a common distance with varying reflectivities, all cameras measured more precisely for closer distances. The only camera that showed a dependency of the noise for different reflectivities was the C70E at the two longer distances.

3.2. Minimum Noise

The accuracy of the sensors were evaluated at the closest distance of 500 mm. For the 100% pattern the integration time was increased to the highest reasonable value where not yet saturation occurred. For the Camcube3 this value was 0.1 ms, for the C70E 4 ms and for the Swissranger4000 1.3 ms. The pixel-wise temporal standard deviation of the acquired data was evaluated for a large region of interest which covered more than 5000 pixels. The results of this evaluation are given in Fig. 4 and Table 1. In addition to the minimum noise that could be measured, we acquired all reflectivity patterns from the noise variation evaluation with the fixed integration time to estimate the stability of the noise when acquiring objects with heterogeneous reflectivity.

	Camcube3	SR4000	C70E
mean noise	4.8 mm	1.6 mm	0.9 mm
integration time	0.1 ms	1.3 ms	4.0 ms

Table 1: Standard deviation at 500 mm for 100% white-level.

3.3. Amplitude Dependency

For investigation of the amplitude dependent distance deviation a calibration was done [Zha00] and a per-pixel fixed pattern offset was computed. After computation of the intrinsic parameters the normalized projection vectors $\mathbf{v}_i$ for each pixel i of the cameras were computed. The world coordinate $\mathbf{c}_i$ corresponding to the measured distance value d_i is computed as

$$\mathbf{c}_i = d_i \mathbf{v}_i. \tag{1}$$

The camera was set up to face a flat wall at the respective working distance. The integration time was adjusted to maximize the quality of the acquisitions and a reference frame (averaged over 50 datasets) was recorded. Assuming the cameras focal point $\mathbf{o}$ at the origin of the coordinate system, the position of the wall can be described by a plane P. The position of P is defined by its normal $\mathbf{n}$ and a reference point $\mathbf{p}$. For any pixel i, the intersection $\mathbf{I}_i$ between $\mathbf{v}_i$ and P computes as

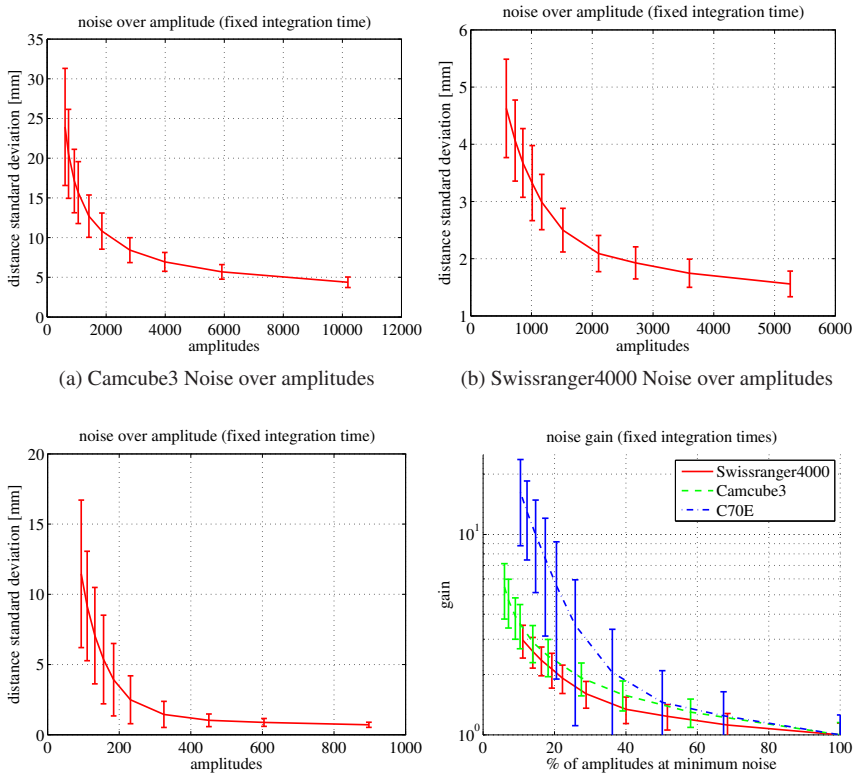
$$\mathbf{I}_i = -\mathbf{v}_i \frac{\mathbf{n}^T \mathbf{p}}{\mathbf{n}^T \mathbf{v}_i} \tag{2}$$

and the distance between $\mathbf{o}$ and $\mathbf{I}_i$ is the expected ground truth radial distance that should be measured by the ToF-sensor. Due to the effects that where stated in this paper the measured point $\mathbf{c}_i$ will deviate from the expected value $\mathbf{I}_i$. The remaining offset

$$e_i = \|\mathbf{I}_i - \mathbf{c}_i\|_2 \tag{3}$$

was computed for all pixels using the measurements from the averaged reference acquisition. In this way a flat measurement of the wall can be assured and relative distance errors which come from changes in the amplitude values can be investigated. A 2-D sketch of the computation of e_i is given in Figure 5.

