Faculty for Physics and Astronomy

University of Heidelberg, Germany

Diploma Thesis
in Physics

submitted by

Holger Rapp
born in Düsseldorf, Germany

September 2007

Experimental and Theoretical Investigation of Correlating TOF-Camera Systems

This diploma thesis has been carried out by **Holger Rapp** at the

Insitut für Wissenschaftliches Rechnen (IWR)

under the supervision of

Prof. Bernd Jähne

Experimentelle und theoretische Untersuchung von korrelierenden TOF-Kamera-Systemen

Diese Arbeit untersucht Time-of-Flight (TOF, Flugzeit basierte) 3D Bildgebungssysteme. Sie enthält die Beschreibung des mathematischen Modells, das nötig ist um die systematischen Fehler und die statistischen Unsicherheiten solcher Kameras vorherzusagen. Um die Fehler experimentell zu bestimmen und um das Modell zu testen wurde für diese Arbeit ein Versuchsstand aufgebaut. Kapitel 2 enthält eine detaillierte Beschreibung dieses Aufbaus.

Drei Kameras wurden experimentell untersucht: die PMD[vision] 19k, die SwissRanger SR-3000 und die Effector O3D Kamera. Alle Kameras haben einen maximalen Messbereich von 7,5 m. Diese Arbeit beschreibt die Experimente, die Ergebnisse und die sich daraus ergebenden Folgen und schließt mit einer ausführlichen Diskussion der Resultate. Möglichkeiten, die systematischen Fehler zu korrigieren werden in der Diskussion präsentiert.

Diese Arbeit brachte drei gemeinsame systematische Fehler zutage: Die periodischen Abweichungen aufgrund der anharmonischen LED-Modulation erzeugt einen periodischen Fehler in der Tiefenmessung von ca. 80-200 mm (je nach Kamera); die Inhomogenität der Pixel verfälscht die Messung um ca. 20 mm und der von der Integrationszeit abhängige konstante Offset liegt zwischen 35 und 100 mm. Die statistischen Schwankungen bei 30% der maximaler Amplitude liegen zwischen 9 mm und 23 mm. Des weiteren wird eine Methode vorgestellt um, soweit technisch möglich, überbelichtete Pixel zu detektieren und zu entfernen.

Mit der vorgeschlagenen Kalibration konnte der absolute systematische Fehler aller gut ausgeleuchteten Pixel der SwissRanger SR-3000 von maximal 300 mm (Standardabweichung: 40,81 mm) auf unter 16 mm (Standardabweichung: 3,16 mm) reduziert werden.

Diese Arbeit entstand in enger Zusammenarbeit mit Industriepartnern im Rahmen des vom BMBF (Bundesministerium für Bildung und Forschung) getragenen Lynkeus-3D Projekts (`http://www.lynkeus-3d.de`). Die Untersuchungen dieser Arbeit führten zu der Entdeckung und Behebung eines Konstruktionsfehlers in einem der Kamera-Systeme.

Experimental and Theoretical Investigation of Correlating TOF-Camera Systems

This thesis investigates Time-of-Flight (TOF) 3D imaging systems. A mathematical model is developed to predict the systematic errors and statistical uncertainties of such cameras. In order to determine the errors experimentally and to test the model, a custom experimental setup has been built for this work. Chapter 2 provides a detailed discussion of this experimental setup.

Three camera systems are investigated experimentally: the PMD[vision] 19k, the SwissRanger SR-3000) and the Effector O3D. All cameras have a maximum measurement range of 7.5 m. This thesis discusses the experiments, the results and the implication of this tests and concludes with a critical discussion of the results. Possible ways to correct the revealed systematic errors is presented in the discussion.

This work reveals three common systematic errors: the variation due to the anharmonic LED modulation provokes a periodic depth error of around 80-200 mm (depending on camera), the inhomogeneity of the pixels accounts for around 20 mm and the constant offset depending on the integration time was found to vary between 35-100 mm. The statistical variances at 30% of the maximum amplitude was found to be between 9 mm and 23 mm. Moreover, a technique to detect and remove overexposed pixels whenever possible is presented.

With the proposed calibration, the absolute systematic error could be reduced in a sample calibration for the SwissRanger SR-3000 from maximal 300 mm (standard deviation: 40.81 mm) to below 16 mm (standard deviation: 3.16 mm) for all well exposed pixels.

This work has been done within the framework of the Lynkeus-3D project (`http://www.lynkeus-3d.de`) supported by the BMBF (Bundesministerium für Bildung und Forschung) and in close cooperation with industry partners. The investigations of this work led to the detection and the mending of a construction error in one of the camera systems.

Contents

List of Figures

List of Tables

Introduction

Over the years image processing tasks have become more involved and interested in the third dimension. Systems that only deliver gray (intensity) images from a scene are decreasing and are replaced by systems delivering more information per frame. This process is most obvious with RGB digital camera systems and only logically advances into systems that acquire depth data – future systems will likely acquire RGBD data frames with a comparable resolution to todays 2D camera systems.

Many principles have been proposed for 3D measurement techniques by optical means. The big categories herein are Time-of-Flight (TOF) measurement ([1]), Triangulation methods ([2] gives an overview) (e.g. stereo vision) and Shape-from-Shading ([3]) (e.g. reflectometry and deflectometry). But none of these techniques have found a broad application in industry. This is due to many reasons, for example high complexity (e.g. stereo vision), small application field (e.g. deflectometry) or a sophisticated and bulky setup (some Shape-From-Shading methods).

This work will discuss a new subclass of the TOF techniques: correlating TOF 3D measurement systems. These systems are a promising new technology combining gray and intensity information in a small camera system with active illumination suitable to use with any standard PC. The technology relies on new semiconductors that correlate (and therefore compare) reference and optical signal directly on chip. This increases precision and decreases size and costs compared to a common system which correlates after recording.

While the correlating TOF on-chip technology is already in use and its earliest scientific introduction was already around 10 years ago ([4]) it remains a field of active research. There is especially a lack of systematic investigations of errors and statistical properties of these systems. This work provides such a systematic study: it is based on a detailed mathematical model of a TOF-Camera, which includes an error propagation model from the measured intensities to the estimated distances. This model's prediction of errors is investigated in detail with a custom made test stand with motor-driven linear tables and with three TOF-Camera systems. The conclusion gives an outlook to open questions and further research.

This work is divided into five main parts. The first chapter details the working principle of correlating TOF and provides the mathematical model needed to follow the further work. Chapter 2 concentrates on the experimental setup that was build in the scope of this thesis; this directly leads to the third chapter which describes the data preprocessing and enhancing steps done for this work. The fourth chapter presents the experimental results, the systematic and statistical errors revealed and investigated in this research and the last chapter contains a discussion of the results and findings and their implications. Possible future research topics and open questions as well as the shortcomings of the investigated systems and the principle itself are also discussed here.

Appendix A lists all experiments made for this work.

Appendix B gives a quick introduction for people who want to work with these types of cameras. This is a practical approach to kick start anyone who wants reliable data quickly or who is unsure if these systems will work for him. It only touches everything briefly, but references to the corresponding sections are given, therefore it is an ideal path to quickly find the information you are interested in. If you read nothing else, make sure to read this!

Chapter 1

Theory

The theoretical basics needed to understand and follow the further progression of this work will be explained and discussed in this chapter.

In section 1.1, the basic working principle of correlating TOF 3D-measurement is summarized in detail. The focus is on modern systems that correlate before recording, therefore directly on chip.

In section 1.2, the exact measurement technique and the underlying mathematics are presented in a congruent mathematical model. Implied systematic errors are presented as well as a discussion of the statistical error propagation.

Section 1.3 concludes this chapter by presenting the mathematical means of calculating the intensity of a flat light source at any given position in space. This equips us with all theoretical knowledge to investigate depth estimation from amplitude decrease.

1.1 Principle of TOF-Systems

Figure 1.1 shows a schematic of a correlating TOF-System which correlates on-chip. Such a system always includes an active modulated light source to illuminate the scene, normally in the infrared spectrum (with wavelengths of around 850 nm). The light doesn't need to be coherent since no interference is needed for the measurement, instead the amplitude of the signal is modulated with the fixed frequency ν. Therefore, cheap and very general light sources can be used. Currently, all systems use LEDs (Light emitting diodes). Future systems will likely also deploy other light sources, e.g. vertical lasers which offer a higher optical power per watt and a more linear and shorter response time to voltage regulation[5].

The light travels from the camera to the scene, gets reflected there and returns back again. It is then recorded in the camera. This measured signal has a different phasing than the departing. The phase shift φ_d is directly associated with the distance d between camera and object according to

$$\varphi_d = \frac{4\pi\nu}{c_{\text{light}}} d, \tag{1.1}$$

with the speed of light in air c_{light}.

The phase shift cannot be measured directly, instead it is done through correlation (see

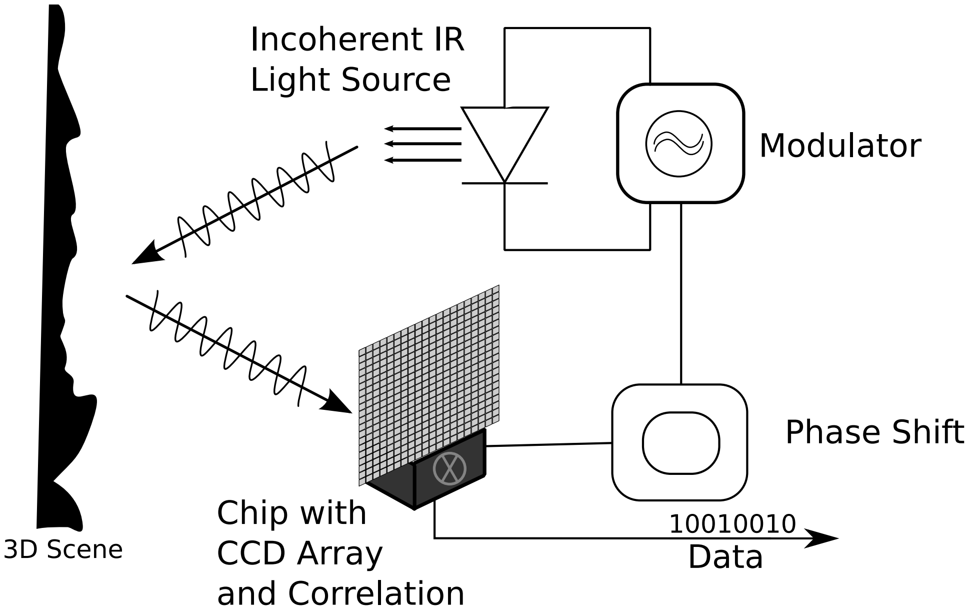


Figure 1.1: Principle of correlating TOF-System

next section) 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 exposure time is equivalent to the integration time in the mathematic expression (see equation (1.2)). A TOF-System therefore directly measures the correlation function (CF) of the emitted and recorded signals and delivers these values as the most fundamental data (raw data).

The CF contains information about the returning optical signal: the constant DC offset c, the modulation amplitude A and the phase shift φ_d, from which the distance d between the camera and the measured object can be computed according to (1.1). The shape of the CF is theoretically known if the exact form of the light modulation is known. But because the CF can only be sampled at a small number of points, the parameters c, A, and φ_d are inferred from a regression on three or more sample points. The more sample points are acquired, the more exact the inferred parameters will be. Each sample point is taken at a different phase position of the CF. This is easily done, since the phase position can easily be changed by shifting one of the correlation functions by a constant phase α_n. By taking at least 3 such sample points the CF can be reconstructed. The process is schematically presented in Fig. 1.2 for 4 sample points. The four top plots show an optical sinus (red) and a reference square (green) wave with four different values for α_n. The blue plot below is the product of the two, the area suggests the integration taking place. The value received through integration is recorded in the lower plot. This gives the sample points through which the known shape of the CF gets fitted (red curve).

This example with two differently shaped curves may seem artificial, but it is closer to reality then calculating with two sinus waves or two square waves: in real systems, the reference wave is usually a square, but due to the non linear reaction of LEDs, the outgoing optical signal (and therefore the recorded incoming) is more sinusoidal. This signal then

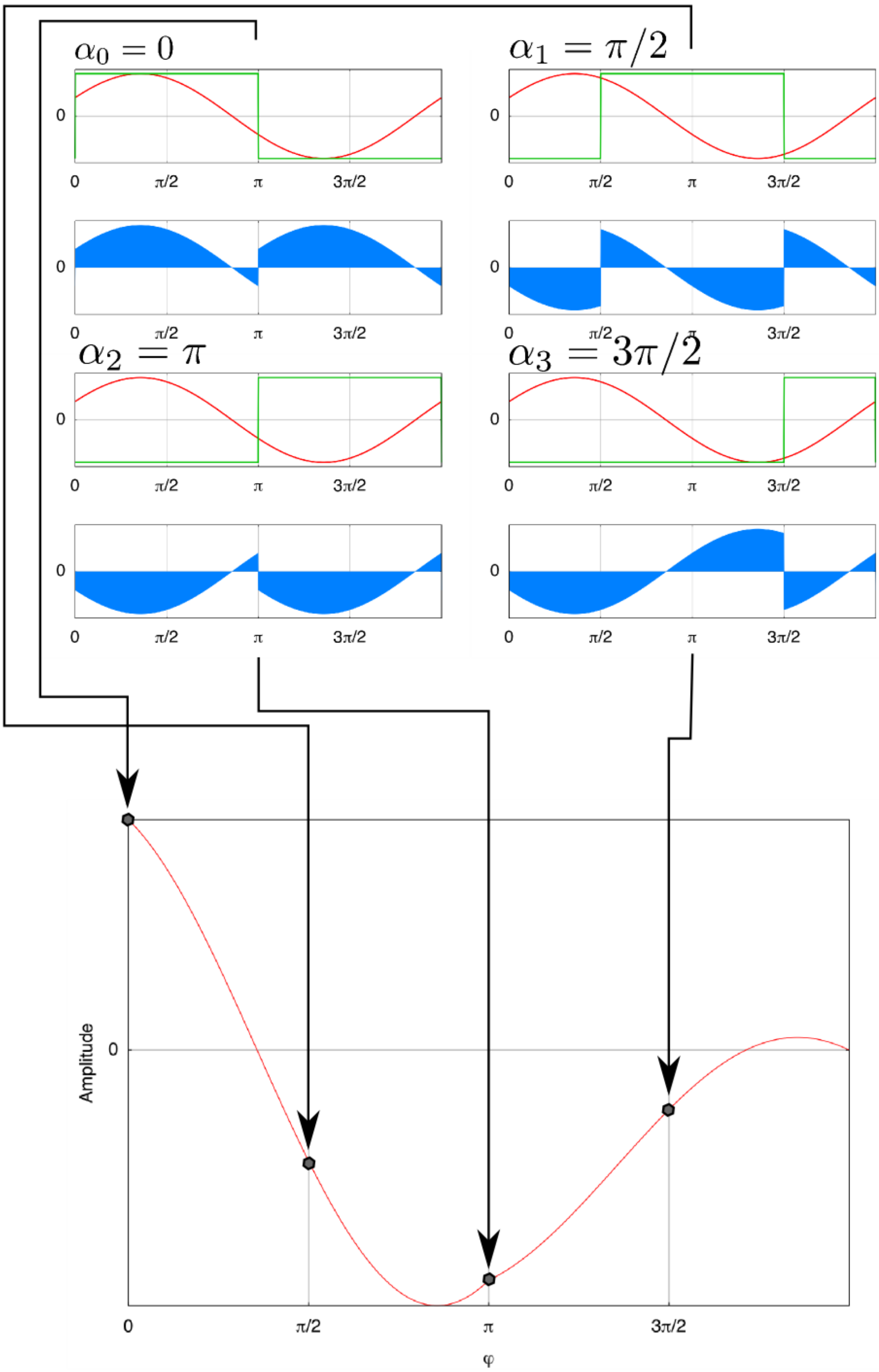


Figure 1.2: Schematic reconstruction of the CF through discrete measurements.

gets correlated with the reference square wave.

1.2 Camera Model

The following camera model gives a brief yet complete overview about the involved mathematics. An exact understanding of the correlation and the principle how it is measured is important to understand the systematic error analysis.

The model approximates the light source as a point light. But to increase the optical power, all systems use more than one point light source, therefore they really illuminate the scene with a more or less homogeneously radiating extended field. The point light approximation is therefore only valid for the far field; how big the distance has to be depends heavily on the geometry of the cameras illumination unit. The proportionality

$$A \propto \frac{1}{d^2}$$

between the amplitude A and the distance d is also only valid for a point light. The exact proportionality is discussed in section 1.3.

A mathematical model of TOF-Camera systems was first published by [6] and [7]. This thesis follows the mathematically equivalent discussion of [8] which is shorter and more flexible and elegant due to the use of complex notation.

1.2.1 Distance Calculation

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

$$\begin{aligned} I_n(\nu, \alpha_n) &= \frac{1}{t'_1 - t'_0} \int_{t'_0}^{t'_1} \mathcal{I}(\nu, t + t_d) \cdot \mathcal{O}(\nu, t + \frac{\alpha_n}{2\pi\nu}) \,\mathrm{d}t && (1.2) \\ &= \frac{1}{t_1 - t_0} \int_{t_0}^{t_1} \mathcal{I}(\nu, t) \cdot \mathcal{O}(\nu, t + \frac{\alpha_n}{2\pi\nu} + t_d) \,\mathrm{d}t, && (1.3) \end{aligned}$$

with the second formula being only a shifted version of the first, therefore $t' - t'_0 = t - t_0$. Since both functions are modulated with the same frequency, it is reasonable to expand both functions into Fourier series:

$$\mathcal{O}(\nu, t) = \sum_{j=-\infty}^{\infty} \mathcal{O}_j e^{ij\omega t}, \qquad I(\nu, t) = \sum_{k=-\infty}^{\infty} \mathcal{I}_k e^{ik\omega t}. \tag{1.4}$$

Here, the angular frequency $\omega = 2\pi\nu$ was introduced to conform to the notation common in physics literature. Equation (1.2) therefore becomes:

$$
\begin{aligned}
I_n(\nu, \alpha_n) &= \frac{1}{t_1 - t_0} \int_{t_0}^{t_1} \left(\sum_{k=-\infty}^{\infty} \mathcal{I}_k e^{ik\omega t} \right) \cdot \left(\sum_{j=-\infty}^{\infty} \mathcal{O}_j e^{ij\omega(t + \frac{\alpha_n}{\omega} + t_d)} \right) \mathrm{d}t && (1.5) \\
&= \sum_{k=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} \mathcal{I}_k \mathcal{O}_j e^{ik\omega(\frac{\alpha_n}{\omega} + t_d)} \frac{1}{t_1 - t_0} \int_{t_0}^{t_1} e^{ij\omega t} e^{ik\omega t} \, \mathrm{d}t. && (1.6)
\end{aligned}
$$

To further simplify this expression, we take a closer look at the last term

$$
\frac{1}{t_1 - t_0} \int_{t_0}^{t_1} e^{ij\omega t} e^{ik\omega t} \, \mathrm{d}t. \tag{1.7}
$$

It trivially calculates to 1 for $j = -k$, for $j \neq -k$ it becomes

$$
\frac{1}{t_1 - t_0} \cdot \frac{e^{i(k+j)\omega t_1} - e^{i(k+j)\omega t_0}}{i(j+k)\omega} \approx 0, \tag{1.8}
$$

because of

$$
(t_1 - t_0)\, \omega = (t_1 - t_0) \cdot 2\pi\nu = 2\pi \frac{(t_1 - t_0)}{T_{\mathrm{mod}}} \gg 1
$$

for typical values of the integration time $(t_1 - t_0) \approx 2 \cdot 10^{-3}\,\mathrm{s}$ and $T_{\mathrm{mod}} \approx 50 \cdot 10^{-9}\,\mathrm{s}$. Therefore we neglect all terms with $j \neq -k$ and equation (1.6) therefore becomes

$$
I_n(\nu, \alpha_n) \approx \sum_{k=-\infty}^{\infty} \mathcal{I}_{-k}\, \mathcal{O}_k\, e^{ik\omega(\frac{\alpha_n}{\omega} + t_d)}. \tag{1.9}
$$

Calculations for Harmonically Modulated Signals

Assuming that both signals have a sinusoidal form,

$$
\begin{aligned}
\mathcal{O}(\nu, t) &= \mathcal{O}_0 + \mathcal{O}_1 \cdot \cos(\omega t + \Phi_0 + \alpha_n) && (1.10) \\
I(\nu, t) &= \mathcal{I}_0 + \mathcal{I}_1 \cdot \cos(\omega t + \Phi_0 + \omega t_d) && (1.11)
\end{aligned}
$$

the correlation frames are easily calculated using trigonometric expressions:

$$
\begin{aligned}
I_n(\nu, \alpha_n) &= \frac{1}{t_1 - t_0} \int_{t_0}^{t_1} \Big(\mathcal{I}_0 + \mathcal{I}_1 \cdot \cos(\omega t + \Phi_0) \Big) \\
&\quad \cdot \Big(\mathcal{O}_0 + \mathcal{O}_1 \cdot \cos(\omega t + \Phi_0 + \alpha_n + \omega t_d) \Big)\, dt && (1.12) \\
&= \frac{1}{t_1 - t_0} \Bigg[\int_{t_0}^{t_1} \mathcal{I}_0 \mathcal{O}_0 \, dt \\
&\quad + \int_{t_0}^{t_1} \mathcal{I}_1 \mathcal{O}_1 \cdot \cos(\omega t + \Phi_0) \cdot \cos(\omega t + \Phi_0 + \alpha_n + \omega t_d)\, dt \\
&\quad + \underbrace{\int_{t_0}^{t_1} \Big(\mathcal{I}_0 \mathcal{O}_1 \cos(\omega t + \Phi_0 + \alpha_n + \omega t_d) + \mathcal{O}_0 \mathcal{I}_1 \cos(\omega t + \Phi_0) \Big)\, dt}_{=0} \Bigg] && (1.13)
\end{aligned}
$$

The last two terms vanish because integration spans over a full period. This is because $t_1 - t_0 = n \cdot T$ with $n \in \mathbb{N}^+$ and $T = 2\pi/\omega$. Using the well known trigonometric expression

$$\cos(\alpha + \beta) = \cos(\alpha)\cos(\beta) + \sin(\alpha)\sin(\beta) \tag{1.14}$$

we get:

$$\begin{aligned} &= \mathcal{I}_0\mathcal{O}_0 + \frac{1}{t_1 - t_0}\int_{t_0}^{t_1} \mathcal{I}_1\mathcal{O}_1 \cdot \\ &\quad \Big(\cos(\omega t)\cos(\Phi_0) - \sin(\omega t)\sin(\Phi_0)\Big) \cdot \\ &\quad \Big(\cos(\omega t)\cos(\Phi_0 + \alpha_n + \omega t_d) - \sin(\omega t)\sin(\Phi_0 + \alpha_n + \omega t_d)\Big)\,\mathrm{d}t. \end{aligned} \tag{1.15}$$

Reusing the argument about the integration interval after factoring out removes all terms containing $\cos(\omega t)$. Further combining the remaining terms with quadratic occurrences yields:

$$\begin{aligned} &= \mathcal{I}_0\mathcal{O}_0 + \frac{1}{t_1 - t_0}\int_{t_0}^{t_1} \mathcal{I}_1\mathcal{O}_1 \cdot \Big(\\ &\quad \cos^2(\omega t)\cos(\Phi_0)\cos(\Phi_0 + \alpha_n + \omega t_d) \\ &\quad + \sin^2(\omega t)\sin(\Phi_0)\sin(\Phi_0 + \alpha_n + \omega t_d)\Big)\,\mathrm{d}t \end{aligned} \tag{1.16}$$

We can now carry out the integration using the fact that

$$\int_a^b k \cdot \sin^2(\omega t)\,\mathrm{d}t = \int_a^b k \cdot \cos^2(\omega t)\,\mathrm{d}t = k\frac{b-a}{2} \tag{1.17}$$

as long as $b - a = n \cdot T$, which is the case here. Applying the trigonometric function (1.14) to the result of the integration finally gives us the solution for the correlation frame.

$$= \underbrace{\mathcal{I}_0\mathcal{O}_0}_{:=c} + \underbrace{\frac{\mathcal{I}_1\mathcal{O}_1}{2}}_{:=A}\Big(\cos(\Phi_0)\cos(\Phi_0 + \alpha_n + \omega t_d) + \sin(\Phi_0)\sin(\Phi_0 + \alpha_n + \omega t_d)\Big) \tag{1.18}$$

$$= c + A \cdot \cos(\alpha_n + \underbrace{\omega t_d}_{:=\varphi_d}) \tag{1.19}$$

$$= \frac{A}{2}\Big(e^{-2\pi i \frac{n}{N}} e^{-i\varphi_d}\Big) + c + \frac{A}{2}\Big(e^{2\pi i \frac{n}{N}} e^{i\varphi_d}\Big) \tag{1.20}$$

Given that N correlation frames are acquired, the offset c, the amplitude A and the phase delay φ_d can be calculated with:

$$A = \frac{2}{N}\Big|\sum_{n=0}^{N-1} I_n e^{-2\pi i \frac{n}{N}}\Big|, \quad \varphi_d = \arg\Big(\sum_{n=0}^{N-1} I_n e^{-2\pi i \frac{n}{N}}\Big), \quad c = \frac{1}{N}\sum_{n=0}^{N-1} I_n \tag{1.21}$$

We'll show next that this solution is optimal in the least square sense, given that $N \geq 3$, otherwise the system would be under-determined (this was first shown by [6]). For a good introduction to least square fitting, see [2], chapter 17.4. We'll use the notation provided there adapted to complex notation.

