A theoretical and experimental investigation of the systematic errors and statistical uncertainties of Time-Of-Flight-cameras

H. Rapp*, M. Frank, F.A. Hamprecht and B. Jähne

Digital Image Processing,
Interdisciplinary Center for Scientific Computing,
University of Heidelberg,
Im Neuenheimer Feld 368,
69120 Heidelberg, Germany
Fax: +49 6221548850
E-mail: holger.rapp@iwr.uni-heidelberg.de
*Corresponding author

Abstract: The following paper presents a model to predict the systematic errors and statistical uncertainties of Time-Of-Flight (TOF) 3Dimaging systems. The experimental data obtained with a custom build test setup show that the SD of the depth signal rises approximately quadratically with the depth. The most significant systematic depth error is periodic with an amplitude of around 50mm. It is provoked by the inharmonic correlation function. The inhomogeneity in each pixel (fixed pattern) accounts for a depth error of about 20mm, while illumination and reflectivity variations cause depth errors of less than 10mm, provided that no overflows occur.

University in 1977, 1980, and 1985, respectively, and a habilitation degree in applied computer science from the University of Hamburg-Harburg in 1992. Between 1980 and 1987, he was a Research Assistant in the Institute for Environmental Physics at the Heidelberg University in 1984, a Visiting Scientist in IMST, University of Marseille in Marseille, and a Heisenberg Fellow in 1986 and 1987. Between 1988 and 2004, he was a Marine Research Physicist in the Scripps Institution of Oceanography, University of California, and, since 1994, he has been a Professor of physics and computer science at the Interdisciplinary Center of Scientific Computing and the Institute for Environmental Physics and Head of the research groups on image processing and small-scale air–sea interaction. His research interests include small-scale air–sea interaction processes, image processing, and the application of image processing techniques in science and industry.

Fred A. Hamprecht studied chemistry in the Swiss Federal Institute of Technology (ETH) in Zurich and received his PhD from the same school. After a brief period at the Seminar for Statistics, ETH Zurich, he became a Robert Bosch-endowed Professor for Multidimensional Image Processing at the University of Heidelberg in 2001, and an Affiliate Professor of Pathology at Children's Hospital, Boston in 2006. He is a Member of the Steering Committee of the German Association for Pattern Recognition (DAGM), and of the Advisory Board of the Heidelberg Image Processing Forum. He is interested in all aspects of data analysis, ranging from signal and image processing to supervised, semi-supervised and unsupervised learning. With his group, currently he develops automated diagnostic systems with applications both in industrial quality control and medicine, where particular emphasis is put on a principled treatment of spatial and temporal context.

1   Introduction

The invention of Time-Of-Flight (TOF)-camera systems with a modulated light source for illumination which measure the phase-shift directly by correlation on chip (see Schwarte et al., 1997; Luan et al., 2001) makes depth-measuring imaging systems almost as simple as standard intensity measuring cameras. For practical usage, it is important to specify in detail the systematic and statistical errors of this new type of cameras. While simple linear models for TOF-cameras are well known and some limited studies of the accuracy of the depth measurement are available (e.g. Lindner and Kolb, 2006; Strehler, 2007; Kahlmann, Remondino and Ingensand, 2006), a systematic investigation is still missing.

   This paper discusses the first steps of such a systematic study. It is based on a detailed model of a TOF-camera, which includes an error-propagation model from the measured intensities to the estimated distance (Section 2). A test stand with motor-driven linear tables was used to acquire data with targets at precise distances from the camera (Section 3.3). The statistical and systematic errors in the distance measurements are analyzed in detail (Section 4.2 and Section 4.1) as well as the pixel-base non-uniformity of the depth signal (Section 4.1.3).

2   TOF-camera model

A correlating TOF-system (Figure 1b) always includes an active modulated light source to illuminate the scene, normally in the infrared spectrum (with wavelengths of around 850 nm). Currently, all systems use LEDs. Future systems will likely also deploy other light sources, e.g. vertical lasers.