To obtain the ground truth position of the wall relative to the camera, checkerboard patterns where attached and the cross-section points where extracted. For each distance, 13 - 19 images of the checkerboard were recorded. The squares of the pattern were of size 160 mm $\times$ 160 mm for the longest distance, 80 mm $\times$ 80 mm for the middle and



(a) Camcube3 Noise over amplitudes

(b) Swissranger4000 Noise over amplitudes

(c) C70E Noise over amplitudes

(d) Relative noise gain regarding to relative changes of amplitudes.

Figure 4: Minimum noise (standard deviation) measured for several white-levels at 500 mm distance with fixed integration time. The error bars denote the standard deviation of the noise for all evaluated pixels. The numerical amplitude values are not comparable between the cameras. For (d) the amplitudes have been normalized to the value at the minimum noise (100% white-level).

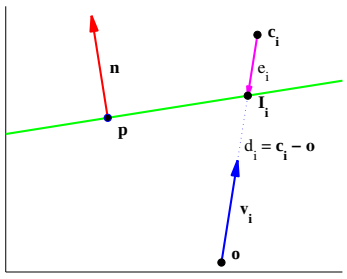


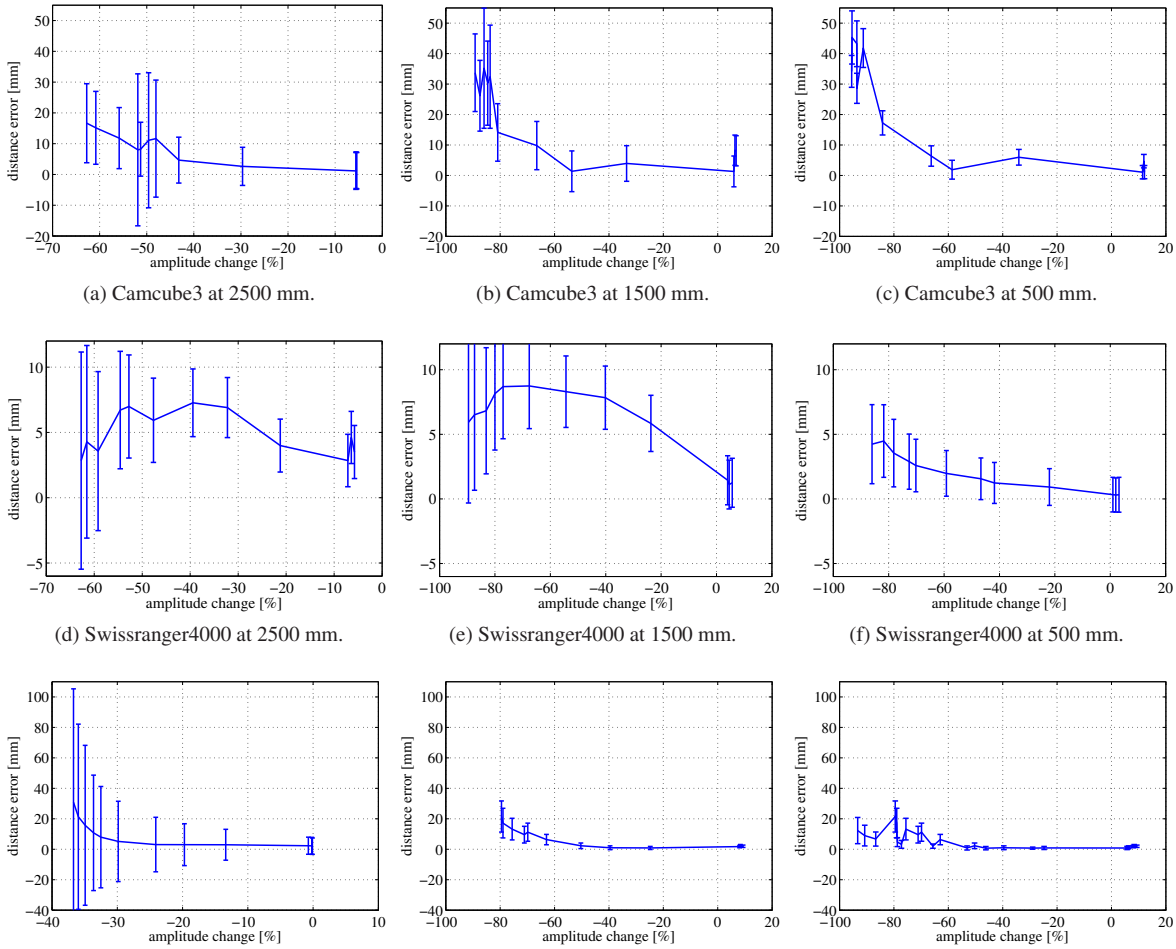
Figure 5: Computation of distance offset e_i.