Writing the correlation frames as derived by equation (1.20) in a matrix notation yields an over-determined linear system of equations:

$$\underbrace{\begin{pmatrix} 1 & 1 & 1 \\ u^1 & \bar{u}^1 & 1 \\ \vdots & \vdots & \vdots \\ u^{N-1} & \bar{u}^{N-1} & 1 \end{pmatrix}}_{:=\,\boldsymbol{M}} \cdot \underbrace{\begin{pmatrix} \frac{A}{2} z \\ \frac{A}{2} \bar{z} \\ c \end{pmatrix}}_{:=\,\boldsymbol{p}} = \underbrace{\begin{pmatrix} I_0 \\ \vdots \\ I_{N-1} \end{pmatrix}}_{:=\,\boldsymbol{d}} \tag{1.22}$$

with $z := e^{i\varphi_d}$ and $u := e^{\frac{2\pi i}{N}}$. To get the general least square solution

$$p_{\text{sol}} = (\boldsymbol{M}^* \boldsymbol{M})^{-1} \boldsymbol{M}^* \cdot \boldsymbol{d}$$

we need to calculate the Moore-Penrose inverse of $\boldsymbol{M}$ which is trivial in this case, because M only contains conveniently spreaded roots of unity, therefore the Moore-Penrose inverse simply is:

$$(\boldsymbol{M}^* \boldsymbol{M})^{-1} \boldsymbol{M}^* = \frac{1}{N} \boldsymbol{M}^* \tag{1.23}$$

and therefore the least square solution becomes

$$p_{\text{sol}} = \frac{1}{N} \boldsymbol{M}^* \boldsymbol{d} = (Az, A\bar{z}, c)^T, \tag{1.24}$$

which is equivalent to the solution stated in equation (1.21), since $\varphi_d = \arg(Az)$ and $A = |Az|$.

Error propagation This paragraph calculates the error propagation for sinusoidal shaped signals for the results presented in the previous paragraph. Gaussian error propagation is used here, another approach is taken by [9] which discusses the statistical error propagation very detailed and for many sample points N.

Here, for reasons of shortness a N of 4 is assumed, meaning 4 sample points are taken by the system[1]. All calculations can theoretically also be carried out with more sample points and for more complex signal forms, but this gets extremely bulky.

The error propagation is done using the well known Gaussian theory. This theory implicitly contains a linear approximation which proves valid though in experiments (see 4.4).

Expanding the results of equations (1.21) for $N = 4$ directly yields the following results:

$$A = \frac{1}{2}\Big|I_0 - I_2 + i(I_3 - I_1)\Big|, \qquad \varphi_d = \arctan\left(\frac{I_3 - I_1}{I_2 - I_0}\right), \qquad c = \frac{1}{4}(I_0 + I_1 + I_2 + I_3) \tag{1.25}$$

Note though that the arctan function must be used with special care to guarantee values in the full unambiguity range of $[0, 2\pi]$[2].

Mathematically, this means to interpret the raw values as a complex vector in the plane, the phase shift φ_d is then the angle of rotation of the vector, the amplitude A the length. We also use this model for following error propagation calculation.

---

[1] Not completely incidentally; this is the number of samples all currently existing systems use

[2] Matlab and NumPy both offer a function called `arctan2` for this purpose

In the next step, we introduce a function f which maps the raw data I to the calculated data vector $(A, \varphi_d, c)^T$. f is thus equivalent to the equations in (1.21), it only uses a matrix notation and the complex interpretation introduced just now:

$$f : \mathbb{R}^4 \to \mathbb{R}^+ \times [0, 2\pi] \times \mathbb{R} \tag{1.26}$$

$$I \to (A, \varphi_d, c)^T. \tag{1.27}$$

Next, we separate f into two functions χ_1 and χ_2 to simplify the calculation of the Jacobian of f. With this separation, f is of the form

$$f = \chi_2 \circ \chi_1 \tag{1.28}$$

with

$$\chi_1 = \begin{pmatrix} \frac{1}{2} & 0 & -\frac{1}{2} & 0 \\ 0 & -\frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \end{pmatrix} \quad \text{and} \quad \chi_2(A, \varphi_d, c) = (\Phi^{-1}(A, \varphi_d), c). \tag{1.29}$$

Herein, Φ is the polar coordinates map $\Phi(A, \varphi_d) = (A\cos(\varphi_d), A\sin(\varphi_d))$.

For the error propagation we need the Jacobian of f:

$$Df(I) = D\chi_2(\chi_1(I)) \cdot D\chi_1(I) = D\chi_2(\chi_1 I) \cdot \chi_1 I. \tag{1.30}$$

Using the following relation

$$D(\Phi^{-1})(\Phi(A, \varphi_d)) = (D\Phi(A, \varphi_d))^{-1} \tag{1.31}$$

it directly calculates to

$$Df(I) = \frac{1}{2} \begin{pmatrix} \cos(\varphi_d) & -\sin(\varphi_d) & -\cos(\varphi_d) & \sin(\varphi_d) \\ -\frac{1}{A}\sin(\varphi_d) & -\frac{1}{A}\cos(\varphi_d) & \frac{1}{A}\sin(\varphi_d) & \frac{1}{A}\cos(\varphi_d) \\ \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \end{pmatrix} \tag{1.32}$$

Supposing now that all correlation frame values acquired by one system have the same variance σ^2 – which is a safe assumption for static scenes since all data is generated through the same process with the same electronics in the same surroundings – we can calculate the relation between σ and the estimated variances of the results through the well known Gaussian error propagation formula:

$$\mathrm{Var}(A, \varphi_d, c) = Df(I)\mathrm{Var}(I)Df(I)^T = Df(I)Df(I)^T\sigma^2 = \mathrm{diag}\left(\frac{1}{2}, \frac{1}{2A^2}, \frac{1}{4}\right)\sigma^2 \tag{1.33}$$

which especially yields the very interesting result that the statistical depth error is directly related to the amplitude. This will become of importance later on, because this relation implies that the amplitude can be used as a confidence information for the depth measurements (see chapters 3 and section 4.3).

Non symmetric Amplitude Modulation

The previous sections provided the theory and the model for TOF-Cameras which are modulated with a sinusoidal signal shape. In general, the cameras are not modulated this way and often also the recorded optical signal shape and the reference signal shape differ. For these cases, the above solutions are no longer valid.

A solution is suggested by [6]; it was first stated in trigonometric form in [10]. Again, [8] delivers a more elegant but equivalent approach. The calculations – which are not restated here, since they do not contain practical information for this work – are oblong and their solution is quite similar to the solutions found for sinusoidal shaped signals above – but they of course contain the higher Fourier modes. It is restated here for completeness in complex form as found in [8]:

$$A_k = \frac{2}{N}\Big|\sum_{n=0}^{N-1} I_n e^{-2\pi i k \frac{n}{N}}\Big|, \quad \varphi_d, k = \arg\Big(\sum_{n=0}^{N-1} I_n e^{-2\pi i k \frac{n}{N}}\Big), \quad c = \frac{1}{N}\sum_{n=0}^{N-1} I_n \tag{1.34}$$

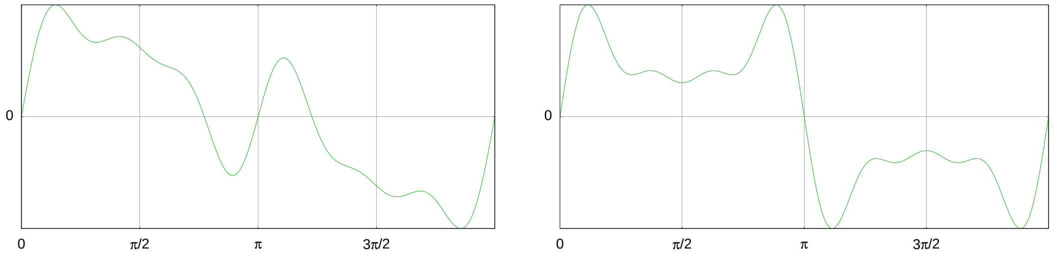
Note that $N \geq 2l + 1$ must be true, otherwise the minimized equation system is underdetermined.

In practice, these equations are not used since they require more sample points and a higher calculation effort. The following thoughts lead to a more simple approach.

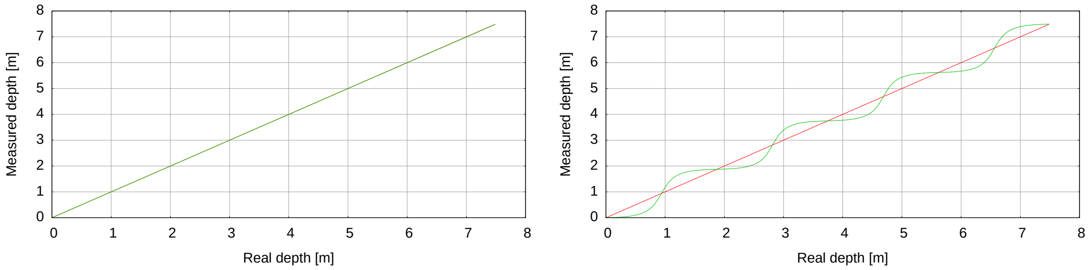
If either the optical or the reference signal is not symmetric (e.g. have odd Fourier harmonics) or if they differ in shape (e.g. one is a rectangle function, the other a sinus – therefore introducing odd Fourier harmonics into the CF) the CF will have a different shape than found in equation (1.19) for sinusoidal signals. This results in periodic systematic errors in the phase calculation when carried out with the equations in (1.21) and therefore in the resulting depth. This effect was mentioned by [7], a mathematical explanation and discussion can be found in [8].

The impact on distance calculations is shown in figure 1.3. The plot shows two signal forms with higher Fourier harmonics and the resulting depth compared to the correct depth when autocorrelating each of this functions with itself and calculating the depth from the result using the equations (1.21) which implicitly rely on sinusoidal shaped signals. Even harmonics do not introduce an error in the depth calculation as can be seen in the two plots on the left. Only even harmonics were used here and the lower plot shows that no systematic error is introduced: the green line lies perfectly on top of the red line.
The plots on the right side are created using a function with odd harmonics. The lower plot shows significant errors in the depth calculation – the measured depth 'wiggles' around the real depth. But it also shows that if the higher modes are sufficiently suppressed – as it is the case in all TOF-Systems and in this example – the bijective nature between calculated depth and real depth is not broken. Therefore, a simple lookup table can be used to correct measured data in field. This solves the problem very easily, all mathematical approaches would need more sample points and more involved and expensive calculus.

Explanation of the wiggling error This paragraph explains the origin of the wiggling error for $N = 4$ sample points. We use figure 1.4 as example. In this figure, two distinct optical functions $\mathcal{I}$ (one with 3rd fourier mode and one with 5th) are investigated. The top



(a) $f(x) = \sin(x) + \frac{3}{4}\sin(2x) + \frac{1}{2}\sin(4x) + \frac{1}{4}\sin(6x)$ (b) $f(x) = \sin(x) + \frac{3}{4}\sin(3x) + \frac{1}{2}\sin(5x) + \frac{1}{4}\sin(7x)$



(c) (d)

Figure 1.3: Signal forms (a,b) and resulting depth (c,d) after autocorrelation compared to correct depth for two correlation functions with higher Fourier modes. (a,c) are even harmonics only, (b,d) are odd harmonics only. A modulation frequency of ν =20 MHz was assumed.

plot shows the predicted depth error for this optical function and a square wave reference function which is a more realistic assumption for real-life reference signals (see section 1.1). The plot in the middle shows the correlation function for the investigated case (green) and for the theoretically assumed case (red). The plot below shows the difference of these two CFs.

The theory assumes a sinusoidal CF, thus for other signals the calculations in (1.25) are only correct for those phase shift values φ_d for which the following relation holds:

$$\varphi_d = \arctan\left(\frac{I_3^T - I_1^T}{I_2^T - I_0^T}\right) \overset{!}{=} \arctan\left(\frac{I_3^R - I_1^R}{I_2^R - I_0^R}\right) \tag{1.35}$$

with I_n^T being the theoretical n-th correlation frame and I_n^R being the real. This relation is trivially fulfilled if $I_n^R = I_n^T \; \forall \; n \in \{0, 1, 2, 3\}$ as it is the case for $\varphi_d = \{0, \pi/2, \pi, 3\pi/2\}$. These cases are visualized as green dots in figure 1.4.

The relation is also correct for the non trivial case that the errors of the real correlation frames are in the same relation to each other as the denominator and the numerator in the arctan with the theoretical values above. To clarify this, we take a look at the expression for the real correlation frames. For this we introduce the error (or difference) δI_n^R between theoretical and real correlation function:

$$\delta I_n^R = I_n^R - I_n^T \quad \forall n \in 0, 1, 2, 3. \tag{1.36}$$

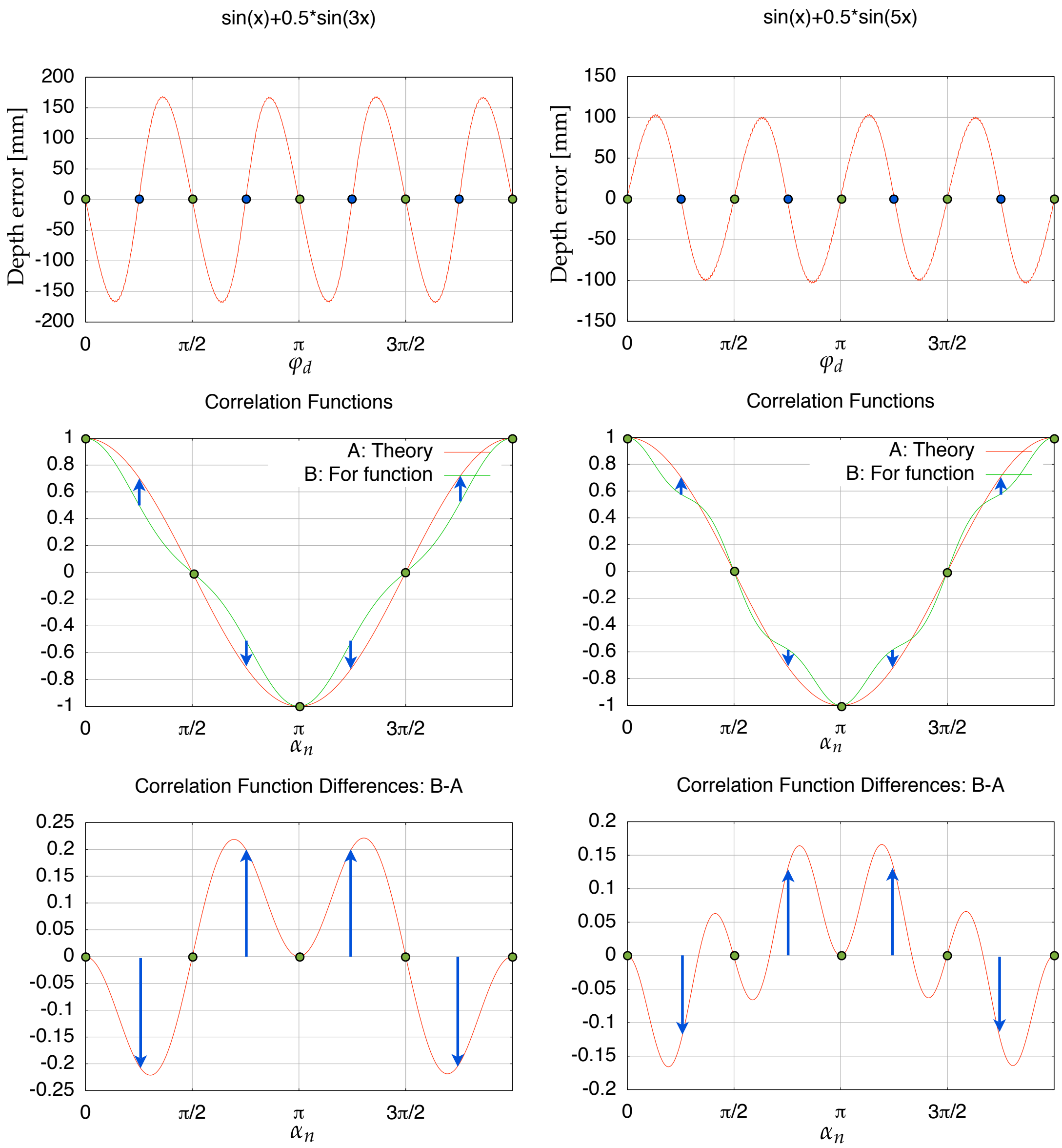


Figure 1.4: Origin of wiggling error. Top plot shows the depth error for the whole unambiguity range for φ_d, middle plot shows correlation function for $\varphi_d = 0$, the CF for the given optical function is green and the CF assumed by theory is red. Bottom plot shows the difference between correlation functions. The colored points and arrows are responsible for the zero-crossing points of the depth error functions.

This error is plotted in the bottom plots in figure 1.4. With it, the equation 1.35 reads:

$$\arctan\left(\frac{I_3^T - I_1^T}{I_2^T - I_0^T}\right) \overset{!}{=} \arctan\left(\frac{(I_3^T + \delta I_3^R) - (I_1^T + \delta I_1^R)}{(I_2^T + \delta I_2^R) - (I_0^T + \delta I_0^R)}\right) := \arctan\left(\frac{a + \delta a}{b + \delta b}\right) \tag{1.37}$$

where $a := I_3^T - I_1^T$, $b := I_2^T - I_0^T$, $\delta a = \delta I_3^R - \delta I_1^R$ and $\delta b = \delta I_2^R - \delta I_0^R$. For the relation

$$\arctan\left(\frac{a}{b}\right) \overset{!}{=} \arctan\left(\frac{a + \delta a}{b + \delta b}\right) \tag{1.38}$$

to be true, either $\delta a = \delta b = 0$ must be true (this is the trivial case already discussed above) or

$$\frac{\delta a}{\delta b} = \frac{a}{b} \tag{1.39}$$

This is the case for the zero-crossing points marked with blue dots in figure 1.4. Note that these cases are still special because here $\frac{a}{b} = \frac{\delta a}{\delta b} = -1 \in \mathbb{Z}$. For higher fourier modes also rational fractions appear.

More quantitative examples for the wiggling error are shown in figure 1.5, also here a modulation frequency of 20 MHz was assumed. The optical signal $\mathcal{I}$ is given below each plot, the reference signal is again assumed to be a square wave. Then the predicted measured-depth signal is calculated from this correlation function and the corresponding depth error is plotted.

The top two plots show that the 3rd and the 5th fourier mode are responsible for a wiggling error with a wavelength of approximately 1.9 m, the two plots below show that the 7th and 9th fourier induce a wiggling with a wavelength of approximately 0.8 m. This continues for higher modes, always two fourier modes of the optical signal correspond to one wavelength of the wiggling.

The four lower plots show depth error for combinations of different modes (left) and the irrelevance of even fourier modes (right).

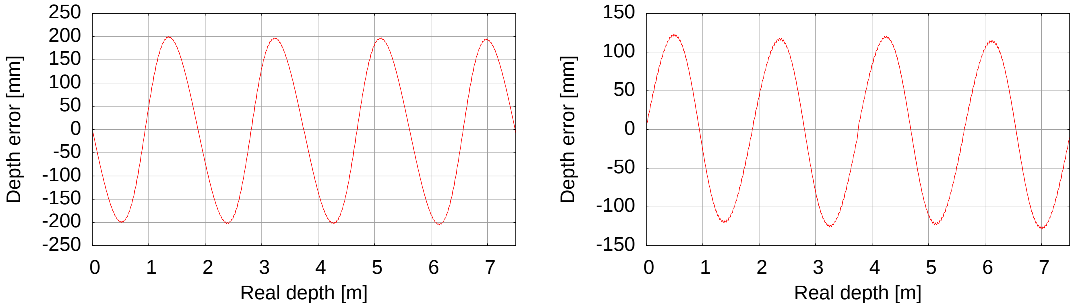
1.3 Amplitude Decrease with Depth

Given a point light in space, it is a well known fact that the intensity of the radiated light drops proportional to the inverse square of the distance:

$$I(\boldsymbol{r}) = \frac{a}{r^2} \tag{1.40}$$

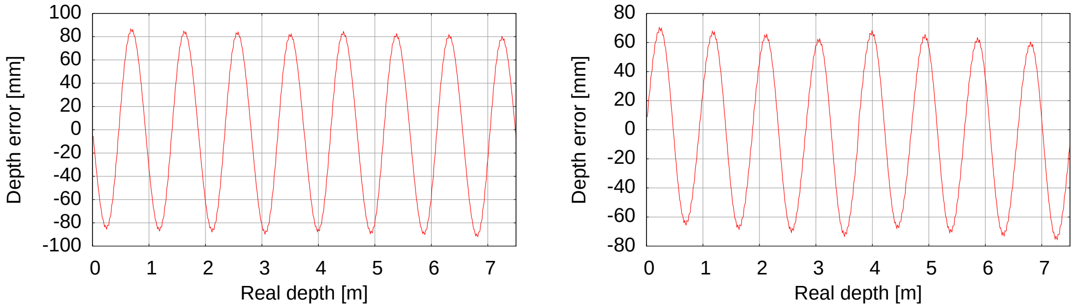
with the proportionality constant a. This information could theoretically also be used to predict the depth from the returned amplitude.

