Master's Thesis

Investigation of Systematic Errors in Time-of-Flight Imaging

conducted at the

Signal Processing and Speech Communications Laboratory
Graz University of Technology, Austria

in co-operation with

Infineon Technologies Austria AG
Graz, Austria

by

Johannes Grünwald

Supervisor:

Dipl.-Ing. Dr. Christoph Steiner

Assessor:

Assoc.Prof. Dipl.-Ing. Dr. Klaus Witrisal

Graz, January 24, 2013

Statutory Declaration

I declare that I have authored this thesis independently, that I have not used other than the declared sources/resources, and that I have explicitly marked all material which has been quoted either literally or by content from the used sources.

Date	Signature

Zusammenfassung

Time-of-flight (ToF) ist ein bildgebendes Verfahren basierend auf Laufzeitmessung, welches die Aufnahme dreidimensionaler Szenen ermöglicht. Neben mehreren Vorzügen gegenüber konkurrierender Technologien besteht ein Nachteil dieser Methode in der potenziellen Ungenauigkeit der gemessenen Tiefeninformation. Insbesondere existieren verschiedene systematische Messabweichungen, welche seit Jahren Gegenstand intensiver Forschung sind. Es gibt breiten Konsens in der Literatur, dass besagte Fehler nur durch aufwändige Kalibrierung behoben werden können.

Anhand eines aktuellen ToF-Sensors wurde eine ausführliche Analyse systematischer Tiefenabweichungen im Kontext der Signalverarbeitung durchgeführt. Aufgrund eines erweiterten Messaufbaus konnten grundlegende neue Einsichten in das Verhalten dieser Fehler gewonnen werden. Der amplitudenbezogene Fehler (*amplitude-related error*) wurde nachweislich auf Reflexionen durch Mehrwegausbreitung zurückgeführt. Sensorseitige Abweichungen ausschließlich aufgrund der optischen Signalamplitude konnten nicht festgestellt werden. Eine amplitudenbezogene Kalibrierung dieses Fehlers kann daher nicht zielführend sein, was im Widerspruch zum aktuellen Stand der Technik steht. Vielmehr müssen neue Strategien hinsichtlich Reflexionskompensation entwickelt werden, um entsprechenden Abweichungen entgegenzuwirken. Weiters kann der Integrationszeit-bezogene Fehler (*integration-time related error*) höchstwahrscheinlich mit Aufwärmungseffekten und damit verbundenen Signalverzögerungen der Beleuchtungseinheit begründet werden. Hier handelt es sich ebenfalls um keine sensorseitige Problematik.

Der wesentlichste Beitrag dieser Arbeit ist jedoch die Entwicklung verschiedener Methoden welche den distanzabhängigen Fehler (*Wiggling Error*) signifikant reduzieren. Es konnte in Theorie und Praxis gezeigt werden dass der Stand der Technik, nämlich die Aufnahme vier äquidistanter Abtastwerte, aufgrund massiver Aliasing-Artefakte den Wiggling Error maximiert und daher eine denkbar ungünstige Variante darstellt. Tatsächlich stellte sich heraus, dass Distanzschätzung basierend auf *drei* Abtastwerten der traditionellen Vierphasenmethode um den Faktor 2.5 (bei 35 MHz Modulationsfrequenz) bis 7 (bei 20 MHz) überlegen ist. Die Berechnungsvorschrift bleibt dabei einfach und ist invariant gegenüber Schwankungen des Gleichanteils und linearer Verstärkung einzelner Pixel. Weiters verringert sich die Systemkomplexität und höhere Frameraten sind möglich. Es konnte keine signifikante Erhöhung der Varianz festgestellt werden.

Hinsichtlich weiterer Reduktion des Wiggling Errors wurde der *adaptive super-resolution* (ASR) Algorithmus vorgestellt. Durch Kombination zweier aufeinanderfolgender Aufnahmen kann die spektrale Auflösung verdoppelt werden, was nur unwesentlich erhöhte Systemkomplexität nach sich zieht. Aliasing-Artefakte können dadurch um den Faktor 9 (bei 5 MHz) bis 34 (bei 20 MHz) reduziert werden, wobei zweiterer Fall einer vollständigen Eliminierung des Wiggling Errors gleichzusetzen ist.