40 mm × 40 mm for the closest distance of 500 mm. For every recording j the corresponding plane P_j with n_j and p_j was extracted using the known intrinsic parameters of the camera and the estimation of the homographies as proposed in [Zha99]. As the pattern was only moved within the plane of the wall, the angles between any pair of normal vectors as well as the distances between the planes are theoretically zero. To reduce the effect of non-optimal estimation of the plane position and orientation, the final plane P was set to be the mean plane over all P_j. The per-pixel offsets e_i where computed for each camera at each distance.

For evaluation of the amplitude-dependent distance errors, papers with four different white-levels where attached to the wall, sequences of 50 frames were averaged and the datasets where calibrated using the projection vectors $\mathbf{v}_i$ and offset corrections e_i (results see Fig. 1). For each white-level, a subset of pixels with similar amplitudes was chosen and the distance

$$\xi_i = \|(\mathbf{c}_i - e_i\mathbf{v}_i) - \mathbf{I}_i\|_2 \tag{4}$$

was computed for all these pixels. The actual distance errors which are shown in Fig. 6, are the mean values of all ξ_i computed for one specific subset. Each point accounts for $100 - 1000$ pixels. Due to the inhomogeneous illumination of the cameras, the computed offset corrections e_i do not



(a) Camcube3 at 2500 mm. (b) Camcube3 at 1500 mm. (c) Camcube3 at 500 mm.

(d) Swissranger4000 at 2500 mm. (e) Swissranger4000 at 1500 mm. (f) Swissranger4000 at 500 mm.

(g) C70E at 2500 mm. (h) C70E at 1500 mm. (i) C70E at 500 mm.

Figure 6: Relative amplitude dependency. The error bars denote the mean standard deviation of all pixels contributing to the point. The standard deviation of ξ_i over all pixels contributing to one point is approximately half of the distance standard deviation. Amplitude changes were scaled to % of the previously determined saturation values (see Fig. 3).

only correspond to single pixels but also to specific amplitude values. As a consequence, the results are given depending on the relative change of the amplitude values (see x-axis of Fig. 6). All three cameras have in common that for lower amplitudes the measurement error increases, but the shape and the magnitude of the error differ. While for the Swissranger4000 the maximum error is still below 10 mm, for the other cameras it increases up to 30 mm and more. Please note that the Swissranger4000 and the C70E show smaller errors in the close range while the Camcube3 exhibits the strongest effects in this distance.

4. Conclusion and Outlook

Time-of-Flight cameras show different systematic behavior depending on what hardware components are used. We have compared three cameras from different vendors and investigated the capability of these sensors to be used as accurate close-range 3-D sensing devices. All cameras measure more precisely at smaller distances, even if the amplitude of the received signal remains the same. In order to achieve accurate measurements, applications using this modality should use short working distances. A dependency of the measured distances on the signal amplitude is evident, but the error differs for all cameras in magnitude and shape. To achieve accurate range-measurements with ToF-cameras, the cameras need to be calibrated specifically for the expected amplitude range and the recorded objects should exhibit a homogeneous reflectivity. In this case, sub-centimeter accuracy is possible with all cameras. With increasing differences of amplitude values within one acquisition, a correction of the am-

	Camcube3	Swissranger4000	C70E
lateral resolution [pixels]	200×200	176×144	160×120
modulation frequency	20 MHz	30 MHz	44 MHz
maximum distance	7.5 m	5 m	3.4 m
illumination wavelength	820 nm	850 nm	808 nm
illumination type	LED	LED	laser
connection to PC	USB	Ethernet	Ethernet
integration time range	0.011 ms – 21 ms	0.3 ms – 25.8 ms	0.1 ms – 40 ms
integration time discretization	0.001 ms	0.1 ms	0.1 ms
solid angle (central pixel)	0.2°	0.22°	0.38°
pixel discretization (at 500 mm)	1.7 mm	1.9 mm	3.3 mm
field of view horizontal	39.8°	40.6°	60.5°
field of view vertical	39.6°	33.0°	45.5°

Table 2: Overview of the evaluated Time-of-Flight cameras

plitude related error becomes necessary for all cameras. An evaluation of the amplitude-dependent error in close-range and with such high accuracy has not been performed before. Post-processing of the data like denoising and application of dedicated calibration models which reduce the amplitude dependency might help to improve the accuracy even further.