The light sources in current systems are no point lights but can rather be modeled as homogeneously radiating areas of rectangular dimensions. To predict the depth from amplitude, we need to compute the theoretical intensity distribution for these systems more precisely. For the calculation, we set the origin in the center of the camera lens. The illumination intensity $I(\boldsymbol{r})$ at the point $\boldsymbol{r} = (x, y, z)$ (axes as in figure 3.5) can then easily be calculated through integration over the dimensions of the illumination unit; if the lens



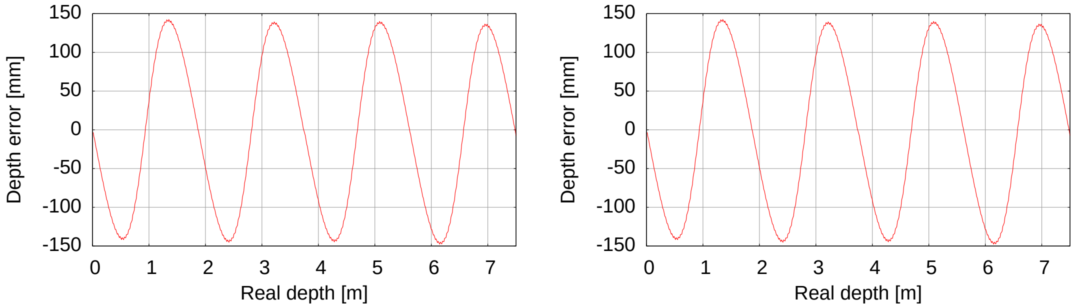
sin(x)+0.5*sin(3x)

sin(x)+0.5*sin(5x)



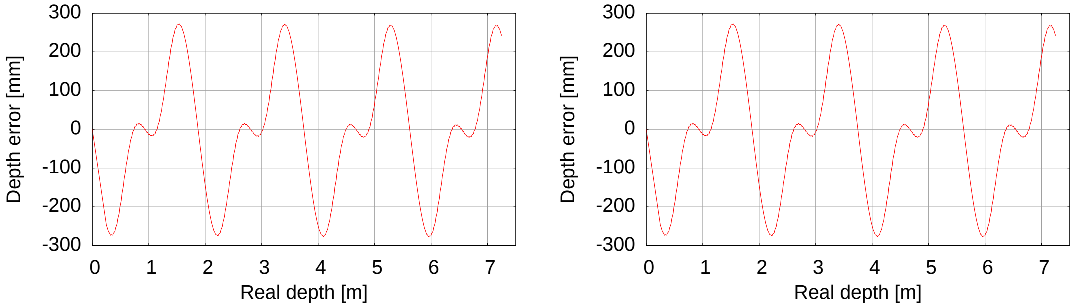
sin(x)+0.5*sin(7x)

sin(x)+0.5*sin(9x)



sin(x)+0.5*sin(3x)+0.25*sin(5x)

sin(x)+0.5*sin(3x)+0.25*sin(5x)+
sin(2x)+sin(4x)+sin(6x)



sin(x)+0.5*sin(3x)+0.75*sin(7x)

sin(x)+0.5*sin(3x)+0.75*sin(7x)+
sin(2x)+sin(4x)+sin(6x)

Figure 1.5: Wiggling examples for the given optical signal function $\mathcal{I}$, the reference function $\mathcal{O}$ was a square wave. A modulation frequency of 20 MHz was assumed.

is inside the illumination unit (like with the SR-3000 and the O3D), the area covered by it is omitted in the integration:

$$I(\boldsymbol{r}) = a \cdot \int_{-l_x}^{l_x} \int_{-l_y}^{l_y} \frac{dx'\,dy'}{(\boldsymbol{r'} - \boldsymbol{r})^2} \tag{1.41}$$

$$= a \cdot \int_{-l_x}^{l_x} \int_{-l_y}^{l_y} \frac{dx'\,dy'}{(x'-x)^2 + (y'-y)^2 + z^2} \tag{1.42}$$

One of the two integrals can be carried out analytically. This yields the result:

$$I(\boldsymbol{r}) = a \cdot \int_{-l_y}^{l_y} \left[\frac{\arctan\left(\frac{x-x'}{(y'-y)^2+z^2}\right)}{\sqrt{(y'-y)^2+z^2}} \right]_{-l_x}^{l_x} dy' \tag{1.43}$$

The second integral must be carried out numerically with the exact dimensions of the light source known.

Chapter 2

Experimental Setup

This chapter describes the experimental setup built and used for this thesis. This comprises a detailed overview of the third party products used and the assemblage of the experiment as well as the discussion of the unique features of the experimental rig (section 2.2). Also, an overview of the software written for each part of the hardware is given.

The chapter continues with a presentation of benefits and shortcomings of the custom made targets in section 2.3.

It concludes with the discussion of the three correlating TOF-Systems investigated in this work – the PMD[vision] 19k, the SwissRanger SR-3000 and the Effector O3D – in section 2.4. The similarities and differences of the camera systems are presented and the unique special features of each camera are discussed.

2.1 A Word on Software

There are currently no standards defined for software to communicate with TOF-Camera systems. But for ease of use with different systems, a common interface was needed and therefore was developed.

The software used as a host for data acquisition and basic processing was Heurisko®[1]by AEON Verlag & Studio. Most software was written as plugin-DLLs for this comprehensive image processing software solution.

As rapid prototyping environment and more flexible solution for automated measurement the free programming language Python[2] together with the SCIPy (Scientific Python) and PIL (Python Image Library) modules was chosen. This approach proved to be very useful. Python surpasses the Heurisko solution in speed and agility but lacks a collection of common image processing algorithms.

Future work using the experimental setup after this thesis will likely need a lot of image processing, therefore the interface plugins for Heurisko are considered a much more important result of this work then the Python code.

---

[1] http://www.heurisko.de

[2] http://www.python.org

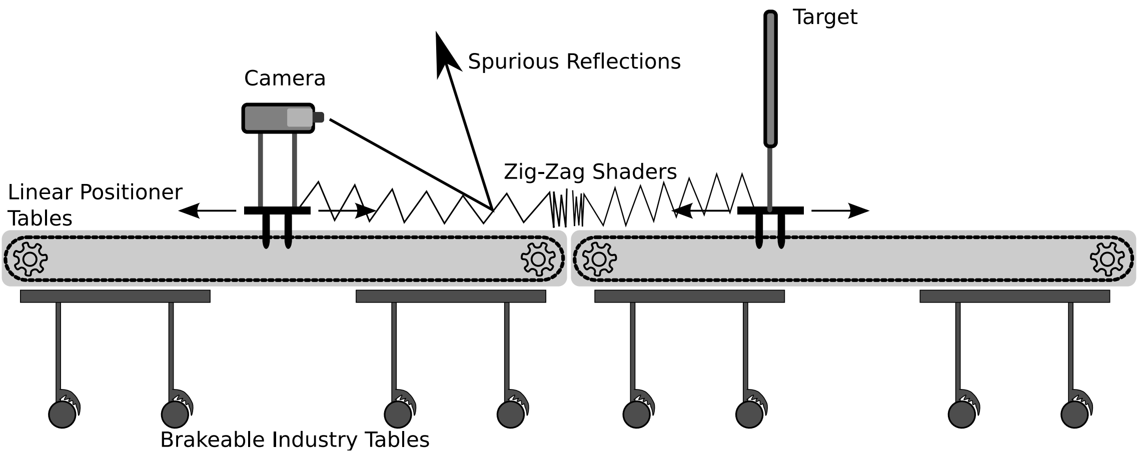


Figure 2.1: Experimental setup

All written software with documentation can be found on the CD-ROM annexed to this work.

2.2 Overview

This section gives a brief overview of the main parts of the experimental setup, each part is discussed in more detail in its own sections below.

A schematic of the experimental setup can be seen in figure 2.1, a photo which contains more details is shown in figure 2.2. The setup consists of two linear positioner tables mounted on top of standard industry tables with brakeable rolls. The first positioner carries the camera, the second the target to be acquired. Between camera and target a zig-zag shader has been installed to avoid spurious reflections of the IR light from the tables.

Not visible in the schematic but in the photo is the cable bearing which guarantees a save movement of the cables while the tables are in motion. Also not visible is the raceway. This device can be seen in the photo as two black-white-yellow rails with ramps. It proved necessary as a guide for the zig-zag shader, because it was always crumbled through friction at high movement speeds.

The whole setup is controlled by a standard PC which positions the tables, acquires the frames from the camera and processes and displays the data. This process is fully automated, so complex and long measurements can be run without any human interaction.

The room has been held dark for all measurements. Furthermore, 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. This provisions ensure the best possible data the system can deliver is acquired – without any errors introduced through the setup or the surroundings. These parameters will likely have an impact though in field and further investigation is needed before deployment of any of the systems.

Figure 2.2: Photo of the experimental setup

2.2.1 Linear Positioner Tables

The main components of the experiment are the two linear positioner table's LCB060 from Parker[3]. These tables are computer controlled (through a special control box connected to a serial port of a standard PC with a custom cable) linear positioner devices with an absolute positioning error of $< 1\,\mathrm{mm}$ each. The positioning error is one order of magnitude smaller then the precision of the measurements of the cameras; the tables position information can therefore be considered as ground truth. The tables have a range of 3 m, can speed up with a precise acceleration and reach an end speed of at most $3\,\mathrm{m/s}$ (depends on bearing).

The tables are mounted on standard industry tables with brakeable rolls. This allows them to be flexibly connected and rearranged to each other which is especially interesting for measurements of moving targets (see figure 2.3 for an example). The tables also make it easy to move the whole setup to another location – for example outdoors for bright daylight measurements.

If the tables are aligned as shown in figure 2.1 – as they were in all experiments for this work – the setup allows for a sub-millimetre precise depth positioning in the range of $d_0 \approx 0.2\,\mathrm{m} < d < 6\,\mathrm{m} + d_0$.

---

[3] http://www.parker-eme.com

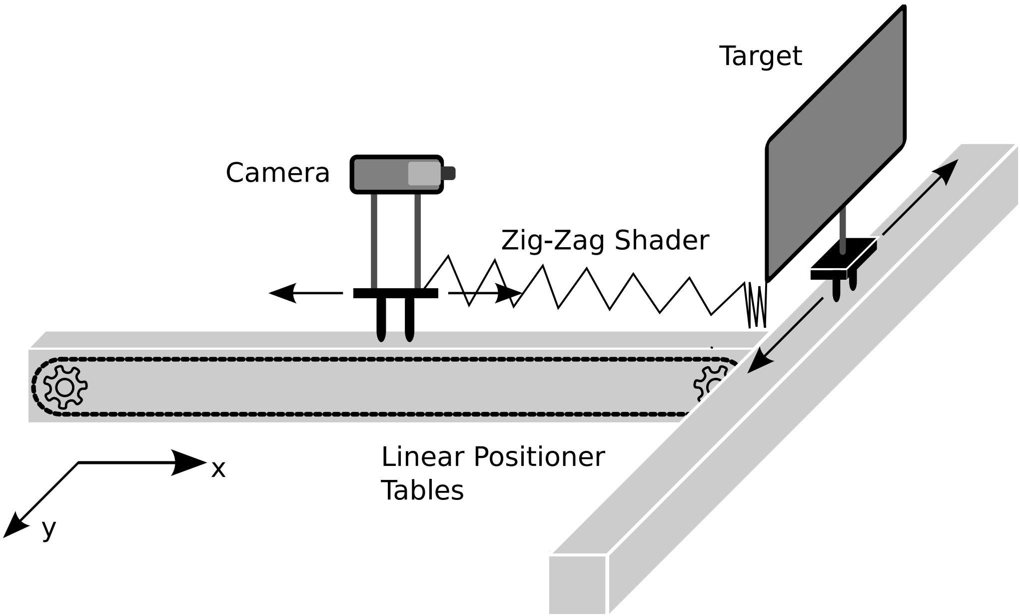


Figure 2.3: Example of flexibility: experimental setup for cross movement acquisitions

Software The tables come with a Win32 executable to configure the basic parameters like torque, gear transmission rate and frame of reference. To control the movement of the tables, a command string in ASCII described in the handbook must be written to the serial port the device is connected to. To simplify this work, the author has written a Win32 DLL to encapsulate the complex communication with the device (see compax3com.dll and compax3com.h). Based on this DLL, a Python class (compax3.py) and a Heurisko acquisition device (compax3.dll) with similar interface have been developed.

2.2.2 The Zig-Zag Shader

The zig-zag shader has been developed and deployed because of the problem visualized in the top picture of figure 2.4. Here a beam of the active illumination of the camera falls on the positioner table, gets reflected and falls on the target and from there back to the camera. The problem is that the camera cannot detect if the beam took the direct way to the target and back or if it was reflected. Therefore the data of this pixel gets deteriorated and we can't detect the depth correctly.

The problem is easily solved through a zig-zag shader (bottom picture) for two reasons. The main reason is that beams from the camera that hit the shader would either reflect back into the room and not into the recorder unit (see bottom picture) and therefore not change our target measurement. But even if they would fall back into the camera, the biased pixel would lay on the shader, not on the target. Therefore our target measurements stay reliable. The second reason is that the zig-zag shader is made of black photo pasteboard which absorbs most light. Therefore less light is reflected back than without the shader.

The zig-zag shader was specially made for this experimental setup and basically works like an accordion: when it is stretched out, the zig-zags get flatter and longer, when the

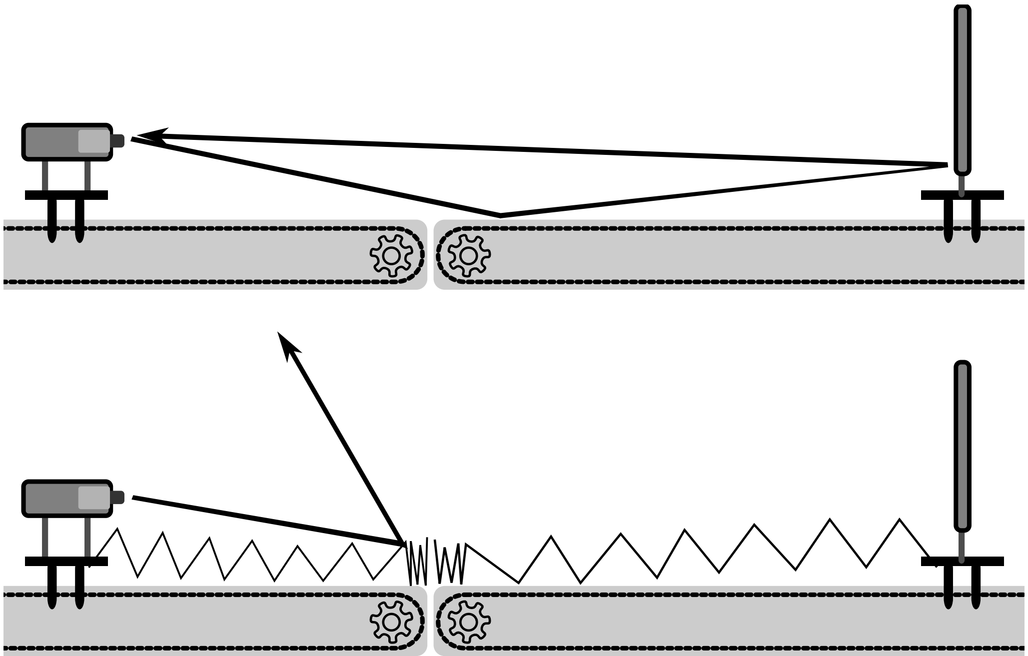


Figure 2.4: Top image shows the reflection problem, bottom picture shows the solution through the zig-zag shader

table moves more to the end, the zig-zag gets compressed and folds neatly into a compact block.
The zig-zag shader is attached to the moving part of the tables and one of its ends. With this layout, the flexibility of the setup is not lost since each table has its own shader.
To assure that the shader stays close to the tables and not rear up they have been pinched with a fishing line. This simple solution proved durable without affecting mobility.

Figure 2.5 shows a close up shot of the zig-zag shader on the camera table. This is at the end of the table were the shader is attached.

2.2.3 The Cable Bearing

Two cable bearings in chain form have been installed to ensure that all cables leading to the camera devices or future active targets (e.g. light sources) are unharmed by the high movement speed possible. The bearings have been specially mounted to the movable parts of the linear positioner tables as can be seen in figure 2.6.

The cable bearing is completely flat if the table is on distance 0 mm, when it moves to a distance of 3000 mm it rolls up like a chain and takes the cables with it without harming them. Its drawback is that if the chain is rolled in, it has a significant diameter of about 15 cm. Therefore, it was necessary to lift the zig-zag shader by this diameter to make sure that the bearings and the shader would not interfere with each other. The raceway therefore also serves as placeholder.

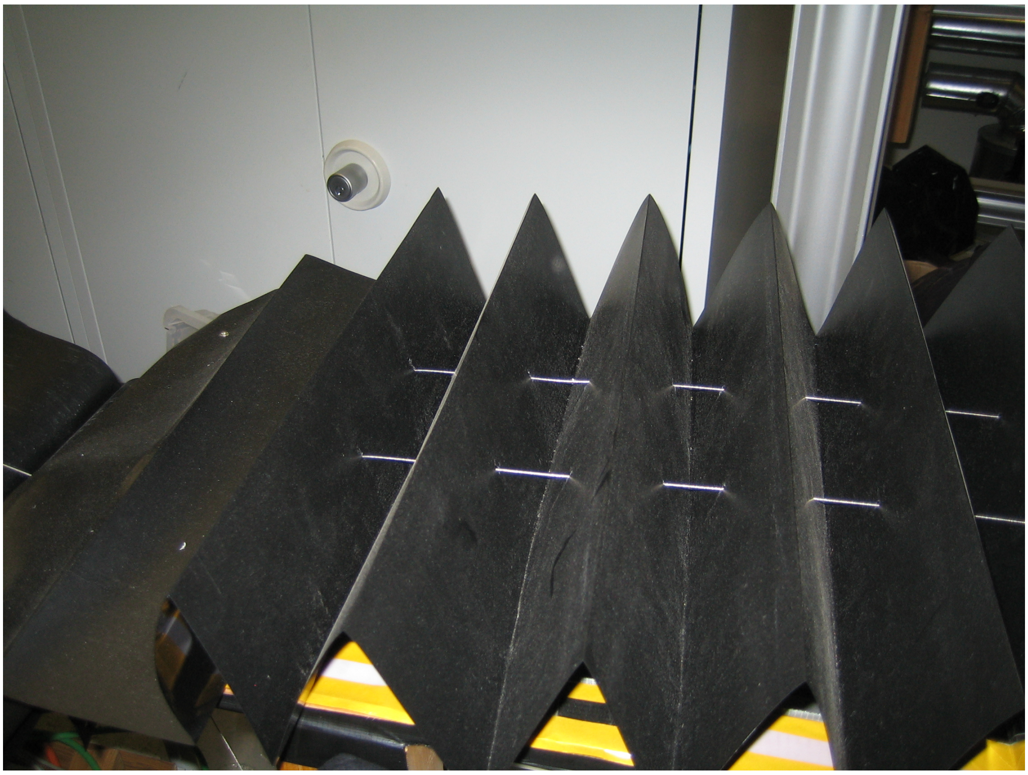
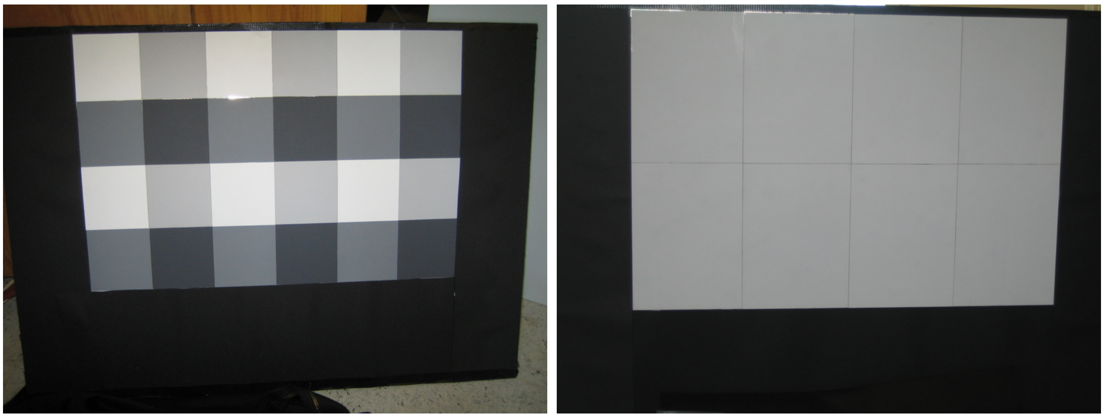


Figure 2.5: Photo of zig-zag shader of the camera table in close-up

Figure 2.6: Photo of connection of cable bearing to movable part of camera table in close-up



(a) Checkerboard Target (b) High-Reflectivity Target

Figure 2.7: Targets used for data acquisition

2.2.4 The Raceway

The raceway is a glider rail for the zig-zag shader which is directly sliding on it when the table moves. It was needed to ensure that the folding and unfolding of the zig-zag shader didn't lead to crumbling due to high friction. It also assures that the zig-zag shader doesn't touch and interfere with the cable bearing: the raceway is high enough so that the cable bearing always stays below it even when it reaches its full diameter.

The raceway consists of two long pieces of wood next to the positioner tables on a constant height. The wood is pasted with Teflon to ensure a low friction for the zig-zag shader.

The raceway can be seen in figure 2.2. It consists of the two black-white-yellow rails next to the linear tables starting with a slope at the right side of the picture.

2.3 Targets

The targets used for the experiments can be seen in figure 2.7.

They were custom-built for this work 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[4]. These cards provide a defined diffuse reflectivity while the specular reflectivity is very low. Thus, they are nearly perfect Lambertian radiators. The high-reflectivity target uses Photo-Cards with 84 % reflectivity, the checkerboard consists of 90 × 90 mm squares with reflectivities of 12.5 %, 25 %, 50 % and 84 % in a regular pattern.