Abstract

Time-of-flight (ToF) is an imaging procedure that allows the acquisition of three-dimensional scenes based on propagation-delay measurements. Whilst having several advantages over competing technologies, one drawback is the potential inaccuracy of delivered depth information. In particular, several systematic errors exist that have been subject to intensive research for several years now. A general agreement exists in literature that these errors can only be overcome by extensive calibration.

With a state-of-the-art ToF sensor at hand, a thorough analysis of systematic depth deviations was conducted from a signal-processing perspective. Due to an extended measurement setup, seminal new insights into the nature of these errors were gained. For example, the amplitude-related error was shown to be a result of multipath reflections. No evidence for sensor-based deviations depending only on the optical signal amplitude was found. In turn, and contradictory to state-of-the-art approaches, calibration with respect to amplitude values is no suitable countermeasure. In fact, new strategies towards compensating multipath reflections must be found in order to cope with deviations of this kind. Furthermore, the integration-time related error was observed not to be an issue of the sensor either but most likely the result of additional delays introduced by the illumination unit due to warm-up effects.

The most important outcome of this Thesis however is the proposal of several strategies that significantly decrease the Wiggling Error (also known as circular distance-error). With consequent alignment between theory and practice, it was thoroughly shown that taking four samples results in a worst-case Wiggling Error since massive aliasing distortions occur. In fact, it was shown that distance estimation based on *three* equidistant samples outperforms the four-phase algorithm by a factor of 2.5 (at 35 MHz modulation frequency) to 7 (at 20 MHz). The estimation formula remains simple and is invariant to offset and linear gain of the pixels. Furthermore, the complexity of the system is reduced and increased frame rates are possible. No significant increase of the variance was observed.

Towards further reduction of the Wiggling Error, the *adaptive super-resolution* (ASR) algorithm was proposed. With only low additional complexity, it allows doubling the spectral resolution by taking information from two consecutive frames. This effectively reduces aliasing distortions between a factor of 9 (at 5 MHz) up to 34 (at 20 MHz), of which the latter case can be considered as a complete elimination of the Wiggling Error.

Acknowledgements

I would like to sincerely thank Dipl.-Ing. Dr. Christoph Steiner at Infineon Technologies Austria for supervising my Thesis, his profound feedback in countless discussions and his active support of my ideas. Furthermore, I am thankful to Assoc.Prof. Dipl.-Ing. Dr. Klaus Witrisal for taking the time and assessing my Thesis, for the constructive meetings and valuable impulses.

I would like to acknowledge the company Infineon Technologies Austria, in particular Dipl.-Ing. Dr. Michael Mark, for the pleasant working atmosphere and the chance to deeply investigate this fascinating topic in the course of my Thesis.

I would like to express my sincere gratitude to my family, in particular to my parents Christa and Konrad Grünwald, who have laid the foundation stone to a wonderful life full of exciting challenges. I am very fortunate to have shared my childhood – and the time beyond that – with my sisters Katrin, Reingard, and Gertraud.

Together with Eva Schiller, who will be acknowledged in private, I am looking forward now to exploring new horizons in South America.

Johannes Grünwald
Graz, January 24, 2013

Nomenclature

ADC	Analog-to-digital converter
ASR	Adaptive super-resolution
CW	Continuous wave
DFT	Discrete Fourier transform
FoM	Figure of merit
FPN	Fixed-pattern noise
FS	Fourier series
LED	Light-emitting diode
LUT	Look-up table
PMD	Photonic mixing device
RMSE	Root-mean-square error
RoI	Region of interest
SNR	Signal-to-noise ratio
SRF	Sensor response function
STD	Standard deviation
ToF	Time-of-flight

Contents

— CHAPTER ONE —

Introduction

1.1 Scenario

About fifteen years after the invention of 3D depth sensing based on the *time-of-flight* (ToF) principle, cameras that implement this technology are still subject to intensive research and development. Recently produced ToF sensor reflect the technological progress as they continuously increase performance whilst already featuring remarkably small package sizes.

One of the obligatory steps between sensor prototype and final product is characterization and, depending on the application, calibration. For ToF imagers, the latter is known to be a tedious process of several minutes per sample due to the fact that a multitude of systematic depth errors exists. In the current context this fact becomes particularly problematic as the manufactured imagers are targeted at the mass market. This potentially requires volumes in an order of magnitude that makes any sample-based calibration method taking more than a few seconds unfeasible.