5. Acknowledgments

The project is supported by the National Research Fund, Luxembourg. The authors gratefully acknowledge funding of the Erlangen Graduate School in Advanced Optical Technologies (SAOT) by the German National Science Foundation (DFG) in the framework of the excellence initiative.

References

[AHK05] ANDERSON D., HERMAN H., KELLY A.: Experimental characterization of commercial flash ladar devices. In *International Conference of Sensing and Technology* (November 2005). 2

[BBHRss] BAUER S., BERKELS B., HORNEGGER J., RUMPF M.: Joint ToF Image Denoising and Registration with a CT Surface in Radiation Therapy. In *Proceedings of the International Conference on Scale Space and Variational Methods in Computer Vision* (2011, in press). 3

[FPR*09] FRANK M., PLAUE M., RAPP H., KÖTHE U., JÄHNE B., HAMPRECHT F. A.: Theoretical and experimental error analysis of continuous-wave time-of-flight range cameras. *Optical Engineering 48*, 1 (2009). 2

[KBKL10] KOLB A., BARTH E., KOCH R., LARSEN R.: Time-of-Flight Cameras in Computer Graphics. *Computer Graphics Forum 29*, 1 (2010), 141–159. 1

[KRI06] KAHLMANN T., REMONDINO F., INGENSAND H.: Calibration for increased accuracy of the range imaging camera SwissrangerTM. In *International Archives of Photogrammetry and Remote Sensing, Part B5* (2006), vol. 37. 2

[Lan00] LANGE R.: *3D Time-of-flight distance measurement with custom solid-state image sensors in CMOS/CCD-technology*. PhD thesis, University of Siegen, 2000. 1

[LK07] LINDNER M., KOLB A.: Calibration of the intensity-related distance error of the PMD TOF-Camera. In *SPIE: Intelligent Robots and Computer Vision XXV* (2007), vol. 6764, pp. 6764–35. 2

[LSKK10] LINDNER M., SCHILLER I., KOLB A., KOCH R.: Time-of-Flight Sensor Calibration for Accurate Range Sensing. *Computer Vision and Image Understanding 114*, 12 (2010), 1318 – 1328. 3

[OLK*04] OGGIER T., LEHMANN M., KAUFMANN R., SCHWEIZER M., RICHTER M., METZLER P., LANG G., LUSTENBERGER F., BLANC N.: An all-solid-state optical range camera for 3D real-time imaging with sub-centimeter depth resolution (SwissRanger). In *SPIE* (2004), vol. 5249, pp. 534–545. 1

[PHS*09] PENNE J., HÖLLER K., STÜRMER M., SCHRAUDER T., SCHNEIDER A., ENGELBRECHT R., FEUSSNER H., SCHMAUSS B., HORNEGGER J.: Time-of-Flight 3-D Endoscopy. In *MICCAI 2009, Part I, LNCS 5761* (2009), pp. 467–474. 2

[PSHK08] PENNE J., SCHALLER C., HORNEGGER J., KUWERT T.: Robust Real-Time 3D Respiratory Motion Detection Using Time-of-Flight Cameras. *Computer Assisted Radiology and Surgery 2008 3*, 5 (2008), 427–431. 3

[SJ09] SCHMIDT M., JÄHNE B.: A Physical Model of Time-of-Flight 3D Imaging Systems, Including Suppression of Ambient Light. In *Dynamic 3D Imaging* (2009), Kolb A., Koch R., (Eds.), vol. 5742 of *Lecture Notes in Computer Science*, Springer, pp. 1–15. 2

[SPH08] SCHALLER C., PENNE J., HORNEGGER J.: Time-of-Flight Sensor for Respiratory Motion Gating. *Medical Physics 35*, 7 (2008), 3090–3093. 3

[SSBH11] STÜRMER M., SEILER C., BECKER G., HORNEGGER J.: Alignment of multiple time-of-flight 3d cameras for reconstruction of walking feet. In *Proceedings of ISB2011* (2011). 2

[Zha99] ZHANG Z.: Flexible camera calibration by viewing a plane from unknown orientations. In *Proceedings of IEEE International Conference on Computer Vision* (1999), vol. 1, pp. 666–673 vol.1. 6

[Zha00] ZHANG Z.: A flexible new technique for camera calibration. *IEEE Trans. Pattern Anal. Mach. Intell. 22*, 11 (2000), 1330–1334. 5