The targets are pasted on an aluminum board and are connected to a frame permanently attached to the target linear positioner table with screws. This allows for a change of target in a matter of a few minutes while guaranteeing the reproducible positioning and a high stability of the targets with respect to the camera.

---

[4] `http://www.fotowand.de`

Figure 2.8: The PMD[vision] 19k

The checkerboard target revealed a big problem of the systems in test phase: some of the squares which are not perfectly aligned leave some silver aluminum looking through. This silver line now reflects the IR light of the camera nearly completely. This basically counts for another reflectivity of approximatively 100 % which often makes overexposed areas in the measurement data that was not intended.

2.4 The Camera Systems

This paragraph will give detailed information about the camera systems investigated in this work and their special properties.

2.4.1 The PMD[vision] 19k

The PMD[vision] 19k camera system by PMDTechnologies GmbH[5] is the oldest model used in this investigation, still it is interesting to investigate, since it already contains all key technology for a TOF-System. A picture can be seen in figure 2.8.

The camera uses LEDs with a wavelength of 870 nm and a total optical power of around 3W. The LEDs are mounted in two arrays, one on either side of the camera which is not optimal, since it introduces near field errors due to TOF differences between the left and the right array that are hard to correct. For all experiments, the modulation frequency was kept at the default value of $\nu = 20\,\mathrm{MHz}$ resulting in an unambiguous depth range of $d_{\mathrm{max}} = \frac{c}{2\cdot\nu} = 7.5\,\mathrm{m}$. The camera acquires samples at four phase shifts for each CF, taking two samples on each measurement (one with α_{n} and one with $\alpha_{\mathrm{n}} + \pi/2$). This redundant information is used to correct inhomogeneities in the chip (see section 3.2, [11] and [12] contain more details), but contains more valuable information for the correction of overexposed pixels.

The correlation inside the camera is performed using a CMOS based optical semiconductor called Photonic Mixer Device (PMD). This technique increases speed and decreases cost

---

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

and noise of the system. The camera has a resolution of 160 × 120 pixels with a frame rate of 5 to 12 fps. The data is digitized with 12 bit and delivered to the host through a firewire interface.

The camera directly delivers the raw channels. The depth information, the amplitude and the intensity information – this means all data – can be gained from the raw channels through the calculations introduced in chapter 1. But the camera delivers all those information also in other data channels, so the user doesn't need to bother with the calculations.

The intensity channel is not usable, because of the high inhomogeneities in the gates (see section 3.2).

Since the camera has no suppression of background illumination (SBI), it is not suitable for measurements in bright daylight which consists to a high part of IR light interfering with the measurements.

Software A Python module has been developed by the author and a Heurisko module has been developed by B. Jähne, M. Schmidt and the author. Both modules define a software interface used for the other cameras as well.

2.4.2 The SwissRanger SR-3000

The SwissRanger SR-3000 (Fig. 2.9) by CSEM is quite similar to the PMD[vision] 19k. It also uses a modulation frequency of $\nu = 20\,\text{MHz}$ and therefore measures the same depth range, its resolution is slightly higher but comparable (176 × 144 pixels) and it has a little higher frame rate of approximatively 20 fps.
Its optical power is lower but its gain is a bit higher – still it can't measure as far as the 19k can. It uses IR-LEDs at a wavelength of $\lambda = 850\,\text{nm}$.

Little is known about the on-chip correlation. Investigation of the delivered data strongly suggests that correlation frames are acquired at four points in time for each image. But if a special semiconductor like the PMD is used is not publicly known but probable.

The camera delivers 4 raw channels[6], a depth and an intensity channel. The intensity channel is the equivalence of the amplitude channel of the 19k, therefore the SwissRanger SR-3000 doesn't deliver real gray value presumably for the same reason as the 19k.

The raw data does not contain as much information as the raw data of the 19k which makes it impossible to correct overexposed pixels with this device (see 3.3 for details).

The SR-3000 has SBI, but experiences of various members of the Lynkeus-3D project showed that it is not good enough for bright daylight measurement. Therefore the SR-3000 is suitable for the same class of problems as the 19k and is therefore its direct rival.

Software A Python module has been developed by the author and a Heurisko module has been developed by B. Jähne, M. Schmidt and the author. The modules follow the interface defined by the 19k modules.

---

[6] as an undocumented feature of the firmware, acquiring the raw channels could be considered as hack.

Figure 2.9: The SwissRanger SR-3000, photo from manufacturer's web page

2.4.3 The Effector O3D

The Effector O3D from IFM Electronic[7] is the most recent advancement in the TOF-Camera sector. For this work, IFM kindly provided a prototype, the product will be introduced into the market in 2008, likely under another name. A picture can be seen in figure 2.10.

This camera plays in another category than the other two cameras. While it uses a similar technique ($\nu = 20\,\text{MHz}$, $\lambda = 850\,\text{nm}$ with IR-LEDs, PMD for correlation) it is not aimed at image processing tasks. Its interface and software is more focused on sensorical aspects. This also shows in its low resolution of 64×50 pixels. Its optical power is comparable to the SR-3000, but its gain is much better. On the other side it can only measure with integration times up to 5 ms which makes it unsuitable for ranges higher than 2-4 m (depending on the surroundings). Its frame rate is highly dependent on the integration time. This is because the LEDs are the main heat source in the device which has no active cooling. Therefore the software adjusts the frame rate and only turns the LEDs on when a frame is acquired, the time between frames is used to cool down the device. Therefore the frame rate is directly controlled by the internal heat management of the camera.

The camera was developed using a chip from PMDTechnologies, therefore the correlation on chip is done exactly as in the 19k camera. The camera delivers depth information and amplitude (called intensity in their software). The gray information is not delivered, for the same reasons as with the other cameras. The raw data contains as much information as the 19k's, therefore overexposed pixels can easily be detected.

The camera also contains some more intelligence to enhance the data. In its standard acquisition mode, it acquires two pictures per delivered frame, one with a high integration

---

[7] http://www.ifm-electronic.com

Figure 2.10: The Effector O3D

time and one with a low one. The pictures are then combined taking only well exposured pixels from each picture. This effectively increases the dynamic range of the camera through means of software. This function was turned off for all experiments because systematic errors in relation to integration time were investigated and after combining the pixels, it is impossible to determine which pixel was from which integration time.

Software A Python module has been developed by the author. No Heurisko module is available at this time due to the high programming effort needed. Partners in the Lynkeus project are working on a software DLL which should ease the work to write such a module.

2.4.4 The Systems in Direct Comparison

The investigated systems are compared side-by-side in table 2.1.

	PMD[vision] 19k	SwissRanger SR-3000	Effector O3D
Resolution	160 x 120	176 x 144	64 x 50
Pixel Dimensions	40 x 40 μm	40 x 40 μm	100 x 100 μm
Focal Length	12.0 mm	8.0 mm	8.6 mm
Light Source	2 LED arrays	1 LED array	1 LED array
Modulation Wavelength	870 nm	850 nm	850 nm
Optical Power	≈ 3 W	< 1 W	≈ 1 W
Modulation Frequency	20 MHz	20 MHz	20 MHz
FPS	max. 15	max. 40	between 1 and 100
Connection	FireWire, Ethernet	USB 2	Ethernet
Dimensions	220 x 210 x 55 mm	60 x 50 x 65 mm	55 x 45 x 85 mm

Table 2.1: Features of the investigated TOF-Camera Systems

Chapter 3

Data Preprocessing

3.1 Improving the Data through Averaging over Time

Noise reduction is an essential operation before systematic errors can really be investigated. A proper, simple and well understood way to reduce noise is to average image frames over time. With depth data from correlating TOF-Systems though, naive averaging is not likely to reduce noise and improve the data, especially for low exposed pixels.

When the depth data is first calculated from the raw channels and then simply summed over and divided, the confidence information from the amplitude is completely ignored. This means that wrong data is weighted equally as correct data. This will bias the averaged frame, especially in areas of low exposure where wrong data appears more frequently.

A simple solution to this problem is to take not the mean but the median of the data. This gives better results, because the median is especially suited to ignore outliers in data sets. This is a general approach and doesn't take into account the additional data available from this systems. Therefore a better solution is to enhance the data through weighting the depth frames with the amplitude information. After correcting overexposed pixels (see section 3.3), well exposed pixels are statistically more likely to contain valid depth information (see 1.2.1) and should therefore get a stronger weight.

To reduce calculation complexity, this work uses another equivalent technique. Instead of averaging after calculation, we take the mean of the raw channels before calculating depth and amplitude. This is the same as frames with a low amplitude only add a smaller raw value to each frame, therefore depth information with a low amplitude are weighted less due to the nature of the calculation process. This technique has the advantage that it can already be used while acquiring data and by averaging the raw channels, all calculated data is also already averaged. This reduces the number of calculations from N to 1 for each calculated data channel.

3.2 Correcting Gate Inhomogeneities

The Effector O3D and the PMD[vision] 19k internally do not only take four but eight measurement points. These samples are always taken in pairs at four points in time (in two different physical gates per pixel), therefore they theoretically contain redundant information and do not improve the depth calculation at all. Instead this data is used to correct inhomogeneities.

When gate A measures the correlation frame for phase shift $\alpha_n^\mathrm{A} = 0$ the gate B in the same pixel measures the value shifted by half a period, thus for phase shift α_n^B. This means theoretically:

$$I^\mathrm{A}_{\alpha_\mathrm{n}} = I^\mathrm{B}_{\alpha_\mathrm{n}+\frac{\pi}{2}} := I_\mathrm{n} \tag{3.1}$$

with I^B and I^A being the correlation frame measured on the first (gate A) and the second (gate B) gate of the pixel respectively. But due to technical reasons, each gate in each pixel has a constant offset error $\delta_\mathrm{A,B}$ and a different gain $g_\mathrm{A,B}$:

$$I^\mathrm{A}_{\alpha_\mathrm{n}} = g_\mathrm{A} \cdot \hat{I}^\mathrm{A,B}_{\alpha_\mathrm{n}} + \delta_\mathrm{A} \tag{3.2}$$

and likewise for B.

For the calculations in equation (1.21) only the differences of two correlation frames are needed. To calculate for example $d := I_0 - I_2$ there are two possibilities:

$$d = (g_\mathrm{A} \cdot \hat{I}^\mathrm{A}_0 + \delta_\mathrm{A}) - (g_\mathrm{B} \cdot \hat{I}^\mathrm{B}_0 + \delta_\mathrm{B}) \tag{3.3}$$

$$d = (g_\mathrm{B} \cdot \hat{I}^\mathrm{B}_\pi + \delta_\mathrm{B}) - (g_\mathrm{A} \cdot \hat{I}^\mathrm{A}_\pi + \delta_\mathrm{A}). \tag{3.4}$$

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

$$d = \frac{1}{2}(g_\mathrm{A} \cdot \hat{I}^\mathrm{A}_0 - g_\mathrm{B} \cdot \hat{I}^\mathrm{B}_0) - (g_\mathrm{A} \cdot \hat{I}^\mathrm{A}_\pi - g_\mathrm{B} \cdot \hat{I}^\mathrm{B}_\pi). \tag{3.5}$$

The gain differences are not corrected by this, they account for constant offsets in the amplitude data. But the gain differences are not as big, therefore this much smaller error is easier to tolerate.

This correction was done for all experiments of this work with the O3D camera and the 19k.

Note that the SR-3000 camera most likely also contains a correction for the pixel inhomogeneities, but it is not publicly known how it works.

3.3 Removing Overexposed Pixels

Removing overexposed pixels is normally not a difficult task in image processing. Simple 2D cameras deliver a fixed maximum value on overexposure, some more sophisticated systems are able to set a flag to mark overexposure. But due to the more involved principle of on-chip correlation, TOF-Systems make it often much more difficult to detect overexposed pixels, sometimes it is even impossible.

Theoretically, when a pixel is overexposed, all sample points should deliver the same value thus all raw data channels should be equal. This also results in a zero amplitude which is – given that the amplitude is used as a confidence value – optimal: Overexposed pixels are already sorted out by the confidence information.

Unfortunately, due to the correlation, there are exposure intervals where some sample points are overexposed and some are not. This results in a non-zero amplitude and also in faulty depth data but it's not obvious in all cases that the data is bad.

This section demonstrates methods to detect overexposed pixels for the three cameras described in section 2.4. The problem is discussed using a set of data points with increasing depth from the target. With this experiment, overexposure can easily be identified in the data: if the camera is near to the target, the pixels are overexposed. If the camera moves away, overexposure ends at a definite position. It is important to detect overexposure without this help for in-field-deployment, therefore wherever possible, a suggestion is made how to identify overexposed pixels.

3.3.1 PMD[vision] 19k

The 19k camera offers in its delivery status four raw channels. Since these are calculated internally with eight raw values and since those eight values contain more information, it is advisable to patch the camera to deliver all eight raw values. A firmware patch is available from PMDTec.

Also, using only four channels the camera behaves very similar to the SwissRanger SR-3000. A safe detection of overexposure is therefore also not possible (see section 3.3.3).

With eight channels though, the problem is easily solvable. Figure 3.1 shows three plots of the same experiment. The distance to the target increases in x direction. At the top the eight raw channels are plotted for all positions. The red lines are gate A and the green gate B at four sample points each. We can see that there is an area were all sample points are obviously overexposed because the red and green lines lay exactly on top of each other. But when we compare the amplitude plot in the middle, we see that the data is still faulty up to the black line. Not until then starts the amplitude to drop at a range as is physically expected. The unnatural rise and fall can also not be explained by near field effects but is completely evoked by overexposure. This can be seen when plotting other integration times: the amplitude shows a similar behavior at the end of overexposed areas.

The plot at the bottom also shows that the data is faulty up to this point, only then does the wiggling start (see section 4.3.1). Left from the line, the depth data is completely random.

Since in the eight-channels-mode all channels rise monotonically with exposure[1] we can easily find a criteria for overexposed pixels: A pixel is overexposed for the PMD[vision] 19k when at least one raw data channel is below 2500.

3.3.2 Effector O3D

The Effector O3D is the easiest of the three cameras when it comes to detect overexposed pixels. The Figure 3.2 shows again raw data, amplitude and depth error for one integration time. It is most obvious in the top plot that for overexposed data the channels either deliver a value close to the minimum (in this case around 514) or a value close to the maximum of $2^{16} - 1 = 65535$.

---

[1] Note that the camera delivers a maximum value for zero exposure.

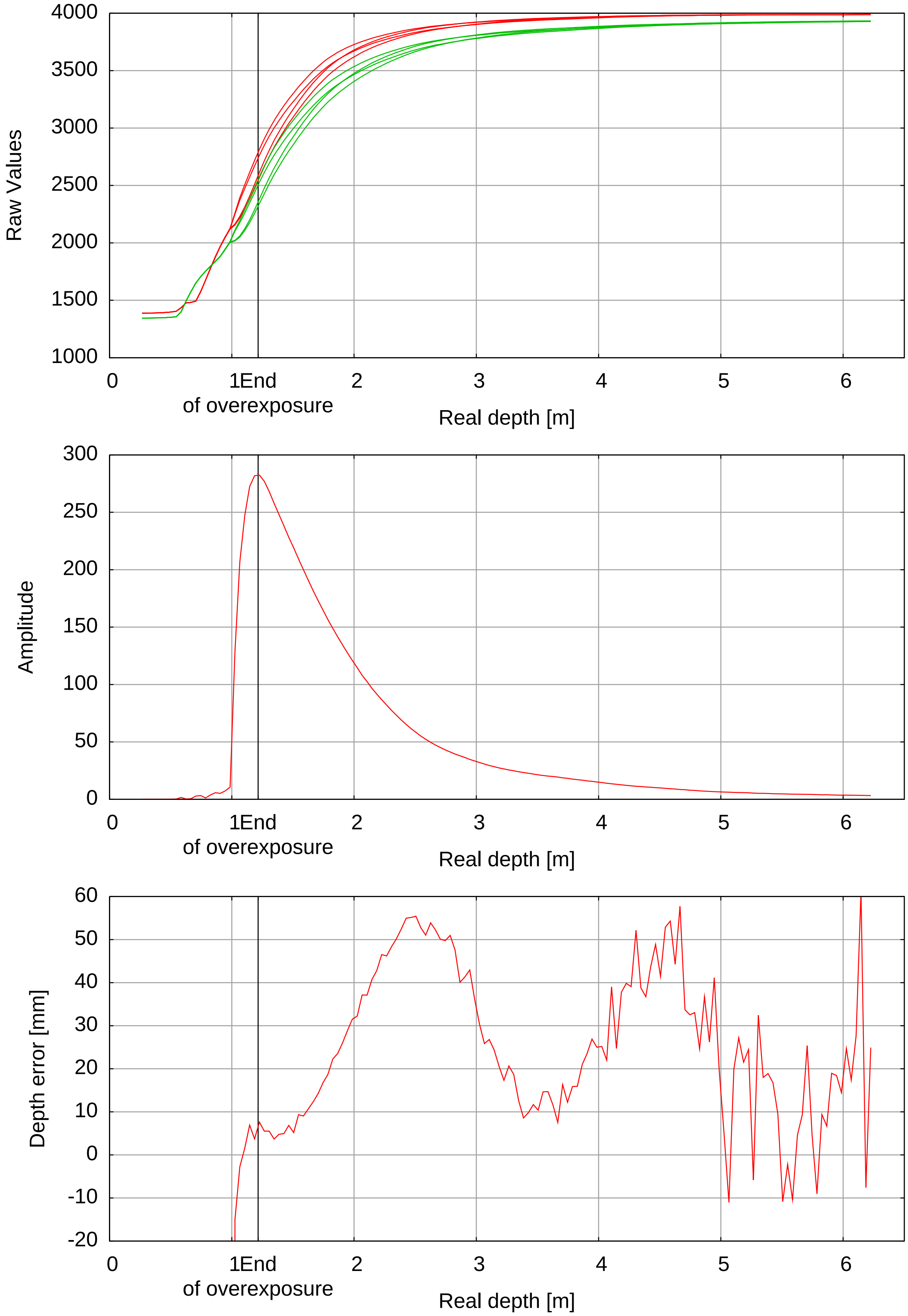


Figure 3.1: Overexposure of the PMD[vision] 19k displayed with the high-reflectivity target and an integration time of 8.0 ms

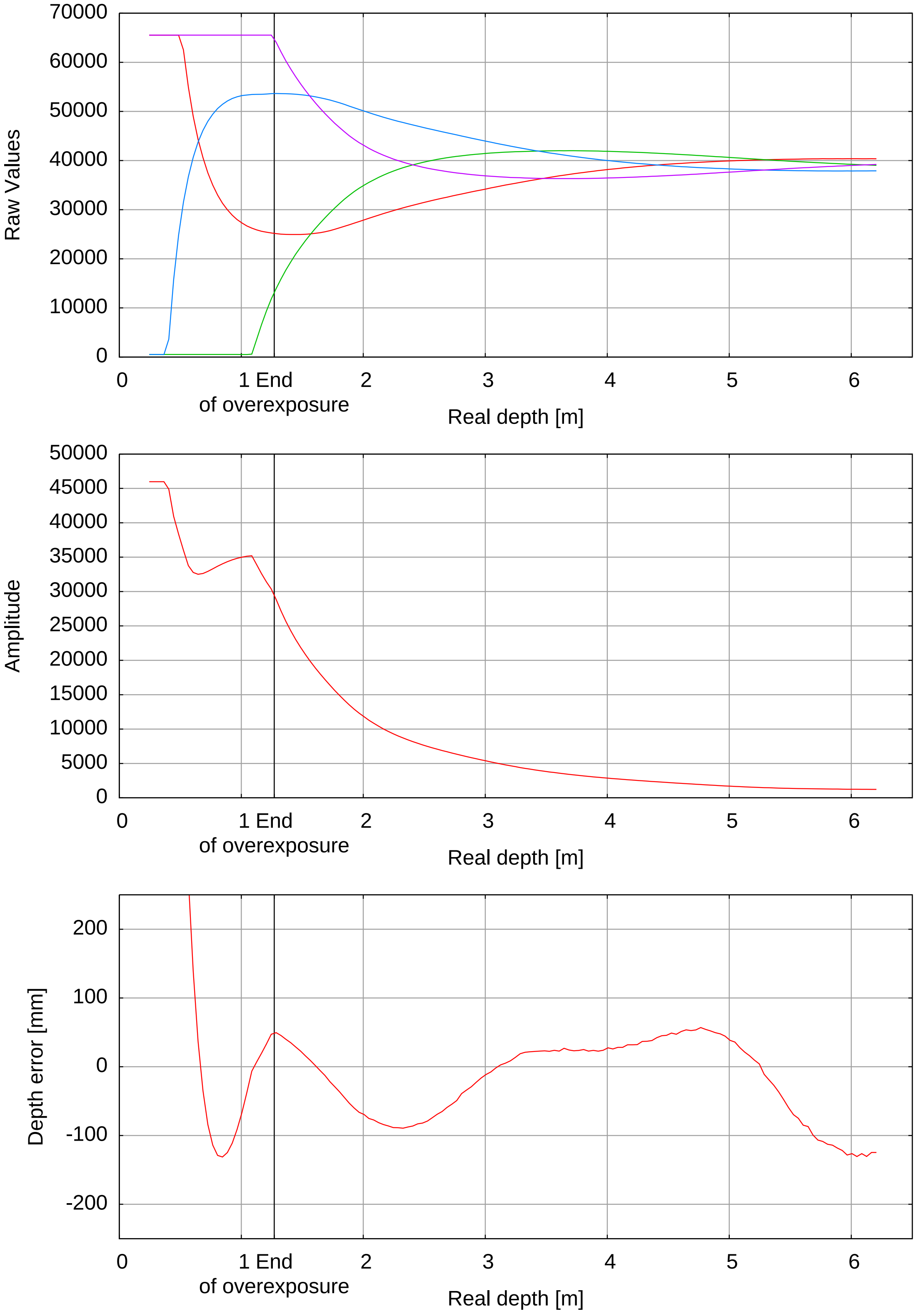


Figure 3.2: Overexposure of the Effector O3D displayed with the high-reflectivity target and an integration time of 5.0 ms

A safe bet is to say a pixel is overexposed for the Effector O3D when at least one raw data channel is above 65530 or when at least one raw channel is below 520.

Very interesting is also the area shortly before the end of overexposure. In this, only the purple raw channel is still overexposed. The amplitude and depth error plots already show an almost correct behavior. But the steepness of the curves changes again when the purple raw channel also leaves overexposure. This shows that strict rules for detecting overexposed pixels must be available; analyzing the proximity of a measured point is not sufficient to guarantee well exposed pixels.

3.3.3 SwissRanger SR-3000

The Figure 3.3 shows the analogous plot for the SwissRanger SR-3000. This proves to be a much harder problem to solve. Even though there is an area where one or more raw channels lay on top of each other in the overexposed area, they separate quite soon and show a random behavior. The amplitude plot doesn't help either because the amplitude takes every possible value between zero and its maximum in the overexposed area.

It is therefore not possible to detect overexposed pixels with the SR-3000. In further sections and data discussions, the overexposed pixels have been filtered manually for this camera.