1.2 Motivation

Basically, calibration is the process of compensating systematic measurement deviations via knowledge gained from reference data. Depending on the level of understanding relationships and root causes of the respective error, more or less reference data is required. As a simple example, if some linear dependency between temperature and depth deviation can be proven, this only requires the determination of two parameters in the calibration process, i.e., offset and gain. In turn, a linear compensation model can be established by means of these parameters. On the other hand, without any knowledge of this linearity one would need to acquire reference data over the whole operational temperature range in order to compensate deviations. Thus, generally speaking calibration effort can be reduced by increasing the structural knowledge of systematic errors – which is exactly the key motivation of this Thesis.

Investigating systematic measurement deviations can have appreciated side effects. For example, if root causes can be revealed as being implementation-specific, this is a valuable feedback for the hardware design team. It is then a matter of discussing theoretic and practical aspects in order to elaborate solutions to this issue. In the best case, errors that require extensive calibration

in the first place can be turned into ones that can be compensated in an entirely algorithmic manner.

1.3 About this Thesis

Conducted as a collaboration between Infineon Technologies Austria AG and the Signal Processing and Speech Communications Laboratory (SPSC) at Graz University of Technology between April 2012 and January 2013, this Master's Thesis concludes the Author's studies of Telematics, with specialization in digital signal processing and system-on-chip design.

The structure of this Thesis follows the logical process of the undertaken investigations. First, a literature study was carried out with emphasis on systematic errors and calibration approaches in the context of ToF imaging. Thus, as an introductory step towards the topic, Chapter 2 provides the quintessence of this study, i.e., a concise overview of the technology and a short discussion of its advantages and disadvantages. In a next step, the mathematical framework was derived, as presented Chapter 3. The experimental part of this Thesis is covered by Chapter 4 and Chapter 5, where in the first the measurement setup and the experiments to be made are described, and in the second the results are discussed. In Chapter 6 several approaches that reduce systematic errors are presented and evaluated. Finally, a summary of the Thesis and an outlook on further investigations are given in Chapter 7.

— CHAPTER TWO —

Time-of-Flight Technology in a Nutshell

In this chapter, the general principle of time-of-flight imaging is outlined. First, an overview and basic relations are given in Section 2.1 with a focus on the key feature of ToF imaging sensors, the *photonic mixing device* (PMD). Main advantages and disadvantages compared to competing 3D-sensing technologies are covered in Section 2.2 and Section 2.3. The major part of this introduction however focuses on measurement deviations of all kinds that have been elaborated in literature so far. This should provide a solid basis for the practical examination of systematic depth errors. Finally, this chapter concludes with some recommendations for further reading in Section 2.4.

2.1 Overview

Time-of-flight is an imaging technology that allows the acquisition of depth information. Basically, this is done by literally measuring the *time of flight* of an emitted (and then reflected) light beam. The distance d to the targeted object is directly related to this time delay t_d via

$$d = \frac{1}{2}\,\mathrm{c}\,t_d \tag{2.1}$$

where $\mathrm{c} = 3 \cdot 10^8\,\mathrm{m/s}$ is the speed of light and the factor of $1/2$ accounts for the fact that the distance is covered twice by the light beam. Besides several possibilities to implement this principle, one approach (and the one in focus of this work) is to modulate the emitted signal with a continuous wave (CW) of frequency f_{mod} and estimate the resulting phase shift φ_d. In particular, this phase shift is measured between the *optical* signal after it is received back at the sensor, and the *reference* signal of same shape but zero delay. If this phase shift can determined, time delay and distance can then be reconstructed as

$$t_d = \frac{\varphi_d}{2\pi f_{\mathrm{mod}}} \quad \Rightarrow \quad d = \frac{\mathrm{c}}{4\pi f_{\mathrm{mod}}}\,\varphi_d \tag{2.2}$$

within a certain ambiguity range of $\mathrm{c}/(2f_{\mathrm{mod}})$ due to the inherent 2π-periodicity of phase information. It is now evident that *distance* estimation refers to *phase* estimation in this context, which is basically the main objective of ToF imaging with CW-modulated signals. For better illustration, the basic principle of ToF is sketched in Fig. 2.1.