Figure 1   Investigated TOF system and measurement principle (see online version for colours)
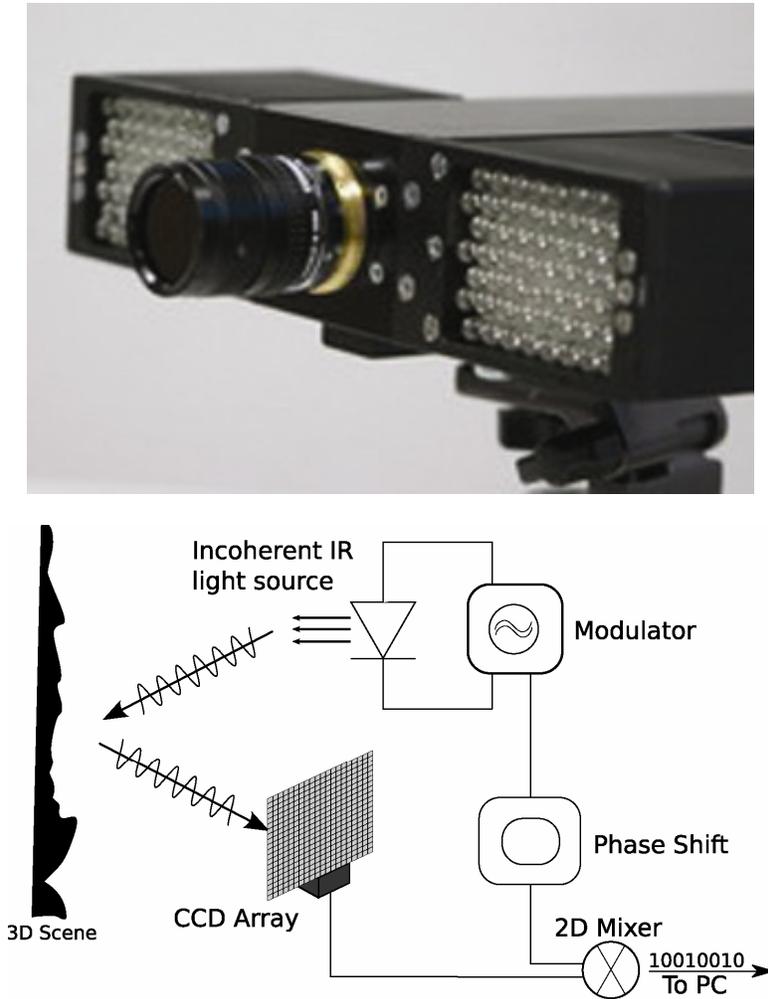
This light travels from the camera to the target and back again and thus experiences a phase shift

$$\varphi_d = \frac{4\pi v}{c}d,\tag{1}$$

which is directly proportional to the distance d.

The outgoing optical signal is amplitude modulated with a fixed frequency v, is reflected by the scene and returns to the camera. Inside the camera – directly on chip in most modern systems – the returning optical signal is correlated with the electrical reference signal which is in phase with the modulated outgoing light. The system therefore directly measures the Correlation Function (CF) of the emitted and recorded signals.

The CF contains information about the returning optical signal: the constant DC offset c, the modulation amplitude A and the phase shift from which the distance d between the camera and the measured object can be computed according to Equation (1). The shape of the CF is known if the exact form of the light modulation is known. Because the CF can only be sampled at a small number of points, the parameters c, A and φ are inferred from a regression on three or more sample points.

A mathematical model of the TOF-camera was first published by Xu (1999) and Schneider (2000). This paper follows the mathematically equivalent discussion of Plaue (2006) which is shorter and more flexible and elegant due to the use of complex notation.

Given some modulation function $O(v, \text{t})$ with fixed frequency v, the recorded intensity $I(v, \text{t})$ will have the same frequency and shape. The n-th correlation frame is then calculated with various constant phase shifts α_n:

$$I_n(v,\alpha_n) = \frac{1}{t-t_0} \int_{t_0}^{t} I(v,t)\, O(v,t+\frac{\alpha_n}{v}+t_d)\, \mathrm{d}t \tag{2}$$