3.4 Gauging the Phase Shift

Each camera system has an internal offset error in the calculation of the phase shift. This offset error is introduced by electronic run time delays inside the camera and is different for each individual camera system.
This error can be mathematically expressed as a zero point shift on the unit circle: while theoretically a depth of zero should be a phase shift of zero, all cameras will deliver another value φ_d^0. But starting from this value, the phase shift will monotonously increase with the depth until $d = d_{\max} = \frac{c}{2\nu}$ in which the phase will have the value $\varphi_d^0 + 2\pi$.

This phase offset must be corrected by gauging with one known distance point in the calculation, otherwise, the phase calculation in (1.21) will wrap too early which leads to a jump in the depth calculation. This is visualized in figure 3.4. By subtracting the zero point φ_d^0 before taking the modulo 2π this problem can easily be avoided. The gauging doesn't need to be exact, even if it is only approximate, the wrap will occur at much higher real distance so it will impact only if the full range of $d_{\max}$ is to be measured. The remaining offset error can be corrected later while correcting the pixel offset errors (see section 4.3.3).

Note that all cameras implicitly subtract this offset in the data delivered as depth information. PMDTec also gives a hint how to correct for this offset after calculation of depth[2].

---

[2] The miniSDK function `pmdGetOffset` is addressed, but caution: it delivered a wrong value with our 19k

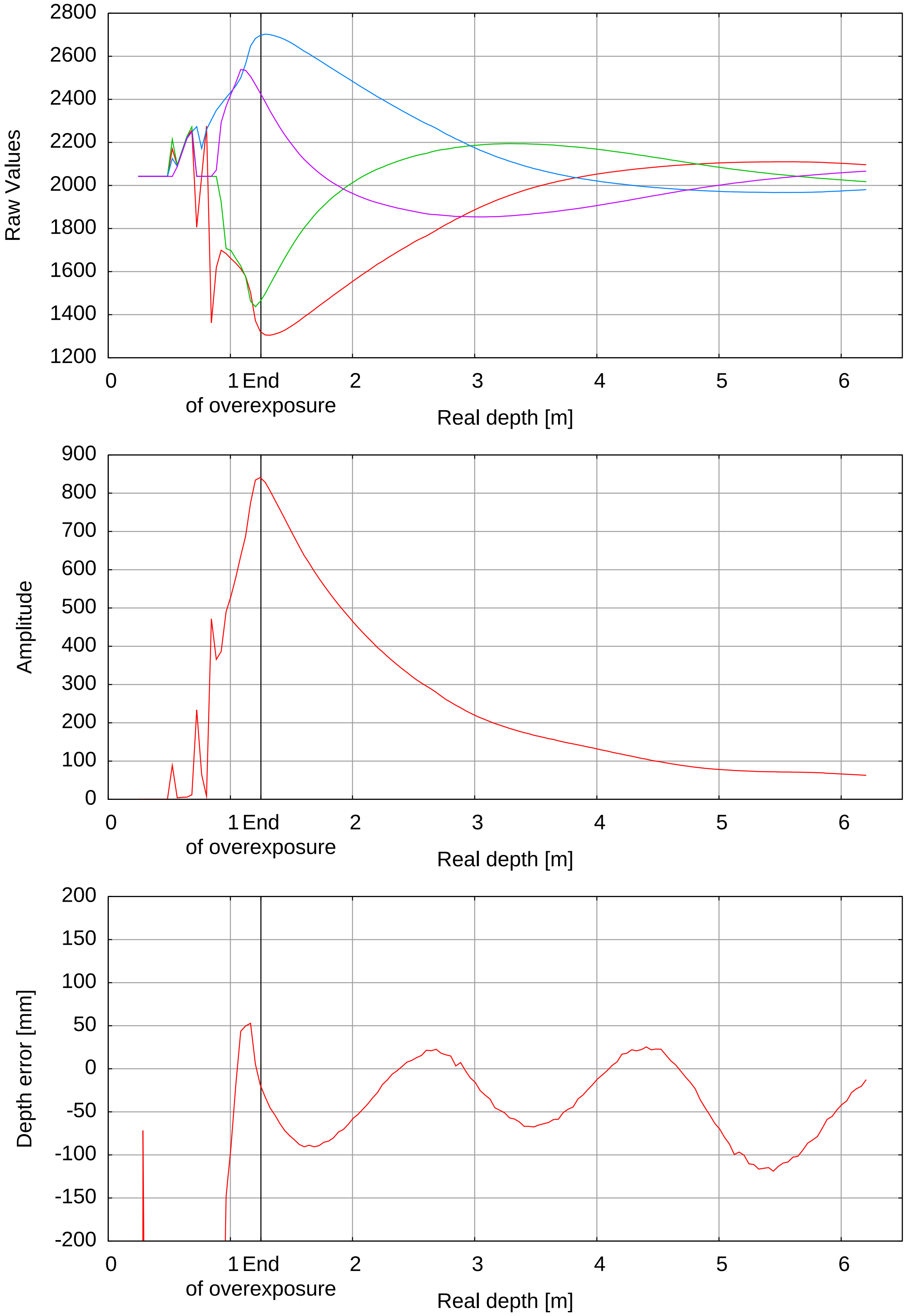


Figure 3.3: Overexposure of the SwissRanger SR-3000 displayed with the high-reflectivity target and an integration time of 30.0 ms

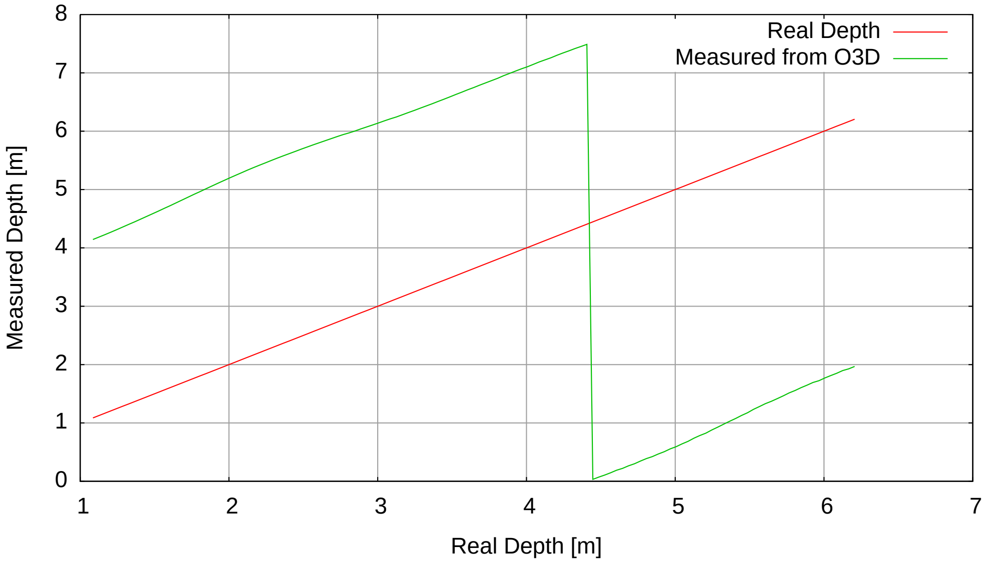


Figure 3.4: Example of a depth calculation without correcting the phase offset φ_d^0 first.

3.5 Radial to Orthogonal Distance

The TOF-Cameras - like most 3D measurement techniques - are limited to acquire the radial distance from the lens to the scene. But in most cases, the orthogonal distance is the desired information.

Using Figure 3.5 as geometric reference, we'll derive a simple transfer mechanism based on the model of a pinhole camera. Note that this is only approximatively correct since this model completely ignores lens distortion. But with the low resolution of today's systems, this distortion will not have much impact, therefore we do not lose much information. As resolution increases, this model will need adaption and refinement.

Using the nomenclature from Figure 3.5, the camera measures the radial distance d_r. The searched orthogonal distance d_o can simply be acquired by taking the cosine of β:

$$d_o = d_r \cdot \cos(\beta) \tag{3.6}$$

The unknown angle β can easily be derived by using the location of the pixel (n_x, n_y) the orange beam is detected in. This pixel is the one in which the object will appear later in the digital image. Given that the distance a between two pixels is known, the physical position can be calculated to

$$x_p = a(n_x - w/2) \tag{3.7}$$
$$y_p = a(n_y - h/2) \tag{3.8}$$

with w and h being the number of pixels of the camera in x and y direction respectively. This correction is needed because the x_p and y_p must be measured from the optical axes, but cameras usually have their origin in the upper left pixel.

With the location x_p and y_p and the focal distance d_p of the camera known, we can now

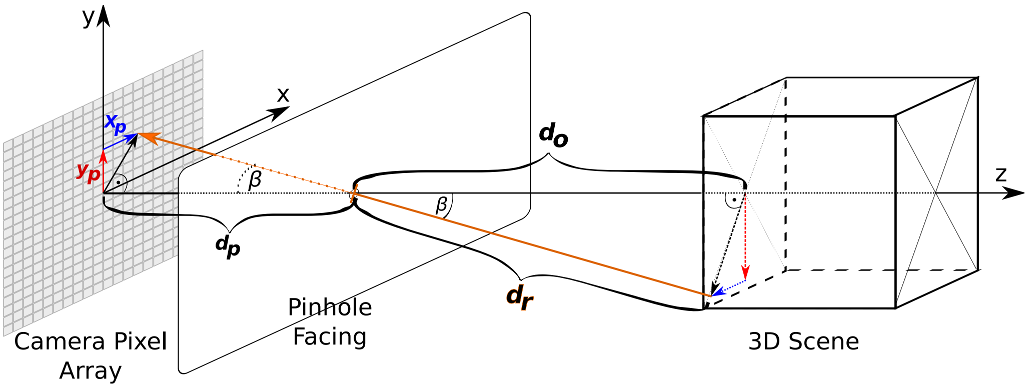


Figure 3.5: Pinhole model for transferring radial to orthogonal distance

easily calculate β with

$$\tan(\beta) = \frac{\sqrt{x_p^2 + y_p^2}}{d_p} \tag{3.9}$$

Therefore, the complete transfer from radial to orthogonal distance is given by

$$d_o = d_r \cdot \cos\left(\arctan\left(\frac{\sqrt{x_p^2 + y_p^2}}{d_p}\right)\right) \tag{3.10}$$

Chapter 4

Results

The previous chapters have introduced the theory, the experimental setup and the data preprocessing steps. This chapter will now present experimental results vested through this knowledge.

The chapter first discusses the amplitude falloff with depth (section 4.1). Then the LED signal shapes of the three devices is presented in section 4.2. The major part of this chapter will describe the systematic errors of the cameras in section 4.3. The last part will present the experimental results of error propagation (section 4.4).

The data for this chapter was acquired with the experimental setup described in chapter 2, all measurements were made at night to ensure no distortion through other IR light sources. The measured depth range was $0.3 \pm 0.1 < d < 6.3 \pm 0.1$ m, the exact start and stop values depend on the length of the individual camera. The system acquired 150 frames each 4 cm and took a mean from these values using the techniques described in section 3.1. If not stated otherwise, the high-reflectivity target was used (see 2.3).

4.1 Amplitude Falloff

Since the amplitude information of the returning signal is readily available and since the theory behind amplitude falloff with depth is widely understood, it is interesting to investigate this also experimentally. This section provides such an experimental overview. The theoretical values in this section are derived according to section 1.3.

The data for the 19k and for the SR-3000 camera have been acquired with the linear positioner tables without zig-zag shader and a photo diode with a suitable IR-Filter in front. The experimental data for the O3D has been acquired with a fast IR-photo diode connected to an oscilloscope and for a much shorter range. This data is not suitable to discuss amplitude falloff since it enfolds only a very small distance range. Therefore, the corresponding figure 4.3 is only included for completeness.

The plot for the 19k can be seen in figure 4.1. Below 60 cm the shadowing effect – the camera case impedes the light to fall directly onto the target due to the unfavorable fixing of the LED arrays – which is not taken into consideration in the theory amounts for a considerable bias. After that, the theory is in good agreement with the experimental data,

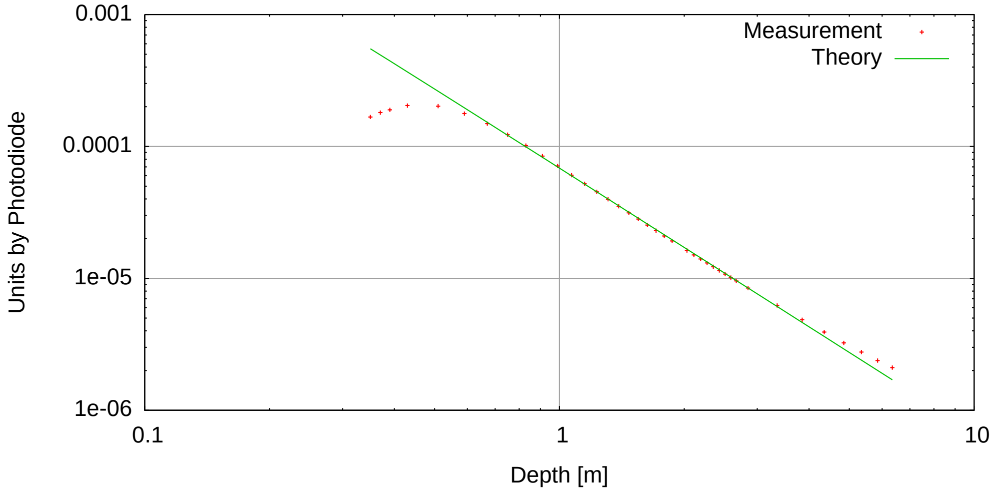


Figure 4.1: PMD[vision] 19k amplitude falloff

only with higher distance, more light intensity is measured than is predicted from theory. This is likely because of reflection from the (unshaded) tables and the surroundings.

The plot for the SR-3000 camera is quite similar and can be seen in figure 4.2. The first values are also lower here, but this can't be amounted to shadowing effects on the camera, because the LEDs are more cleverly mounted here. The effect here is more likely a near field type: the LED arrays are no homogeneous light emitting area as the theory assumes for such small distances. With increasing depth, the values are getting more exact, but the higher values are much worse than the higher values for the 19k. This is because the SR-3000 emits its light into a greater solid angle thus more light is reflected by the surroundings and therefore more light gets reflected back into the camera.

Generally the experiment supports the thesis that depth from amplitude information would support and maybe enhance the depth information from correlation. But further research should be made in this area, especially with scenes containing different reflectivities. This work will not further pursue the data enhancement possibilities through this method.

4.2 LED Signal Shape

Figure 4.4 shows the measured LED signals for all three investigated camera models. The data for this plot was acquired with a fast photo diode – a Femto[1] HCA-S200M-SI was used which is a fast DC diode with a maximum detectable frequency of 200 MHz – connected to an oscilloscope – a DSO5052A from Agilent Technologies[2] which can record signals with a frequency up to 500 MHz. The distance between camera and measurement device was always 45 cm. The plot contains information about the dynamic range of the modulation signal and its shape.

The 19k's plot right at the top shows a very solid LED signal: it doesn't vary much over time and has a good dynamic range with a very low DC part.

---

[1] http://www.femto.de

[2] http://www.agilent.com

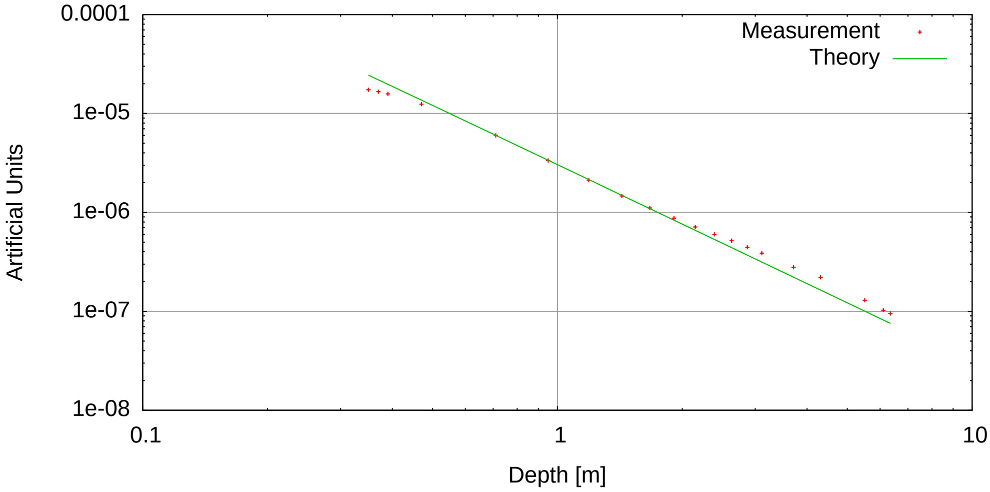


Figure 4.2: SwissRanger SR-3000 amplitude falloff

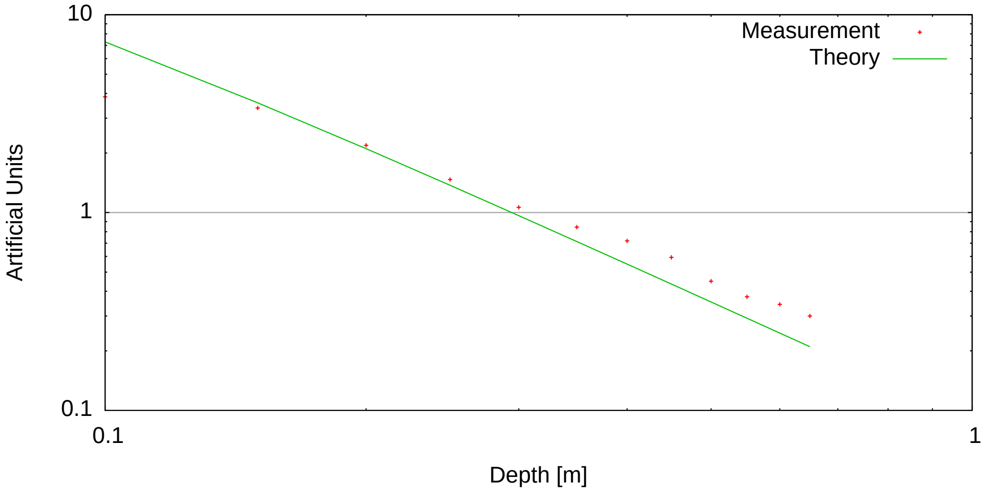


Figure 4.3: Effector O3D amplitude falloff

The SR-3000's signal below looks quite similar. Its DC part is even lower and the signal varies a bit more over time. The most dramatic difference is the much lower overall amplitude. This is due to the much lower energy output of the SR-3000's LEDs.
The O3D's LED signal is not so good. It has a higher amplitude and a quite low DC part, but it is unstable over time – this can be seen in the zig-zagging. Also the form differs from a sinus, especially there's a over-swing in the low third of the curve. This is likely due to the use of surface mounted LEDs in the O3D camera. According to the manufacturer, they are more sensitive to temperature and therefore the signal contains more noise.

Figure 4.5 shows the Fourier transformation for the LED signal shapes. They have been standardized on the first Fourier component so they can be better compared.

The higher harmonics are very strongly suppressed for all cameras, but the 19k shows the best distribution: the impact drops monotonically for the odd components and the 7th and all higher modes are already suppressed below 1% of the first mode.
The SR-3000's signal is not this optimal: while the impact drops, especially the third component – an odd number and therefore with a high impact on the error (see section 1.2.1) – is stronger than 10% of the first mode. Generally, the higher modes are not as strongly suppressed as for the PMD19k.
The O3D camera shows the worst signal as shown already in 4.4. While the impact of the components drop, it doesn't drop monotonically: the 7th and 9th mode are equally weighted in the signal. Generally, the signal and the suppression is quite similar to the one of the SR3000.

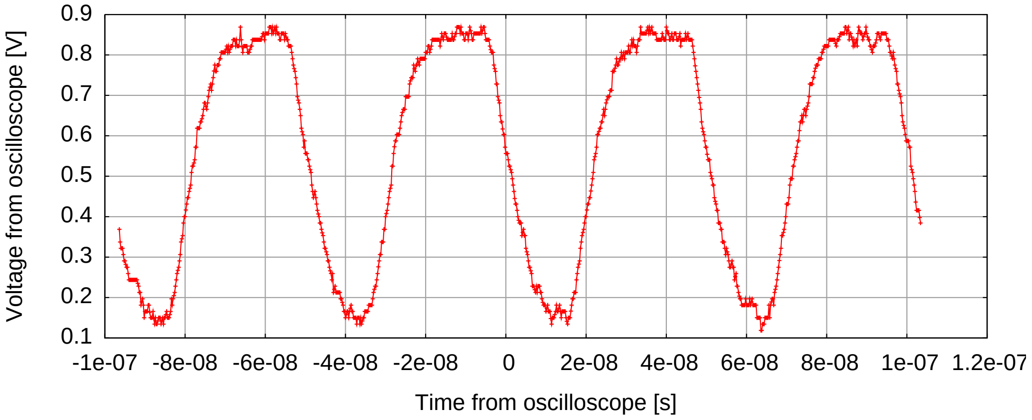
4.3 Systematic Errors

Since the correlating TOF measurement principle is very young, not much is known about its systematic errors yet. Some of them were already mentioned in earlier publications (see e.g. [6, 7]) and also some calibration effort was undertaken (for example by [13, 14, 15]); a systematic approach to the identification and correction of TOF in general has not yet been carried out. This section provides such an analytical investigation.

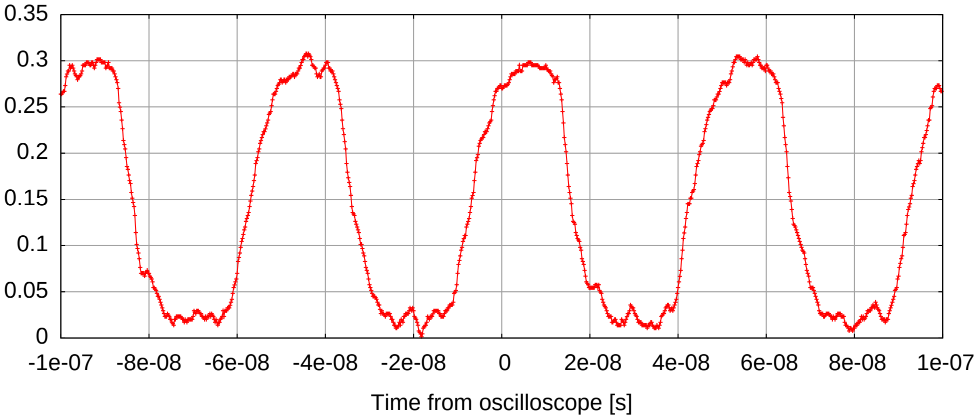
In figure 4.7 measured data for a complete depth range is shown for all three cameras. Only one central pixel was plotted but with different integration times. The data was acquired and averaged as described in chapter 3. Over- and underexposed pixels have already been removed in the plot, therefore the different curves span different depth ranges. This plot already shows two distinct systematic errors.

The first one is the "wiggling", this means the error due to the anharmonic properties of the optical and reference signals as described in section 1.2.1. This will be discussed in the next section. The second is the different offset for different integration times which will be discussed in section 4.3.2.

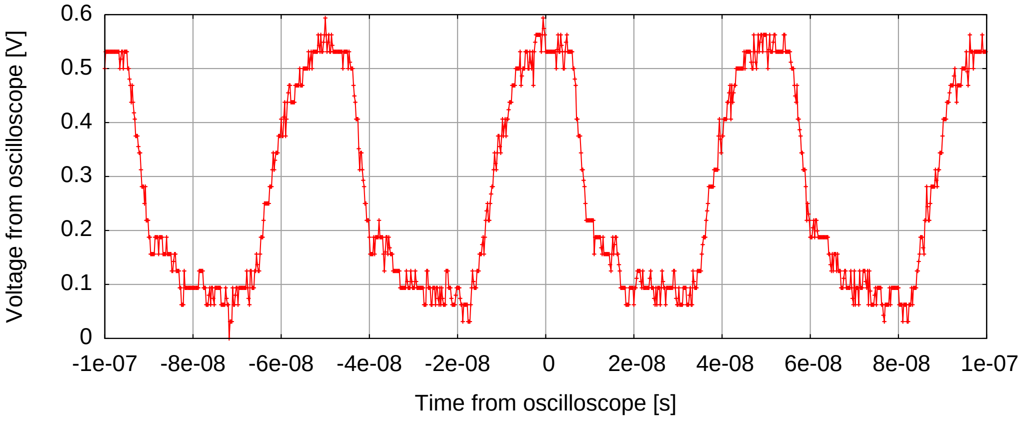
Underexposed pixels are no systematic errors per se since the theory predicts a high statistical error in the depth measurement for low exposure. But since they are usually not usable at all, correcting for underexposure has been included in this chapter in section 4.3.4. Overexposed pixels could also be accounted into systematic errors. Since overexposed pixels do not contain any valuable data at all and must therefore be completely ignored the correcting for them has already been discussed in chapter section 3.3.



(a) 19k

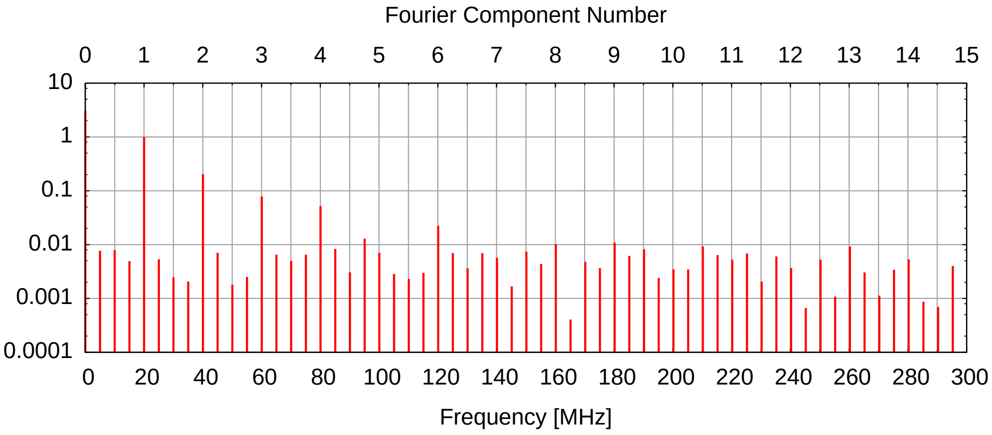


(b) SR-3000

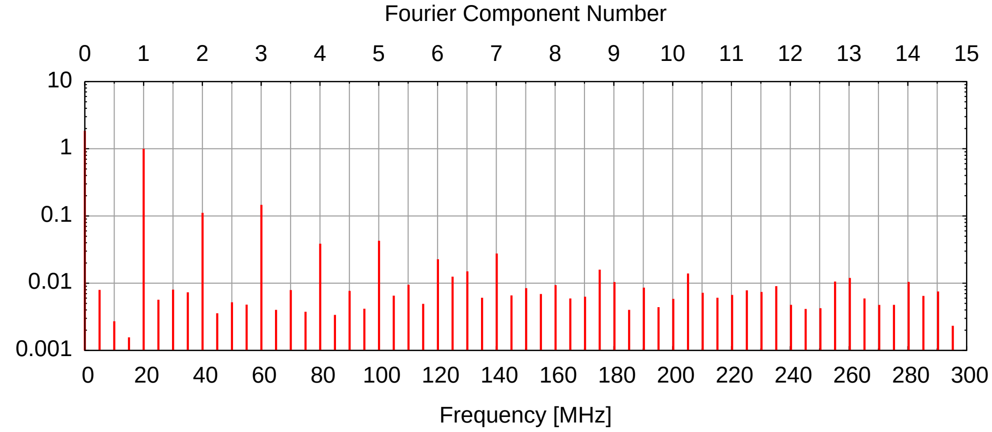


(c) O3D

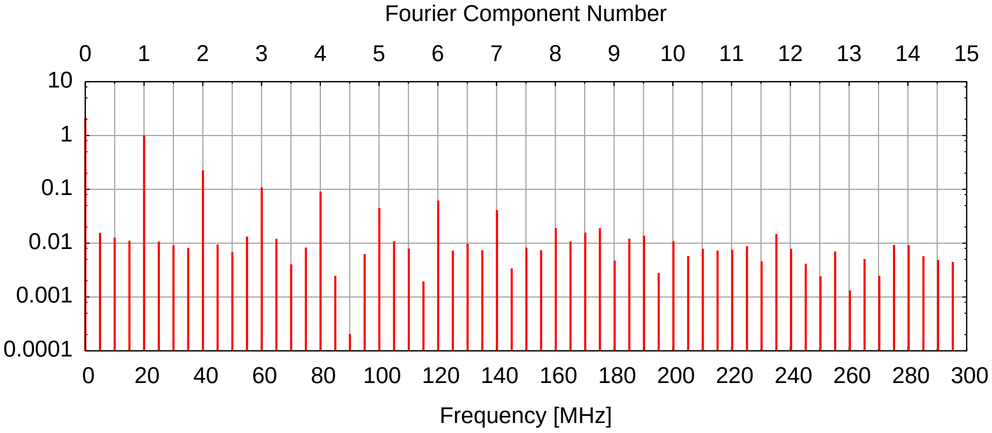
Figure 4.4: LED signal forms as measured by oscilloscope, depth between diode and LEDs was 40 cm



(a) 19k



(b) SR-3000



(c) O3D

Figure 4.5: Fourier analysis of LED signals from plot 4.4. The data has been standardized on the first Fourier component

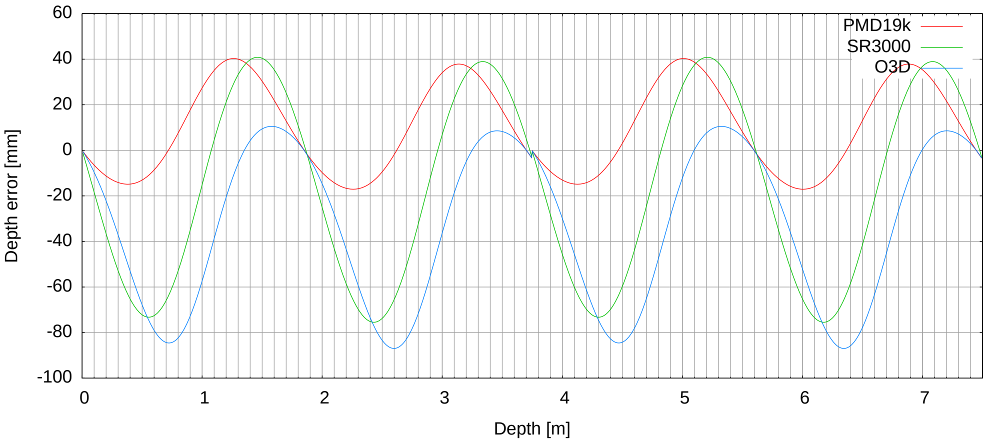


Figure 4.6: Wiggle predictions for all three cameras

The last systematic error discussed in this chapter is the offset introduced by different pixels. Section 4.3.3 treats with this.

4.3.1 Error due to Anharmonic Correlation Functions

The theory of the wiggling error was provided in section 1.2.1. All the LED signals discussed in section 4.2 can be very well fitted with a Fourier series to the 10th order. This gives a very close analytical approximation for the signal shapes of the cameras. Using this analytical expression as theoretical incoming and a square wave as reference function of our camera system and correlating these two functions gives us a prediction of the real correlation function that the camera measures. Therefore we can predict all sample points and use these to calculate the depth information. This yields a prediction of the systematic wiggling error for each camera system. A plot for all three predictions can be seen in figure 4.6.

The predictions for all cameras show a clear periodic behavior. The wave length of this effect is around 1.85 m which corresponds to a combination of the 3rd and 5th fourier mode of the LED signal (see section 1.2.1). The range prediction of the depth error is different for each camera and amounts to around 55 mm for the 19k, 110 mm for the SR-3000 and 95 mm for the O3D camera. It can be seen that all errors approximate the shape of a sine wave – the exact form depends on the weighting of the 5th and 3rd fourier mode of the LED signal relative to each other. The O3D signal is a bit oblate. It is also obvious that the signals do not change much with depth, each period has a very similar amplitude and wavelength. The next section compares this predictions to the measured data.

The measured signal shape seems to be not enough to correctly predict the wiggling errors since all predicted values are quite symmetric – and the real data is not. A reason why the prediction might be incorrect is that the assumption of a square wave reference signal shape is not correct. It is a difficult task to acquire the correct signal form directly from the chip but is a necessary step to enhance the prediction of the wiggling error.

The figure 4.7 now displays the experimental data for comparison with the prediction. The plot shows a complete measurement range for each camera system with different integration times. All overexposed and underexposed pixels have been removed – therefore the functions start and end on different points each. We will discuss this plot for each camera system with focus on the wiggling (not the different integration times) and in comparison with the plot 4.6.

Note that the theory might predict a different phase shift in the wiggling than the reality shows. This is because the LED signal as measured from the oscilloscope has a relative timing information only and not an absolute one. Especially the information how the optical signal lays compared to the reference signal can't be acquired. Therefore the phase delay can't be measured and only the amplitude, the frequency and the shape of the signal can be compared.

PMD[vision] 19k

The first thing that hits the eye in plot 4.7(a) is the jump at around 4 m: this is a sudden change of the depth offset through run time discussed in section 3.4 in the middle of a measurement (a hundred frames were acquired, the 50th frame and all measurements before had one offset, the 51th and all later one 20 cm higher). This happened twice in the scope of this work with the camera and it is suspected that this is a bug in the 19k. It can't be said though if it is a bug in the specific camera used or if this error can be found in all 19k models. All further discussion ignores this jump and discusses the effect for those high integration times as if the offset was still the same.

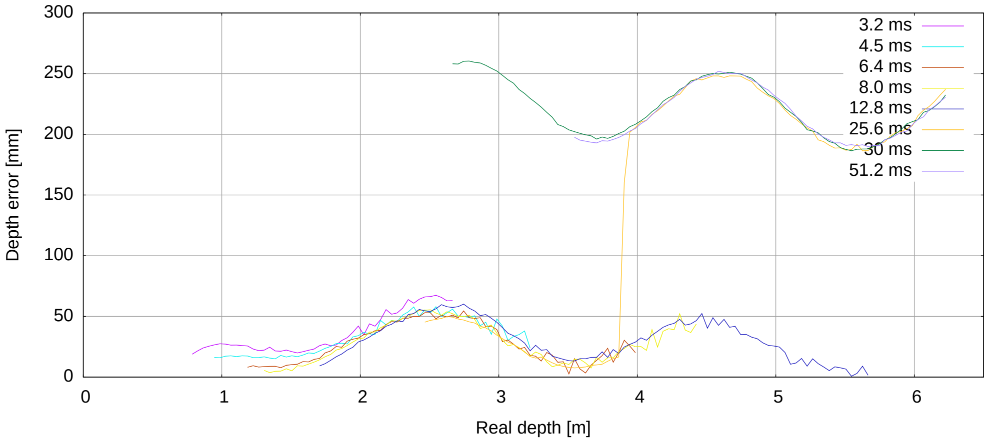
Otherwise the wiggling fits pretty good to the prediction: it has a range of around 80 mm, but if the slope is taken into account and the range is only measured on one period it is around 50 mm which is very close to the 55 mm predicted by theory. Also the wavelength matches pretty good: it is approximately 1.8 m. The measured signal has a slight slope downwards though which was not predicted but overall the theory provides a pretty adequate description for the wiggling of the 19k camera except for the range below 2 m. Here the near field effects dominate – that is the different time of flight of the light sent out by the left and the right LED array of the 19k camera introduce a higher error than the anharmonic modulation of the LEDs.

SwissRanger SR-3000

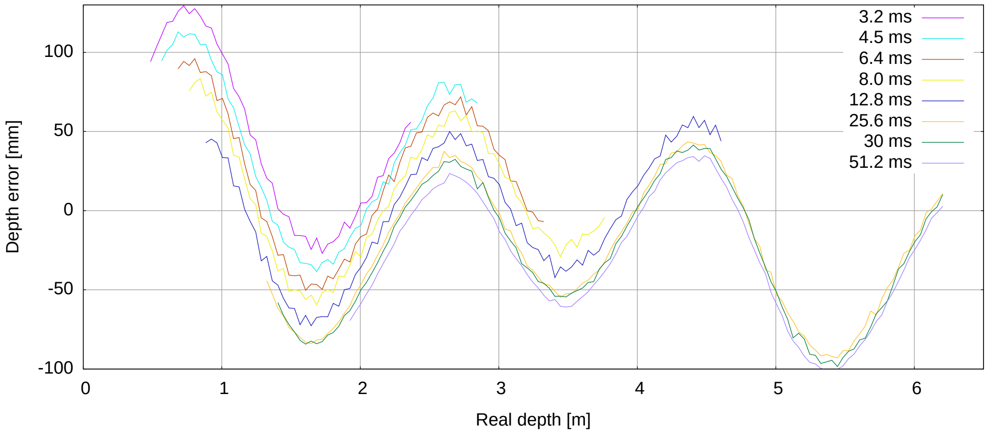
The plot for the SR-3000 camera is shown in 4.7(b). It also clearly shows a sine shaped wiggling effect. Here too, the frequency is quite well matched (wavelength is approximately 1.9 m). The total range of the wiggling comprises around 120 mm which is very close to the theory. The slope of the signal is much stronger and the amplitude of the wiggling varies over depth while the theory predicts a completely self-similar signal. Using the signal as a prediction would over-correct the depth error with higher distances. Still, the theory provides a pretty good prediction of the signal.

Effector O3D

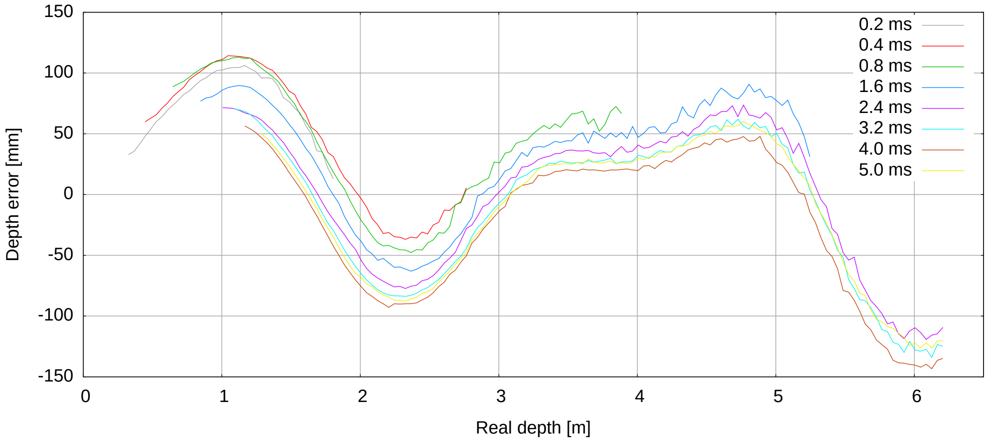
The signal of the O3D camera has a completely other shape than predicted by theory as can be seen in 4.7(c). The wiggling has only an approximated sine shape for the depth



(a) 19k



(b) SR-3000



(c) O3D

Figure 4.7: Depth error to real depth for various different times, one central pixel

values below 3 m, the shape for higher values is not similar to the theoretical prediction. Only half a wavelength can be seen in the lower area, but it is around 1.2 m which would make a total wave length of 2.4 m, higher than the theory predicts. The total range of the error in this area is around 140 mm while the theory predicts 95 mm.

Interesting to note is that the wiggling is still independent of other factors. It can be seen in this plot that the integration time is not a factor (this is true for all camera models) and later we'll see that also different pixels have the same wiggling error.

Summary

The wiggling prediction is in quite good agree with the experimental data for the SR-3000 and the 19k cameras. The predictions are close but not good enough to provide a basis for depth corrections. It is likely though that the predictions can be enhanced to this level with a higher sampling of the LED signal and when the reference signal in the camera is known.
The O3D camera shows fewest correspondence with the theory. But the O3D's LEDs' signal shape is also the most complex of all cameras. It is likely that the prediction will enhance as soon as higher Fourier components are known.

4.3.2 Integration Time Offsets

As mentioned before, plot 4.7 shows another systematic error: a constant offset related to the integration time. All curves in all plots lay parallel to each other, only shifted by a constant factor. The relation is not linear as showed in figure 4.8: here the depth error at 2.5 m is plotted against the integration time for all three cameras. The high integration times have been omitted for the 19k because of the spontaneous change of offset that occurred in the measurement (see previous section for discussion).

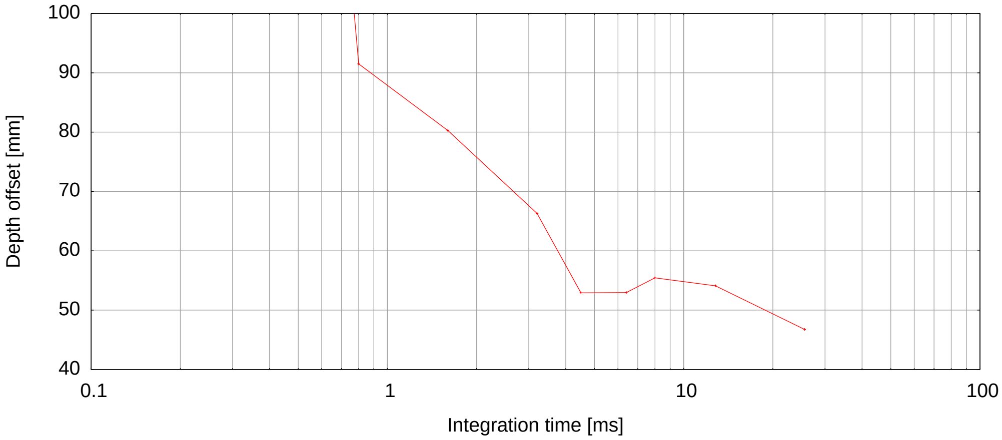
All three cameras show a distinct behavior but all three have in common that the offset is dropping with higher integration times. The 19k shows a very fast drop of the offset at the beginning while the SR-3000 has approximatively the same drop for each decade. The O3D also shows an constant drop per decade for higher integration times, the offset decreases slower for low integration times.

Changing the integration time automatically changes other parameters like amplitude and the time the LEDs are turned on. But the effect discussed here is not related to the amplitude, otherwise it would change with depth – but the offset stays constant over depth for any integration time. More likely, it is related to the recorder electronic or the LEDs.

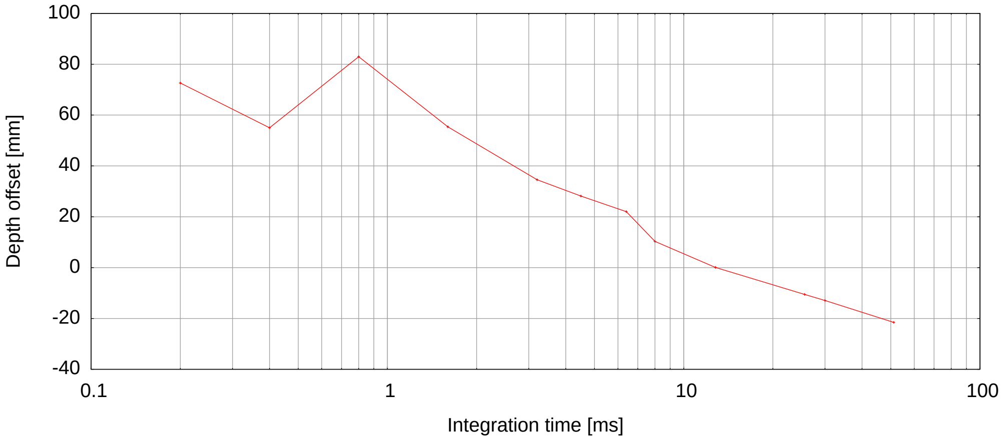
This effect was discussed with PMDTec, but it remains unsure where it results from.

4.3.3 Different Pixel Offsets

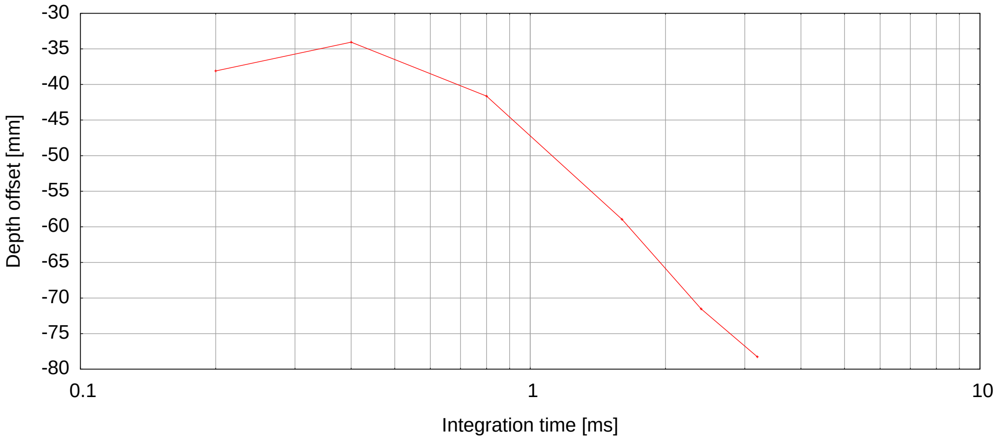
Figure 4.9 shows the effect discussed in this section exemplarily for the SR-3000 camera, but it is visible in all camera systems. This effect is very similar to fixed pattern noise found in 2D cameras. There, the effect results in different sensitivities and gains of the



(a) 19k



(b) SR-3000



(c) O3D

Figure 4.8: Depth error to integration times for d = 2.5 m. The integration times 30.0 and 51.2 ms have been omitted for 19k camera because of the jump (see 4.7 (a))

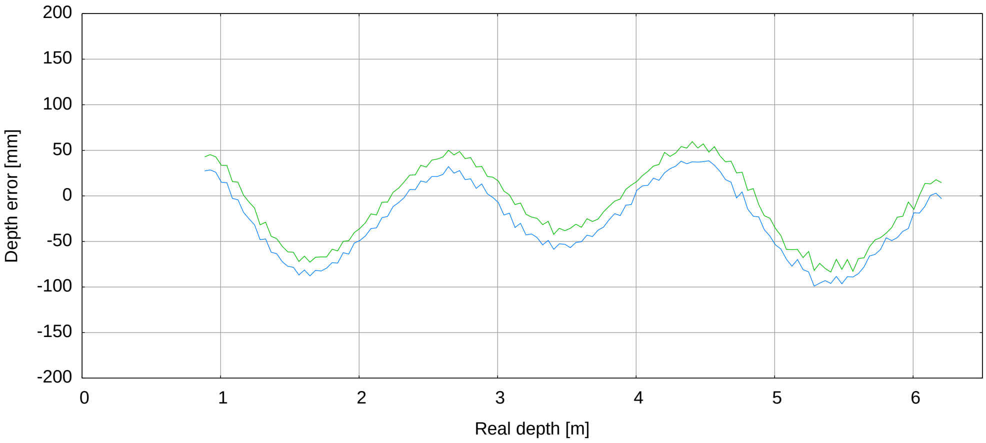


Figure 4.9: Depth to depth error with SR-3000 camera, integration time 12.8 ms, two pixels in close neighborhood.

pixels. The electronics of the TOF-Camera systems also have this problem, but as always two measurements are subtracted from each other in the calculation of the phase shift (see equation (1.21)) this effect vanishes and therefore all offsets disappear. This has been throughouty discussed in section 3.2.