If $N \geq 3$ the optimal solutions for the three unknowns are given by

$$\varphi_d = \arg\left(\sum_{n=0}^{N-1} I_n\, \mathrm{e}^{-2\pi i(n/N)}\right), \qquad c = \frac{1}{N}\sum_{n=0}^{N-1} I_n, \qquad A = \frac{2}{N}\left|\sum_{n=0}^{N-1} I_n\, \mathrm{e}^{-2\pi i(n/N)}\right|. \tag{3}$$

Given that the variance of the I_n are all equal to σ^2, the variances of the parameters become:

$$\mathrm{var}\left(\varphi_d\right) = \frac{\sigma^2}{2A^2}, \qquad \mathrm{var}\,(c) = \frac{\sigma^2}{4}, \qquad \mathrm{var}\,(A) = \frac{\sigma^2}{2}. \tag{4}$$

In all available systems, $N = 4$ and $\alpha_{0,\,1,\,2,\,3} = 0$, $\pi/2$, π, $3\pi/2$. Xu (1999) showed that this solution is exact for sinusoidally modulated signals. If either the optical or the reference signal is not symmetric or if they differ in form (e.g. one is a rectangle function and the other a sinus) the CF will have a different shape. This results in periodic systematic errors in the phase calculation and therefore in the depth. This effect was mentioned by Schneider (2000), a mathematical explanation and discussion can be found in Plaue (2006).

3   Experimental setup and data processing

This section describes briefly the experimental setup and the data acquisition and processing. Though this section contains some informations on how to enhance the data delivered by the PMD[vision] 19k – the camera system used for the acquisition – this is not the main focus of this paper and therefore only discussed briefly. A detailed

description of the experimental setup and an exact explanation of the data processing can be found in Rapp (2007).

3.1   The camera system

All experimental data for this paper was gathered using the PMD[vision][®] 19k camera system by PMDTechnologies GmbH[1] (Figure 1a). The camera uses LEDs with a wavelength of 870 nm and a total optical power of around 3 W. The LEDs are mounted in two arrays, one on either side of the camera. For all experiments, the modulation frequency was kept at the default value of $v = 20$MHz resulting in an unambiguous depth range of $d_{max} = c/2v = 7.5$m. The camera acquires samples at four phase shifts for each CF, taking two samples on each measurement (one with α_n and one with $\alpha_n + \pi/2$). This redundant information is used to correct inhomogeneities in the PMD (see Lange, 2000; Justen, 2001 for details), but contains more valuable information as discussed in Section 3.4.2.

The correlation inside the camera is performed using a Complementary Metal Oxide Semiconductor (CMOS)-based optical semiconductor called PMD. This technique increases speed and decreases cost and noise of the system. The camera has a resolution of 160×120 pixels with a frame rate of 5–12 fps. The data is digitised with 12 bit and delivered to the host through a firewire interface. Since the camera has no Suppression of Background Illumination (SBI), all measurements were performed in a dark room without any IR light source that might interfere.

3.2   Targets

The targets (Figure 2) were custom built for this experiment with the theory of TOF cameras in mind. Therefore, special care was taken to make sure that the reflectivity of the targets at all points is known: the frames were covered with black cardboard to ensure a low reflectivity at the borders, the reflecting areas were made from Photo-Cards by Fotowand-Technic.[2] The high-reflectivity target uses photocards with 84% reflectivity, the checkerboard consists of 90 × 90mm squares with reflectivities of 12.5, 25, 50 and 84% in a regular pattern.

**Figure 2**   Targets used for data acquisition
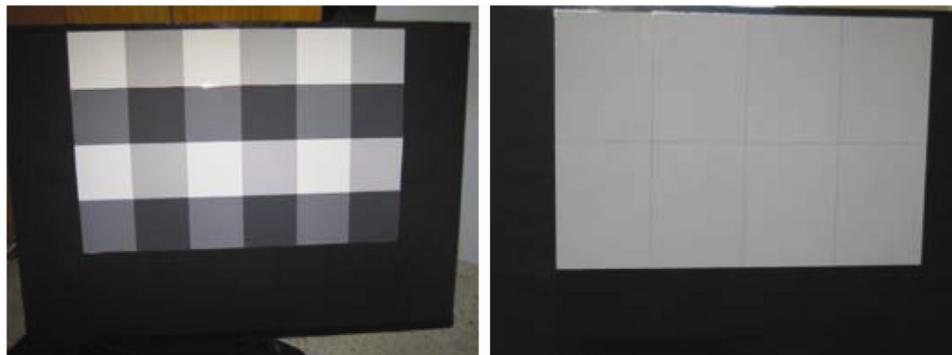
(a) Checkerboard target   (b) High-reflectivity target