So this can not be the reason for the different pixel offsets. But there is another reason why there still are fixed patterns in the depth information of any TOF pixel matrix and it is due to the fact that each pixel has a capacity: the reference signal needed for the on-chip correlation is delayed by the charge time of the capacitor in each pixel. This delay is equivalent to shifting the reference signal by another small phase shift α_p in equation (1.2) which results in shifting the whole correlation function by this value. Therefore, the data for the least square fit is slightly off and this leads to a slightly wrong phase shift φ_d and therefore a slightly wrong depth calculation.

Since the capacity of each pixel is constant, this effect introduces a constant depth error that is visible in all measurements as a fixed pattern noise. All techniques for removing fixed pattern noise in 2D cameras apply here as well, but this problem is generally more good-natured since the pattern doesn't depend on exposure time or amplitude.

4.3.4 Underexposure

As mentioned before, underexposed pixels do not really introduce systematic errors but rather a higher statistical fluctuation. This is in direct accordance with the theory as discussed in section 1.2.1 and as derived in equation (1.33).

Figure 4.10 shows the dependence for the SR-3000, the relations are similar for the other camera types. The integration time was short to ensure that the amplitude gets low very quickly. In the top, the depth error is plotted – the error bars show the standard deviation of the depth measurement – and below is the run of the amplitude for the same measurement. The green line in the top plot shows the complete progression of the measurement while the red line was masked with an amplitude value: only the pixels with an amplitude higher than 1% of the maximal amplitude are used. It is visible that already a lot of data is dropped with this mask. This is due to the quadratic nature of the

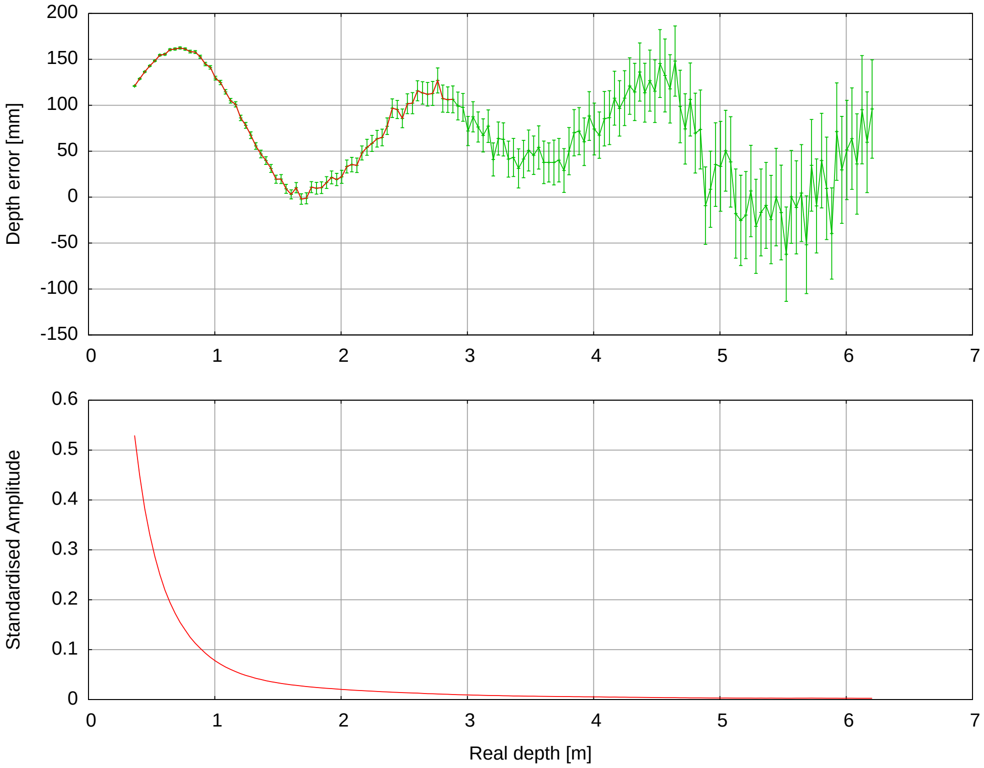


Figure 4.10: Depth to depth error and amplitude with SR-3000 camera, integration time 1.6 ms, one central pixel, the red line in the top plot uses only values with an amplitude > 1% of the maximal amplitude

amplitude decrease with depth. But using this mask ensures a lower statistical error in the depth data and therefore a higher reliability. Note that the statistical error is already very high and increases monotonically with the decrease of the amplitude. The next section provides a more detailed discussion of the statistical properties of the cameras.

4.4 Statistical Errors

Equation (1.33) predicts a relationship between the variance of the depth measurement to the amplitude of

$$\mathrm{var}(d) \propto A^{-2}. \tag{4.1}$$

This relationship has been investigated in figure 4.11. All pixels that show the target have been considered and all available integration times, only correction for overexposed pixels has been done. The red line shows the theoretical slope, the thicker blue line shows the mean value for this amplitude and the thin blue lines show the maximum and minimum values. The results are discussed for all cameras below, the effect is quantified in the overview table 5.1.

PMD[vision] 19k

The 19k camera shows a low variance of the variance, only at high amplitudes and very low ones are outliers in the maximum visible.
The slope is very much as predicted by theory, at least for small amplitudes. As the amplitude rises, the experimental variance drops slower than predicted by the theory. This is likely due to electronical effects that increase with exposure inside the camera – these effects are not taken into consideration in the theory.

The high peaks at around amplitude 100 are interesting too. It is not sure where these effects come from, but they are likely related to electronic amplification. They also suggest that the 19k camera doesn't deliver reliable data in high amplitude range for one shot acquisitions, but the mean data looks smooth and reliable here.

SwissRanger SR-3000

The curve for the SR-3000 cam shows a similar run, but the experimental data diverges earlier from the theoretical prediction, but not as far as for the 19k. There are no outliers except for very low amplitudes.

Also the SR-3000 shows the behavior, that the maximum values of the variance are more pronounced with higher amplitudes. But the reasons for this are even harder to predict than for the 19k camera because it is possible that some considered pixels were still overexposed (overexposure correction for this camera has been done manually with one central pixel as criterion for the reasons described in section 3.3).

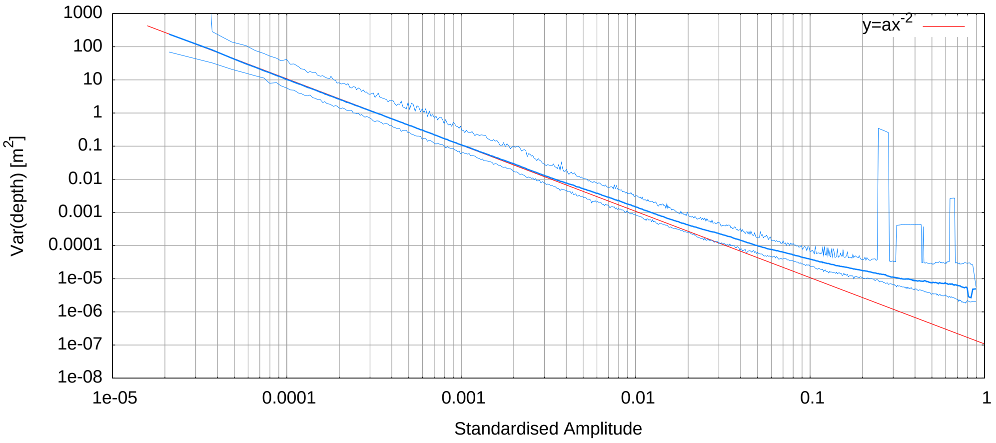
Effector O3D

The plot for the O3D also follows the prediction closely for lower amplitudes but the mean values for the variance are not as smooth as for the other cameras. The O3D also shows the effect that experimental variances are higher for high amplitudes than predicted by the theory. Here again, the same predictions apply as for the 19k.

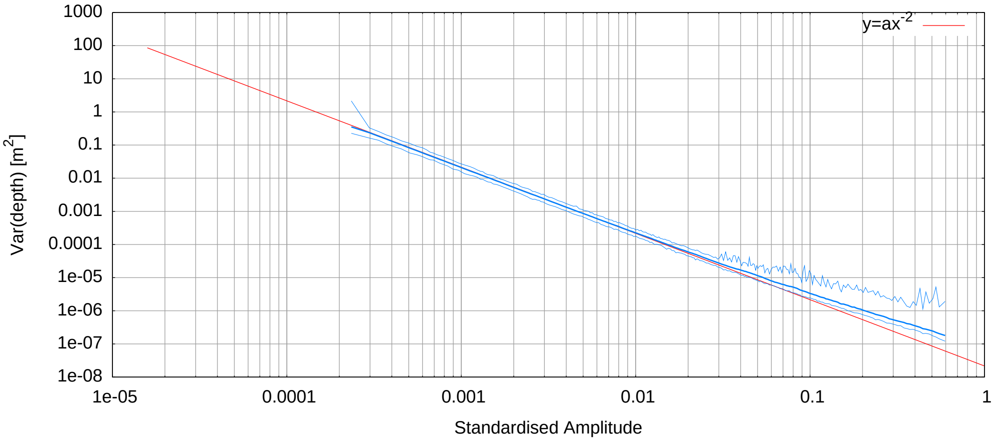
The variance of the variance behaves completely as expected here: it is higher with low amplitudes and shrinks with amplitude increase.

Summary

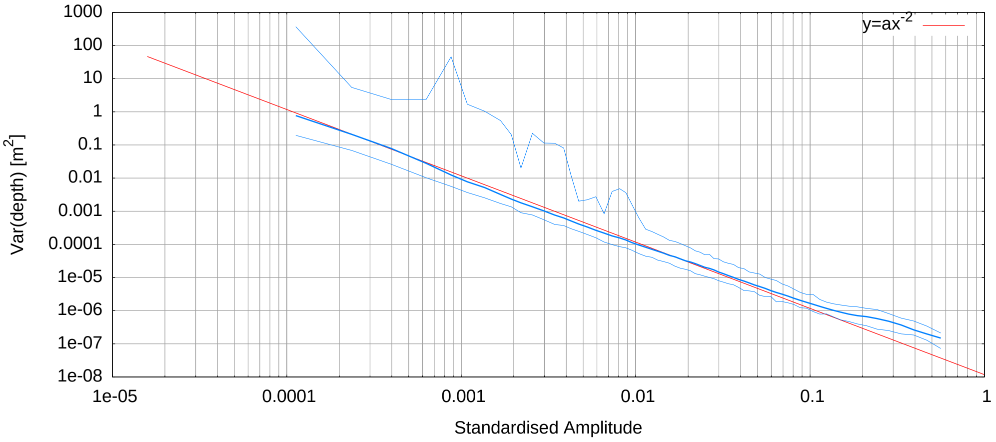
It is important to notice that the variance drops monotonically with the amplitude for all three cameras. Using it as confidence information is therefore valid. Also the deviation from theory occurs late – only with high amplitudes – for which the statistical error is much lower than any systematic error. For most practical purposes the statistical properties of the cameras are well enough described by the theory.



(a) 19k



(b) SR-3000



(c) O3D

Figure 4.11: Amplitude to variance of depth. The red line shows the theoretical slope of var$(d) \propto A^{-2}$, the thick blue line is the mean value and the thin blue lines are the maximum and minimum values of the variance at this amplitude. All pixels on target and all integration times have been considered, overexposed pixels were removed.

Chapter 5

Discussion and Summary

All experimental data and their interpretation have been presented in the last chapter. The systematic errors found propose a calibration order which is detailed in this chapter. After this, implications of the experimental results will be discussed. This enfolds current limitations of the technique and the systems and open questions for further research.

5.1 Impact of the different errors

Table 5.1 shows a summary of all the different effects investigated in the previous chapter and how much they impact the data of the various camera systems.

The wiggling error is the dominating effect for all three cameras, but its total amount varies very much between the systems. The newest camera – the O3D – shows the strongest deviation. The second strongest effect is the integration time offset. The 19k shows a very good behavior here while for the SR-3000 this effect is nearly as big as the wiggling error. The pixel offset is quite similar for all camera types. From the static effects it has the smallest impact of all.

These three effects are the main errors and easy to correct (see 5.2). The statistical error though is harder to compare. We already saw in section 4.4 that all cameras behave qualitatively as predicted by theory. The quantitative comparison has been done for one pixel that was exposed to 30% of the maximum amplitude. No averaging took place. The table 5.1 lists the mean deviation of this pixel for 10 repeated measurements. The two

Effect	19k	SR-3000	O3D
Wiggling error	80	120	200
Integration time offset	35	100	60
Pixel offset	20	20	20
Statistical error (mean. deviation for 10 frames at 30% of maximum amplitude)	± 23	± 9	± 9

Table 5.1: Overview of the total amount of the depth error of the different effects, all values in mm

newer camera models have a better performance here than the old 19k, but overall it can be said that the systematic errors are dominant.

5.2 Suggested Calibration Technique

In the course of this work, various systematic errors have been found. They have been thoroughly discussed in section 4.3. In order to maximize information and precision of measured data, a calibration is unavoidable. The following is a simple two step calibration, that should decrease the impact of the investigated systematic errors. It is likely that other systematic errors which are yet undiscovered might show up in the calibrated data.

Also note that the impact of constant IR light (sunlight or other industry machinery), the environment (narrow spaces, high-reflecting surroundings) and the temperature of the camera has not been investigated. It has been seen qualitatively that these effects distorts the measurement but the suggested calibration does not take those effects into account.

1. Calibrate integration times

   With one fixed distance to a level target, acquire frames for all integration times you want to use later. Take many frames and average over time to reduce noise. Also correct your data as proposed in chapter 3, especially drop overexposed pixels and calculate orthogonal depth. This step will give a different depth offset for each integration time.

2. Calibrate pixel offset and wiggling error

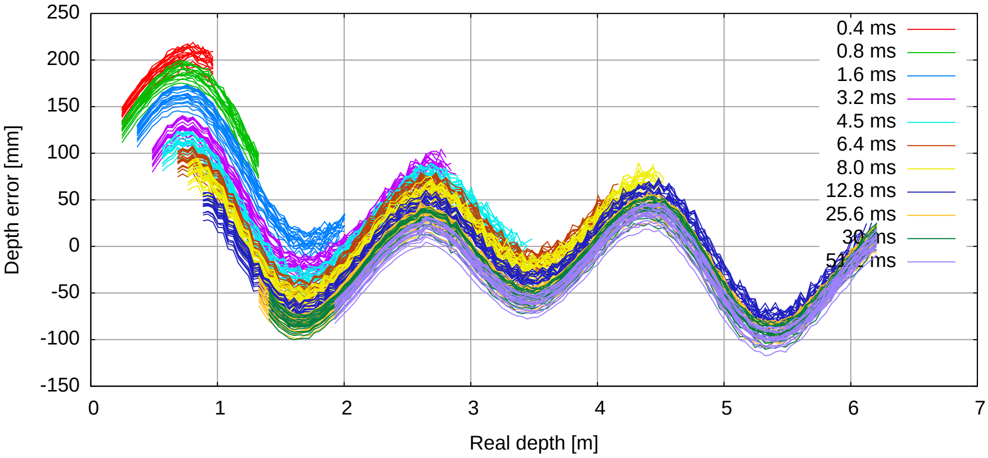
   With one fixed integration time, vary the distance to the level target. You need to calibrate for the complete depth you want to measure later on. Acquire a frame (again with averaging over many frames and correction) for each distance. Correct the integration time offset by subtracting the value resulting from step 1. The acquired frames provide data points for a fitting function which should be used as a lookup table in measurements.

This calibration method corrects at least the integration time offset (section 4.3.2), the per-pixel-offset (section 4.3.3) and the wiggling error (section 4.3.1). It also drastically improves the near field error of the 19k camera.

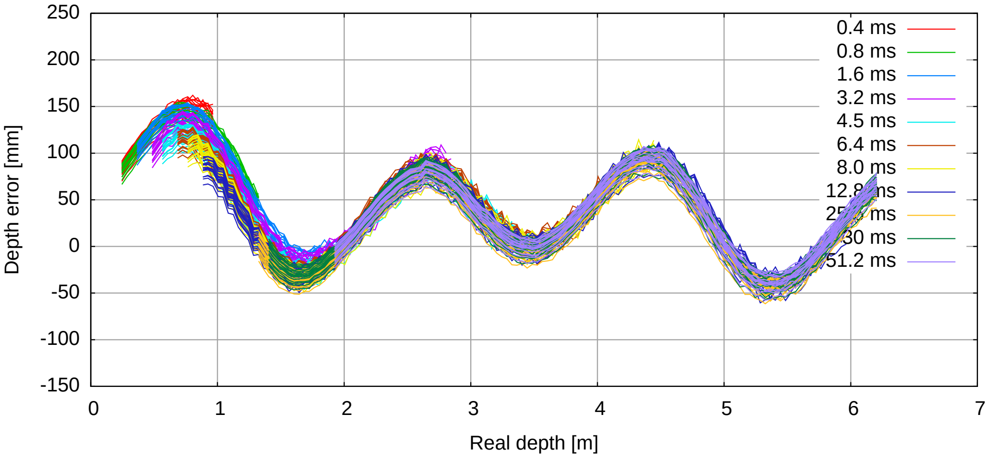
For future works, it is suggested to apply this calibration before further investigating systematic errors.

5.2.1 Sample Calibration

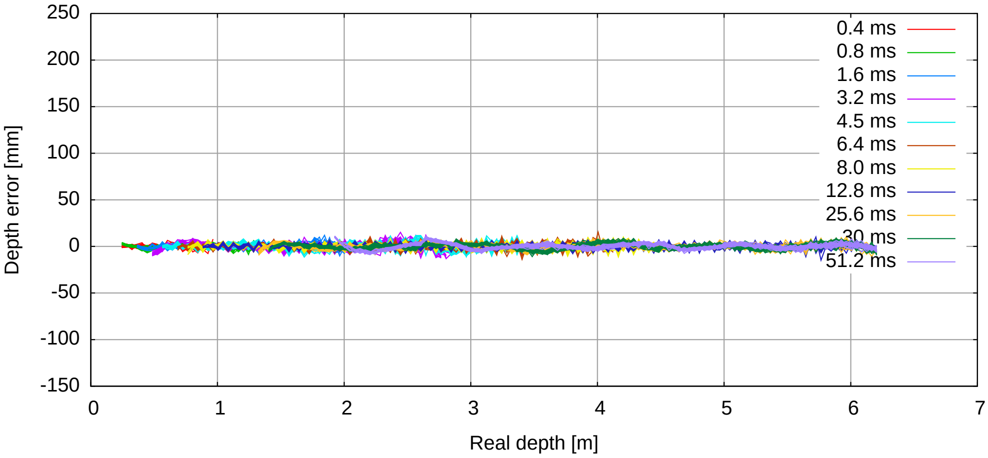
This section presents a sample calibration following the steps described in the previous section. The calibration has been done for 30 central pixels of the SwissRanger SR-3000 with the high-reflectivity target. This is therefore not a complete calibration – for this, all camera pixels must be taken into account – and the results of this calibration have not been tested with data from other surroundings. This sample calibration is thus not



(a) Uncorrected



(b) After integration time offset correction



(c) After wiggling and pixel offset correction

Figure 5.1: SR-3000 camera calibration example

Integration Time [ms]	0.4	0.8	1.6	3.2	4.5	6.4
Depth Offset [mm]	56.91	44.76	19.60	-4.59	-7.67	-20.20
Integration Time [ms]	8.0	12.8	25.6	30	51.2	
Depth Offset [mm]	-25.05	-38.50	-47.65	-53.70	-59.75	

Table 5.2: Integration time offsets, acquired through averaging the depth error at real depth 2.045 m.

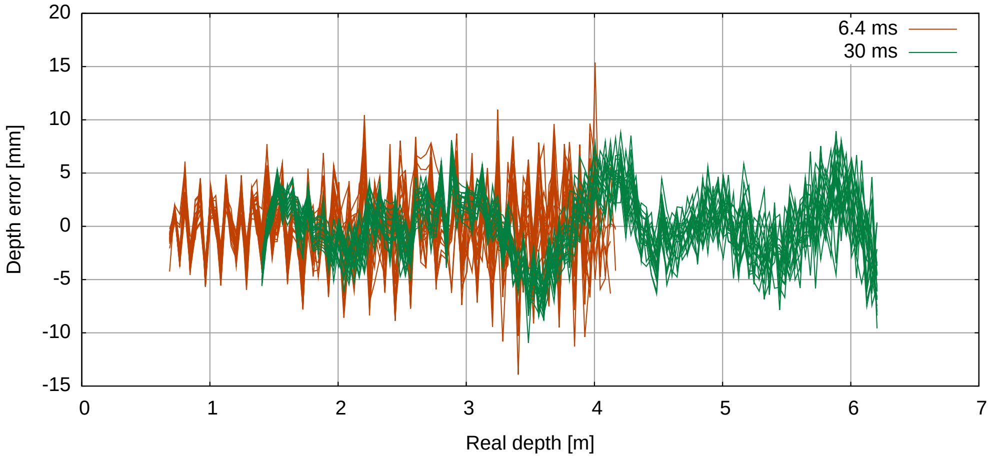


Figure 5.2: Remaining depth error after calibration for integration time 25.8 ms

for employment in field but to illustrate and quantify the error reduction through the recommended calibration.

The plot in 5.1(a) shows the uncalibrated data. Only overexposed and underexposed pixels have been removed. The plot is therefore very similar to the figure 4.7(b), but in the figure 5.1(a), there are all 30 different pixels plotted, each for many integration times. Each time is represented by a color. The plot has the same characteristics as 4.7(b), though: the different integration times show a systematic offset. This offset has been estimated through the mean of the depth error at a real depth of 2.045 m for each integration time. The resulting offsets can be seen in table 5.2. The position was chosen because no integration time is overexposed there. Unfortunately, the lower integration times are quite underexposed already, thus the offset estimation is biased through high variances for them. The offset calculation could be enhanced for them by taking the amplitude as a weight into account.

The plot 5.1(b) shows the data from (a) corrected with the time offsets: every data point has the offset subtracted from it. The error reduction is already quite strong: while we had an error span of 300 mm in (a) the span is now only 200 mm. The main result though is that all integration times lie now very well on top of each other.
In the second step of the calibration, this data is taken and a lookup table is created for each pixel which maps the measured depth to the real depth. The lookup table is implemented here using an unweighted spline interpolation through the data points formed by the measured and time-offset-corrected depth data as x values and the real depth as y values.

This step gives us a per-pixel-correction function which was used to correct the data from

5.1(b). This leads to the data in (c). The total error is now tremendously reduced, it only varies between a maximum absolute error of 15 mm. The absolute error is thus below two centimetre for the whole depth range! The same plot but only for two integration times is shown in figure 5.2. Most interesting to notice is that all 30 pixels for each integration time still lay very well on top on each other – basically each color forms a thick line with less then 2 mm diameter. This suggests that there are still some systematic errors left which were not corrected by this calibration.

This example shows that even a simple calibration process can enhance the data dramatically: the maximum error of uncorrected data is approximately 30 cm with a standard deviation of 40.81 mm, the maximum error after this calibration is around 1.5 cm with a standard deviation of 3.16 mm. The error is reduced by a factor of roughly 20.

5.3 Limitations of Current Systems

With some diligence, reliable data can be acquired even with todays systems. This section discusses the most crucial shortcomings of current correlating TOF systems which provide the most problems.

- Low dynamic

  One of the biggest problems of the current systems is the low dynamic range. With one integration time, only a short depth range can be measured reliably. This is obvious, as the amplitude decreases quadratically with depth and the sensitivity of the current pixels is linear. This problem could easily be fixed by using either one of the following approaches:

  - Logarithmically sensitive pixels

    Currently, the automotive industry is investigating photo diodes that measure logarithmically. This is done by keeping the voltage of the capacitors in the pixels low enough to keep them in the nonlinear area. For a general conception see [16], a more specific approach is described in [17]. A throughout presentation of the current state and future development of High-Dynamic-Range vision which presents many possible implementations is [18]. The very same approaches could also be used in TOF-Camera systems. This would dramatically increase the dynamic range and therefore the depth range of the cameras and it would only contain small changes in the gates of the recording units. The author considers this approach as the most promising one.

  - Per-pixel-integration time

    Another approach which has been used in 2D camera systems for a while is to put some intelligence into the pixels and let them determine their integration time dynamically ([19]). As soon as a pixel is well exposed, it stops counting photons. This approach is sensible, but much harder to implement, at least as long as the offset through integration time (section 4.3.2) is not yet understood. If this offset is only created by the LEDs this approach would be a good solution for the low dynamic range, but if the problem is caused to some part by the recorder on the chip this approach would introduce an offset in each pixel that would change with each frame. This is not feasible, so before this approach can be implemented practically, the integration time offset must be understood and a proper correction must be implemented in the cameras.

  - Multiple integration times per frame

    This technique is already in use by the O3D camera. Instead of acquiring a scene only once with one integration time, it can acquire it with two distinct times – a long and a short one – and only the optimally exposed pixels from both shots will be used. This technique increases acquisition time but also increases the effective dynamic range. But as with the per-pixel-integration approach special care must be taken to account for the different integration time offsets.

- Low resolution

  Since most current systems are engineered with a sensoric background in mind, the resolution of all current camera systems is rather poor. The newest model investigated – the IFM O3D – reduces the resolution even further compared to the two older models. This is an unfortunate development since the TOF-Principle promises fast and reasonable precise depth and gray data for image processing tasks. But image processing needs a minimal resolution to properly detect and segment objects. It is therefore desirable to increase the resolution to at least VGA. The Lynkeus project has as a goal to develop such a camera.

  Increasing the resolution of a chip basically means to reduce each pixel's physical size. This also proportionally decreases the number of photons each pixel receives. Therefore the integration time must be increased to guarantee a good exposure and therefore a low variance. Increased integration time lowers the frame rate though (see next point). A better approach would be a more efficient light source. This would allow to keep integration times steady.

- Low frame rate

  The calculations to get depth information from raw data are rather complex and therefore numerically expensive. Thus it is a rather slow task for the processors in the camera systems. Nevertheless, a stable and high frame rate must be achievable for real-time application. This is possible even for higher resolution as the SR-3000 camera proves: it already has an acceptable frame rate. The 19k camera is handicapped through its internal design: the raw data comes from the chip to an embedded Linux system which serves it to the host computer. This is a slow task. The O3D camera adjusts the frame rate depending on heat. The electronics and the chip could deliver data fast enough, but the camera refuses to acquire frames when the housing is too hot.

  Therefore this problem is more an electronical than a physical one and future camera systems designed for image processing tasks will likely be able to deliver a sufficient frame rate.

- Heat

  Due to the active illumination, the cameras get very hot. Refrained from the fact that the temperature also changes the LEDs' behavior and other parameters, the cameras

even can reach a temperature where they get damaged or destroyed. Currently there are two approaches to this problem. The first one uses fans in the housing of the camera (19k, SR-3000). This has the drawback of a higher mechanical complexity and therefore a high fragility. These cameras can also not be used in dirty or wet industry surroundings.
The other approach is to reduce the time the LEDs are turned on. The O3D only turns the light on when a frame is acquired and waits after each frame till it has cooled down before acquiring another frame. This is a good approach for short integration times (where the LEDs are only on for a short period), but for even medium long integration times (1-5 ms), the frame rate of the O3D drops to below 1 fps because the cooling times need to be that long. This approach is therefore not feasible for most image processing tasks, especially not for real-time applications.

Also here, the best solution would be more efficient light sources. They will reduce heat while keeping the optical output at the same level.

5.4 Open Questions

This work provides first steps for a systematic analysis of the errors and limitations of TOF-Camera systems. Not all interesting points were touched; the following points are surely of some interest. They were not discussed in this work, but are stated and briefly commented here. Other open questions are discussed in more detail below.

- Sunlight

  This work only discussed the power of the systems in optimal surroundings: absolute darkness. While artificial light doesn't pose a big problem for current systems since it does not contain any mentionable IR fraction, sunlight is the strict contrary: it contains a high IR fraction and disturbs the measurements because it effectively reduces the amplitude by increasing the constant DC offset. Qualitative experiments showed that the current systems do not provide a reliable information in bright sunlight. The exact relationship between constant IR light intensity and depth derivation must be investigated in future works.

- Temperature

  It is a known fact that the depth measurement drifts with the camera's temperature (see for example [14]). It is yet to be properly investigated how this error behaves systematically with temperature. Also it would be interesting to see, if this error is mainly dependent on the LEDs' strong relation to temperature or if other effects dominate.

- Environment

  It has been seen in this work that the reflections from surroundings (in this case from the linear positioner tables) have an impact on the precision of the depth measurements. It could be seen qualitatively that the reflections from the tables only add a constant offset to all measurements, but it is likely that more complex environment could change the measurement in unknown ways. It is therefore feasible to investigate this dependency more systematically.

This list and the following discussion are not exhaustive, there might be and are likely to be more systematic errors yet to be discovered and more open questions to be investigated.

5.4.1 Errors Introduced through Scene Reflectivity and Amplitude

This work presented a constant depth offset error depending on integration time in section 4.3.2. It was discussed there that this error does only depend on the integration time and therefore either on the LEDs or on the recorder unit.

The logical step from there is to investigate if there are systematic errors depending on amplitude or reflectivity of the scene. Note that these are two different physical problems and must be investigated separately: the reflectivity of the scene (its materials) might have other effects – like reshaping the optical signal either directly or through reflections – than just lowering the returned amplitude. It is difficult though to separate these two effects and we can expect to see an effect only after correcting for the systematic errors already discovered in this work. It is quite sure that the systematic errors of those two effects are little pronounced since they have not led to problems during this research.

5.4.2 Prediction of the Wiggling Effect

The current explanations why the wiggling of some cameras can't be predicted analytically includes the dependence of the LEDs' signals on temperature and on different mounting types (surface mounted vs. normal LEDs). It is suggested to investigate many LED signal periods (this work only investigated up to ten, future works should consider period numbers in the order of thousands) with many different integration times and also investigate the relation to LED temperature. It is also of importance to measure and investigate higher Fourier modes. This work only measured with 200 MHz which only takes the first ten components into account; but it is possible that the shape of the wiggling error is affected by even higher modes. The LED signals must therefore be investigated with a higher sampling rate. These steps should lead to a more profound understanding of the LED signals.

Also the reference signal as it arrives on the chip should be investigated. This needs a complex setup because of the small dimensions of modern ICs. It might be possible to directly get this data from the manufacturers – they were very supportive during this work. The shape of this reference signal affects the asymmetry of the CF as strong as the shape of the LED signal. Therefore the investigation of this signal is of the same importance as the investigation of the LED signal.

With this expanded knowledge, the author is confident that a better prediction of the wiggling can be achieved.

5.5 Summary

This thesis gave a general conspectus about TOF-Cameras. A theoretical model was presented and verified with three different current TOF-Camera systems and systematic and statistical errors were discussed. Also implications and shortcomings of current systems were shown and suggestions for future development and research topics were given.

The experimental results revealed many systematic errors in current camera systems. The periodic variation due to the anharmonic CF provokes a periodic depth error of around 80 to 200 mm (depending on camera), the inhomogeneity of the pixels accounts for around 20 mm. A constant offset depending on the integration time was found which varies between 35 and 100 mm.

However, most of the errors are very easy to correct: overexposed pixels can be masked and the periodic offset due to the anharmonic CF and the constant pixel offsets can easily be removed with a lookup table, the integration time offsets can be subtracted from the measured data. It has been shown that with a simple two step calibration – the first step calibrates the integration time offset and the second the per-pixel-offset and the wiggling error – the quality of the data can be improved significantly. The calibration reduces the total error to below 2 cm for the SR-3000. Thus, with these simple corrections, reliable 3D data can be acquired even with today's systems.

The cooperation with industry partners in the Lynkeus project was very fruitful and satisfying: the experimental setup was used by partners to make their own systematic investigations and the close partnership with PMDTec helped to find and correct at least one bug in the camera hardware which was revealed through the experiments of this work.

The investigated techniques made a strong impression on the author: The correlating TOF-Camera 3D measurement technology is a young but promising new technology which shows a convincing performance even in a prototype state. Current systems lack in speed and resolution but a lot of work is underway. The implications for image processing tasks will be huge: TOF offers an extra dimension in image data with no extra effort. This will increase the already widespread application field of image processing even further.

The technology stands in the focus of industry and research. Many well known companies are either investigating the use or advancing the development of correlating TOF-Systems. It is therefore reasonable to expect significant progress in the near future.

Appendix A

List of Experiments

Abbreviation	Meaning	Explanation
Cameras:		
19k	PMD[vision] 19k	
SR-3000	SwissRanger SR-3000	
O3D	Effector O3D	
Column Headings:		
IT	Integration times	Integration times used in this experiment (in ms)
DC	Data Channels	Data acquired in this experiment
MR	Measurement Range	Depth range that has been spanned (in m). This is the depth range as delivered by the positioner tables, the initial offset can be found in the column "Offset".
NP	Number of Positions	Number of steps that have been taken in MR
NMF	Number of meaned frames	The number of frames that have been acquired on each stop. The mean and variance has been calculated and this data was saved to work with
Offset		Initial offset between camera and highest point of target (in mm)
Extra frames		Yes, if there have been some single frames (without averaging) saved on each position
Data Channels:		
d	depth	Depth information as delivered by camera (without correction)
a	amplitude	Amplitude information as delivered by camera. Note: this parameter is called intensity in the documentation of the O3D.
i	intensity	Intensity information (gray values) as delivered by camera (without correction)
r	raw values	raw values as delivered by camera without correction, the number gives information about how many raw channels were acquired

Table A.1: List of abbreviations used in the following tables

Running Number	Camera	Target	DC	IT	MR	NP	NMF	Offset	Extra frames
1	19k	High reflectivity	d, a, i, 8r	0.01, 0.05, 0.1, 0.2, 0.5, 5, 12, 20, 30, 50	0 - 6	150	150	222	Yes
2	SR-3000	High reflectivity	4r	0.01, 0.05, 0.5, 1, 2, 5, 8, 12, 20, 30, 50	0 - 6	150	150	203	Yes
3	19k	High reflectivity	8r	0.2, 0.4, 0.8, 1.6, 3.2, 4.5, 6.4, 8.0, 12.8, 25.6, 30, 51.2	0 - 6	150	150	226	No
4	SR-3000	High reflectivity	4r	0.2, 0.4, 0.8, 1.6, 3.2, 4.5, 6.4, 8.0, 12.8, 25.6, 30, 51.2	0 - 6	150	150	205	No
5	O3D	High reflectivity	d, a, 16r	0.2, 0.4, 0.8, 1.6, 2.4, 3.2, 4, 5	0 - 6	150	150	206	No
6	19k	Checkerboard	d, a, i, 8r	0.01, 0.05, 0.1, 0.2, 0.5, 12, 20, 30	0 - 6	301	100	221 and 255	Yes
7	O3D	Checkerboard	d, a, 16r	0.2, 0.4, 0.8, 1.6, 2.4, 3.2, 4, 5	0 - 6	150	150	206	No

Table A.2: List of static depth measurements

Dark Current Measurements:

Camera	DC	IT	Nr. of frames per IT
19k	d, a , i, 8r	0.5, 1, 2, 5, 12, 20, 35, 50	500
SR-3000	4r	0.5, 1, 2, 5, 12, 20, 35, 50	500

Table A.3: List of dark current measurements

Appendix B

10 Rules for Using Correlating TOF-Camera Systems

These rules provide a short overview for everyone who wants to start using a TOF camera system. It represents the combination of the practical experience of the author and the scientific results of this work. This list is not exhaustive but provides a solid introduction and helps to avoid common pitfalls.

Rule 1: Average over time if you can! Averaging will decrease statistical errors of all measurements and therefore makes your data much more reliable (sections 1.2.1, 4.4). All current cameras only deliver reliable data in a small depth interval for each integration time due to the low contrast. It is thus likely that many pixels will be badly exposed in a standard scene. These pixels show a high variance, therefore averaging is mostly not an option but a necessity. If you need real-time data, prefer the SR-3000 cam over the 19k for higher ranges ($> 2\,\mathrm{m}$), because of the higher frame rate (but see Rule 3). For smaller distances consider using the O3D camera. See also Rule 2.

Rule 2: Don't average without thinking! It is possible that naive averaging of depth frames over time might decrease your precision (section 3.1). When averaging over time, use the amplitude as confidence information (Rule 7)! Spacial averaging over pixels without proper gauging will introduce errors because of the low resolution of cameras (two pixels see distinct points in space) and because of different offsets (section 4.3.3).

Rule 3: Correct overexposure! Overexposed pixels do not contain any valid information at all. It is of importance to detect and remove them before processing the data any further. This is possible with the PMD19k and the O3D but not with the SR-3000. Section 3.3.

Rule 4: Correct spherical depth information! One thing to consider is that the depth information delivered by the camera is always spherical. But most of the times, the user will be interested in the orthogonal distance. This is easily calculated with some intrinsic camera parameters known. Section 3.5.

Rule 5: Calibrate your camera for your needs! All cameras show distinct systematic errors that will worsen your data (section 4.3). Especially the wiggling errors (section 1.2.1), the constant pixel offsets (fixed pattern noise in depth data) (section 4.3.3) and the integration time dependent depth offset (section 4.3.2) are the major sources for errors. Therefore calibrate your camera (see Rules 9 and 10) for example as suggested in section 5.2.

Rule 6: There is no rule six!

Rule 7: Use the amplitude information! The measurement principle of TOF-Cameras already provides implicitly a confidence information for the reliability of depth information: the amplitude (sometimes also called intensity) information. It can be used to mask out unreliable data (section 4.3.4), but can (and should!) also be used to weight depth information: the higher the amplitude, the better the depth measurement (section 4.4). Mind overexposure though, see Rule 3.

Rule 8: Mind the low contrast! A big problem is the low contrast of current camera systems. Since the amplitude falls off (at least!) quadratically with depth the amplitude signal is quickly too low for reliable measurements (section 4.4). To increase dynamics, use different integration times (But calibrate first, see Rule 5) and consider averaging over time (Rules 1 and 2).

Rule 9: Mind the sun! Some cameras already contain a suppression of background illumination, but all current camera systems are completely lost against the high IR fraction in sunlight: the modulated amplitude founders in the constant DC fraction of the sun. You can't expect reliable data then. Artificial light only contains a small IR fraction, therefore the cameras perform well in industry surroundings.

Rule 10: Mind the environment! Since the cameras use active illumination (IR light) to measure depth, they are sensible to errors from reflection in their surroundings (section 2.2.2). This can easily be seen by moving the hand close to the LED arrays: the whole scene seems to move closer to the camera in the depth data. Therefore avoid narrow spaces and ensure a direct undisguised look to the scene for the camera.

References

[1] R. Schwarte *et al.*, *Principles of Three-Dimensional Imaging Techniques*, In: Handbook of Computer Vision and Application, Vol 1, pp. 488-509, 1999.

[2] B. Jähne, *Digital Image Processing, 6th Edition* (Springer Verlag, 2005).

[3] R. Klette, R. Kozera, and K. Schlüns, *Reflectance-Based Shape Recovery*, In: Handbook of Computer Vision and Application, Vol 2, pp. 556-615, 1999.

[4] R. Schwarte, Z. Xu, H. Heinol, J. Olk, and B. Buxbaum, *New optical four-quadrant phase-detector integrated into a photogate array for small and precise 3D-cameras*, In: SPIEProc3023 p. 119, 1997.

[5] K. Iga, *Surface-emitting laser – Its birth and generation of new optoelectronics field*, IEEE Journal of Selected Topics in Quantum Electronics Vol. 6, Issue 6, pp. 1201-1215, 2000.

[6] Z. Xu, *Investigation of 3D-Imaging Systems Based on Modulated Light and Optical RF-Interferometry (ORFI)*, PhD thesis, Department of Electrical Engineering and Computer Science, University of Siegen, 1999, discussed in "Zess Forschungsberichte".

[7] B. Schneider, *Der Photomischdetektor zur schnellen 3D-Vermessung für Sicherheitssysteme und zur Informationsübertragung im Automobil*, PhD thesis, Department of Electrical Engineering and Computer Science, University of Siegen, 2000.

[8] M. Plaue, *Technical Report: Analysis of the PMD Imaging System*, Interdisciplinary Center for Scientific Computing, University of Heidelberg, 2006.

[9] M. Frank *et al.*, *Theoretical and Experimental Error Analysis of Continuous-Wave Time-Of-Flight Range Cameras*, To be published.

[10] X. Luan, *Experimental Investigation of Photonic Mixer Device and Development of TOF 3D Ranging Systems Based on PMD Technology*, PhD thesis, Department of Electrical Engineering and Computer Science, University of Siegen, 2001.

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

[12] D. Justen, *Untersuchung eines neuartigen 2D-gestützten 3D-PMD-Bildverarbeitungssystems*, PhD thesis, Department of Electrical Engineering and Computer Science, University of Siegen, 2001.

[13] M. Lindner and A. Kolb, Lateral and depth calibration of PMD-distance sensors, in *In: Proc. Int. Symp. on Visual Computing*, pp. 524–533, Springer LNCS, 2006.

[14] M. Strehler, *Messgenauigkeit und Kalibrierung von Laufzeitkameras*, Fraunhofer Institute for Manufacturing Engineering and Automation IPA, 2007.

[15] Kahlmann, Remondino, and Ingensand, *Calibration for Increased Accuracy of the Range Imaging Camera Swissranger*, In: IEVM06 36 3 136–141, 2006.

[16] U. Seger, U. Apel, and B. Hoefflinger, *HDRC-Imagers for Natural Visual Perception*, In: Handbook of Computer Vision and Application, Vol 1, pp. 223-235, 1999.

[17] J. N. Burghartz *et al.*, *HDR CMOS Imagers and Their Applications*, Institut fur Mikroelektronik Stuttgart, IMS CHIPS, 2006.

[18] B. Höfflinger, *High-Dynamic-Range (HDR) Vision* (Springer Verlag, 2007).

[19] B. Schneider, P. Rieve, and M. Böhm, *Image Sensors in TFA (Thin Film on ASIC) Technology*, In: Handbook of Computer Vision and Application, Vol 1, pp. 262-295, 1999.

Acknowledgments

First to mention are my colleagues here at the Interdisciplinary Center for Scientific Computing (IWR) at the University of Heidelberg. I found a very fun and skilled team of scientists with a lot of different backgrounds. The widespread concern to actively support a becoming physician was a high motivation throughout this work.

A special thanks goes to Prof. Bernd Jähne for offering me the opportunity and to entrust me with this topic. His excellent supervision and mentoring made sure that this work came to a successful end.

I also want to address a special thanks to my roommates M. Frank, M. Schmidt and P. Pavlov. The first two joined me on the work on the TOF technology and discussions with them were always most fruitful. The latter always proved most skilled and willing to help when a problem with any mathematical issue impeded understanding.

An apology goes to A. Herzog and K. Richter for the enduring noise the experiment made in its building and test phase. I also want to thank for the tea and the encouragement when something went wrong. And of course thanks for bearing with all the jokes.

It is my special concern to mention all the helpful individuals who helped me in the industry. First and foremost F. Forster from IFM Electronics who provided a lot of detailed explanation of the Effector O3D which more than compensated for the lack of documentation.
The persons at PMDTechnologies – especially T. Ringbeck and M. Profittlich – were helpful and supportive with all questions about the 19k camera system. Even though not all questions could be ruled out, it was not due to a lack of trying. I am grateful to T. Oggier and M. Paduano from CSEM who made sure, that our SR 3000 camera was repaired quickly and for free (though it was out of guaranty) when the fan broke down.

I'd like to thank all proofreaders who had to bear with my horrible English and convoluted writing style. If you have understood major parts of this work, it is much to their credit.

I thank my family for their sympathy – my parents without their help and support I would not have been able to study physics and my brother for worthwhile discussion and tips.

A special thanks goes to Hanna Podewski. She has been the sunshine of my life ever since I got to know her.

Erklärung

Ich versichere, dass ich diese Arbeit selbständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe.

Heidelberg, den 20. September 2007

Holger Rapp