3.3 Experimental setup

The experimental setup (Figure 3; Rapp, 2007) includes a camera fixed on a 3m long linear position table with an accuracy of < 1 mm. The target is mounted on a similar table which is aligned precisely with the first table. This setup allows a sub millimeter precise depth positioning in the range of $d_0 \approx 0.2m < d < 6m + d_0$. The two tables are mounted on lockable rolls and can be separated to move the complete experiment to other surroundings (e.g. into plain sunlight). To remove systematic errors through unwanted reflections of the IR light from the linear position tables, a zig-zag shading has been installed. The camera and tables are remote controlled through a PC, so automated measurements are possible. The PC also processes and displays the data.

Figure 3 Experimental setup
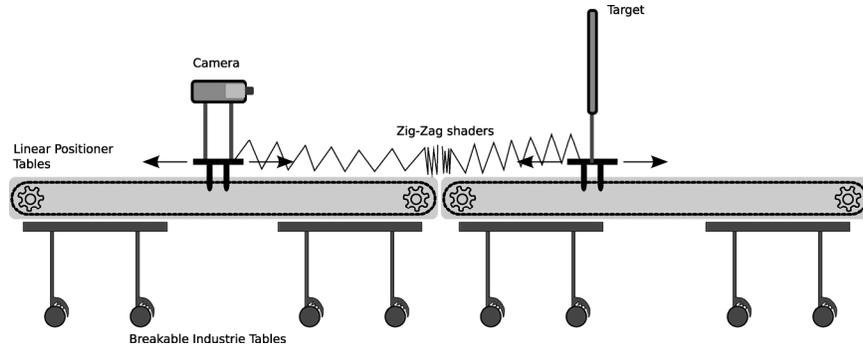
The room has been held dark for all measurements and all objects in the vicinity of the experiments have been covered with black velvet to avoid reflections from the room that could deteriorate the data.

3.4 Processing of the acquired data

The most fundamental data delivered by the camera are the eight raw channels.[3] The camera already delivers depth, amplitude and phase shift for convenience, but most of the problems discussed below can only be avoided by directly accessing the raw data, enhancing it and then doing the math.

3.4.1 Correcting inhomogeneities

As mentioned in Section 3.1, the camera takes two samples on four phase shifts. This redundant information is used to correct inhomogeneities in the two gates of the PMD chip. Theoretically

$$I^A_{\alpha_n} = I^B_{\alpha_n + \pi/2} := I_n \qquad (5)$$

with I_B and I_A being the correlate measured on the first (gate A) and the second (gate B) gate of the PMD, respectively. But due to technical reasons, each gate has a constant

offset error: $I_{\alpha_n}^{A} = \hat{I}_{\alpha_n}^{A,B} + \delta a$ and likewise for B. To calculate for example $d := I_0 - I_2$ there are two possibilities:

$$d = (\hat{I}_0^A + \delta a) - (\hat{I}_0^B + \delta b) \tag{6}$$

$$d = (\hat{I}_\pi^B + \delta b) - (\hat{I}_\pi^A + \delta a) \tag{7}$$

Adding these two equations and dividing by two cancels all constant offsets from the data and yields:

$$d = \frac{1}{2}(\hat{I}_0^A - \hat{I}_0^B) - (\hat{I}_\pi^A - \hat{I}_\pi^B) \tag{8}$$

3.4.2 Detecting and correcting overexposed pixels

Detecting overexposed pixels can be difficult with TOF-cameras. But the PMD-camera offers an easy solution to this problem: the eight raw channels are monotonically increasing with intensity of exposition. All pixels that are below[4] a certain digit (2,500) in all raw channels are over exposed.

This method avoids using the amplitude as an indicator, because there are ranges where the amplitude takes all possible values while the pixel is still overexposed.

3.4.3 Taking a correct mean

For a proper investigation of the systematic errors, the data in all measurements were taken as a mean of 100 measured frames at each stop of the linear position table, if no mean was taken this is noted explicitly in the text.

The mean was taken from the raw data before any calculation took place. This has mainly two reasons. The first is that incorrect circular averaging of the data – which leads to problems where the depth error is higher than the unique measurement range (at low distance or low amplitude) – is avoided. The raw data is normally distributed and therefore uncritical to handle. The second is that by averaging the raw data, the amplitude is implicitly used as weight. This reduces the impact of totally outlying measurements on the mean.

Taking the mean of the raw amplitudes also gives a benefit with the quantisation taking place inside the camera. Taking the mean of some discrete frames will bring a higher granularity into the calculation. Note, though, that the amplitude-dependent quantisation error cannot be avoided in this way, see Frank (to be published) for a more detailed discussion of these problems.

3.4.4 Spherical correction

As with most 3D data acquisition systems, TOF-cameras can only measure the spherical distance to a point in space. Therefore, a plane positioned parallel to the picture plane is detected as a hyperbolic object. To avoid theses errors here, the parabolic depth was transformed to the depth perpendicular to the image plane.

4   Experimental results

4.1   Systematic errors

4.1.1   Near field

The attachment of the LED arrays on the left and the right side of the camera causes a near field effect. Up to $d \approx 1.5$ m, the two LED arrays do not work as a point light with a unique amplitude, instead the light from each array has a different path length and time of flight. This results in a systematic error of exponential shape as can be seen in Figure 4a. The amplitude though behaves as expected as shown in Figure 4b

Figure 4   Plot of four pixels near the target centre and the high-reflectivity target (see online version for colours)
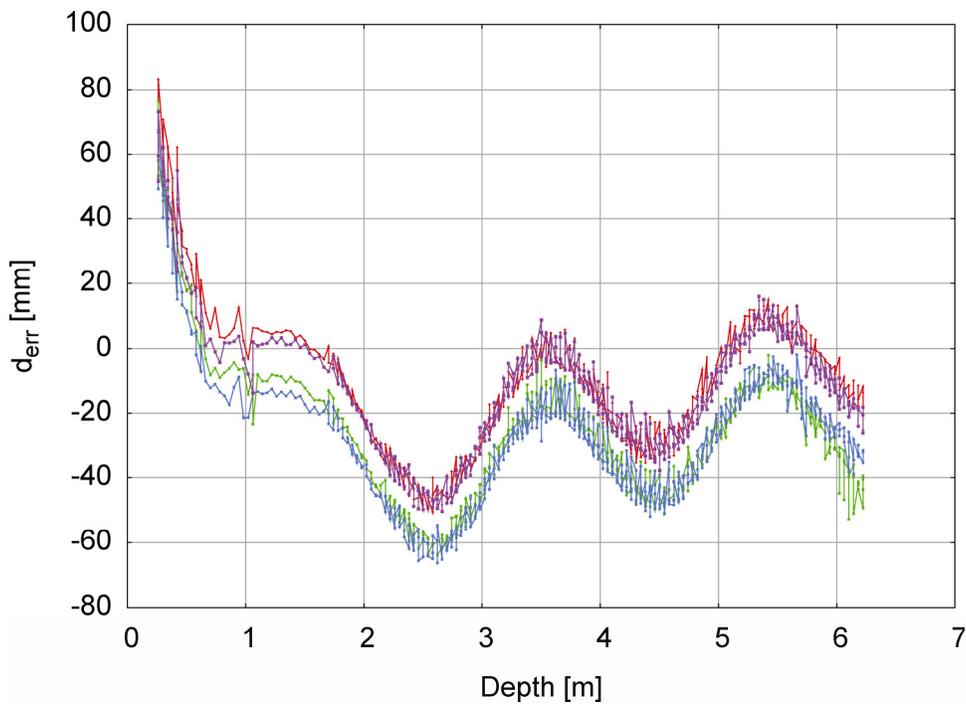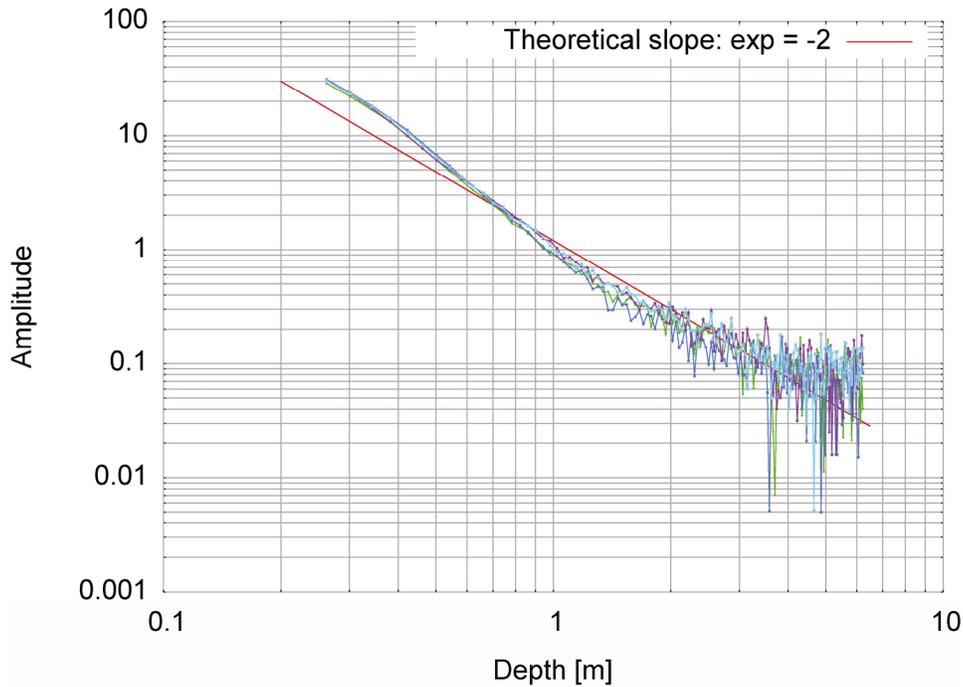
4.1.2 Dependency of depth error on amplitude

In Figure 5, a plot of depth-to-depth error for various integration time is shown. With increasing integration time, the amplitude also increases for a fixed position. There is no fundamental connection visible: the curves lay very well on top of each other, only the 12 ms curve varies a little stronger, it is below the others at $2m < x < 3m$ and above for $4m < x < 5.5$ m, but in the second interval, its variance is already quite high (the variance does depend on the amplitude, see Section 4.2).

Taking the mean depth error in the range between $3.75m < x < 5.75m$ supports this arguments. Table 1 shows that 12 ms has the smallest error, but there is no relationship visible.

Table 1   Mean depth error, high-reflectivity target and mean of depth error in the range $3.75m < x < 5.75$ m

Integration time (ms)	5	12	20	30	50
Mean depth error (mm)	−18.76	−11.09	−14.28	−16.90	−16.78

Figure 5   Plot of depth to depth error for one central pixel with various integration times, A > 3 for all pixels (see online version for colours)
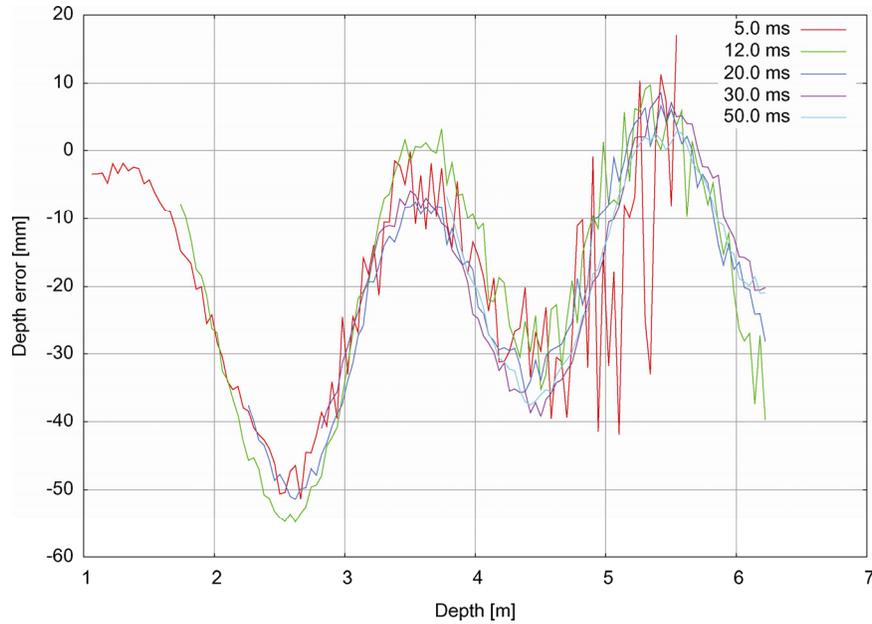
4.1.3 *Inhomogeneities of pixels*

We have already discussed how to handle the inhomogeneities of the gates inside each pixel of the camera, now we will turn to the inhomogeneities of the pixels in respect to each other.

Each CMOS camera has a fixed pattern offset created through the inhomogeneities of the various pixels. In Figure 4a, four pixels are plotted. Apart from the near field, the four lines keep a constant offset to each other, therefore the inhomogeneities of the pixels only results in a constant offset for each pixel in the depth calculation.

4.1.4 *Inharmonic correlation function*

As discussed in Section 2, an inharmonic CF results in a periodic systematic error which depends on the distance of the measured object.

The cameras LED signal is highly anharmonic and it correlates internally with a square wave reference function. The resulting systematic error can be seen in Figure 4(a) as soon as the near field errors loose significance ($x > 1.5$ m). Of the four pixels plotted in the figure, two pairs have the same inhomogeneity each. They stay on top of each other for the whole depth range. This leads to the conclusion that this error depends only on the distance between camera and object and is therefore the same for all pixels – exactly as the theory predicts.

4.2 Statistical uncertainty of depth measurements

According to the model of a point light source and the variance predictions in Equation (4), the following holds:

$$A \propto \frac{1}{d^2} \quad \text{and} \quad \mathrm{Var}(d) \propto \frac{1}{A^2} \Rightarrow \sigma_d \propto d^2 \tag{9}$$

Figure 6 shows the corresponding plot. The relationship holds approximately. But since the LED arrays are no point light source and due to quantisation systematic errors on low amplitudes, the exponent is estimated to around 2.5, while other experiments investigating the depth-amplitude proportionality[5] ask for a smaller exponent of around 1.9.

Figure 6   Plot of standard depth deviation to real depth for four pixels near the target centre, with high-reflectivity target, and the 0.5 ms integration time (see online version for colours)
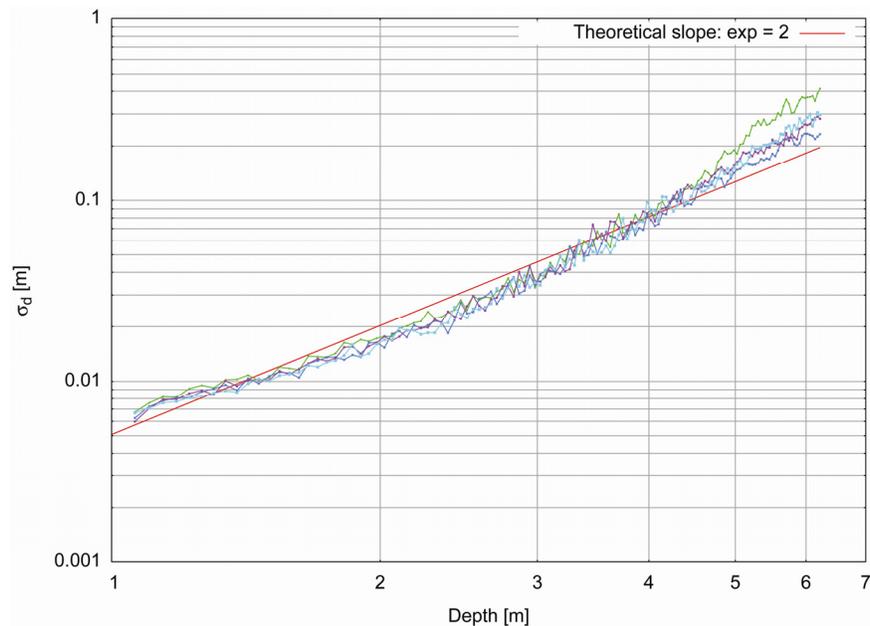
5 Discussion and conclusion

This paper gave a general conspectus about TOF-cameras. A model was presented and verified with experimental data using the PMD[vision] 19k camera, and systematic and statistical errors were discussed. The model proved successful for this camera, but to check its general validity further investigations are required with other TOF-camera systems. The investigated system shows a lot of systematic errors. The periodic variation due to the inharmonic CF provokes a periodic depth error of around 50mm, the inhomogeneity of the pixels accounts for 20mm and the statistical error rises quadratically with the depth. Most of the errors are very easy to correct, though: overexposed pixels can be masked and the periodic offset due to the inharmonic CF and

the constant pixel offsets can easily be removed with a calibration and a lookup table. With these corrections, reliable 3D data can be acquired even with today's systems.

Current TOF-systems are most seriously limited by their narrow depth dynamics at a constant exposure time. A much better performance could be obtained by autoadapting the integration for each individual pixel according to the local irradiance, thus providing a sufficiently high global dynamic range.

The correlating TOF-camera 3D measurement technology is a young but promising new technology

which shows a convincing performance even in a prototype state. It is reasonable to expect significant progress in the near future.

References

Frank, M. (2007) *Investigation of a 3D Camera, Diploma Thesis*.

Justen, D. (2001) *Untersuchung Eines Neuartigen 2D-gestützten 3D-PMD Bildverarbeitungs Systems*. Department of Electrical Engineering and Computer Science, University of Siegen.

Kahlmann, T., Remondino, F. and Ingensand, H. (2006) 'Calibration for increased accuracy of the range imaging camera SwissRangerTM, *ISPRS Image Engineering and Vision Metrology*, Vol. 36, pp.136–141.

Lange, R. (2000) *3D Time-of-Flight Distance Measurement with Custom Solid-State Image Sensors in CMOS/CCD-Technology*. Department of Electrical Engineering and Computer Science, University of Siegen.

Lindner, M. and Kolb, A. (2006) 'Lateral and depth calibration of PMD distance sensors', Paper presented in the Proceedings of the *International Symposium on Visual Computing* (pp.524–533). London, UK: Springer, LNCS.

Luan, X., Schwarte, R., Zhang, Z., Xu, Z., Heinol, H.G., Buxbaum, B., Ringbeck, T. and Hess, H. (2001) '3D intelligent sensing based on the PMD technology', *SPIE Proceedings*, Vol. 4540, pp.482–487.

Plaue, M. (2006) *Technical Report: Analysis of PMD Imaging System IWR*. Heidelberg, Germany: University of Heidelberg.

Rapp, H. (2007) *Experimental and Theoretical Investigation of Correlating TOF-Camera Systems IWR*. Heidelberg, Germany: University of Heidelberg.

Schneider, B. (2000) *Der Photomischdetektor zur schnellen 3D-Vermessung für Sicherheitssysteme und zur Informationsübertragung im Automobil Department of Electrical Engineering and Computer Science*. Siegen, Germany: University of Siegen.

Schwarte, R., Xu, Z., Heinol, H., Olk, J. and Buxbaum, B. (1997) 'New optical four-quadrant phase-detector integrated into a photogate array for small and precise', *3D-cameras SPIE Proceedings*, Vol. 3023, pp.119–128.

Strehler, M. (2007) *Messgenauigkeit und Kalibrierung von Laufzeitkameras*. Fraunhofer Institute for Manufacturing Engineering and Automation IPA.

Xu, Z. (1999) *Investigation of 3D-Imaging Systems Based on Modulated Light and Optical RF-Interferometry (ORFI) Zess Forschungsberichte*.

Notes

[1]http://www.pmdtec.com.

[2]http://www.fotowand.de.

[3]When delivered, the camera operates in the Diff Mode and only delivers four channels. A firmware update gives full access to the eight raw channels. Eight channels are also needed for masking of overexposed pixels (see Section, 3.4.2).

[4]Note that the camera delivers a high digit in the raw channels on low exposure.

[5]see Rapp (2007).