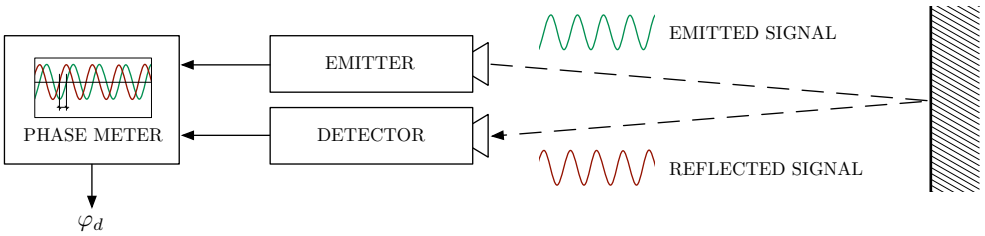


Figure 2.1: ToF imaging principle (cf. Castañeda [1]).

In Tab. 2.1, some characteristic properties of the examined ToF system are provided. It should be remarked that no claims regarding depth resolution and accuracy can be made yet, as these figures are still being evaluated.

Property	**Value**
Distance range	up to 50 m
Modulation frequency	5 - 35 MHz
Signal shape	rectangular
Lateral resolution	288 × 352 pixels
Frame rate	up to 80 fps
ADC resolution	12 Bit

Table 2.1: Characteristic properties of the investigated ToF imager.

2.1.1 The Photonic Mixing Device

One main feature of ToF imaging with CW-modulated signals is the fact that the essential part of the signal processing chain is implemented on the pixel itself: The Photonic Mixing Device (PMD). This pixel has the extended functionality of *mixing* optical signal and reference signal, which will be described more in detail now. The basic principle is sketched in Fig. 2.2.

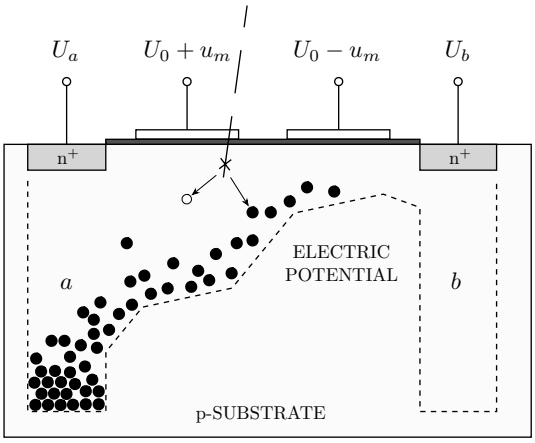


Figure 2.2: PMD pixel overview (cf. Schneider [2]).

Within the light-sensitive poly-silicon layer, incident photons are split up into electron-hole pairs at a certain depth proportional to the wavelength of the light. This process is also referred to as the inner photoelectric effect. In order to measure the intensity of incoming light, the generated electrons have to be separated and collected before they can recombine. This is achieved by a CW-modulated drift field that is generated by two (or more) gate electrodes. Generally, these electrodes are driven in differential manner (denoted as $U_0 + u_m$ and $U_0 - u_m$) in order to construct drift-fields towards the so-called *buckets* a and b. The collected charge carriers in a and b can then be read out after a certain integration time. In order to increase the signal-to-noise ratio (SNR), integration is performed over many periods. It should be remarked that PMD pixel-structures were investigated extensively during the last years. As a consequence, there are publications available that provide an excellent and profound analysis of PMD sensors, far beyond the scope of this Thesis [2–5].

In the context of this work however, one essential relationship between the PMD pixel and signal processing should be understood. It is a perspicuous interpretation that for modulation signals of rectangular shape, in the ideal case the voltages U_a and U_b generated by the accumulated charges refer to samples of the cross-correlation function between incoming light (optical signal) and gate voltages (reference signal). Therefore, it is intuitive that the argument that maximizes this function refers to the phase offset between these two signals, which is in turn the quantity to be measured. For the complete derivation of depth estimation within a mathematical framework, it is referred to Chapter 3.

2.2 Advantages

Classical 3D-vision systems traditionally suffer from several shortcomings that limit their field of application. Mostly, these limitations directly result from the functional principle of the system and are difficult (if not impossible) to compensate. For example, stereo-vision systems need a sufficiently long baseline between the cameras in order to reconstruct the scene with certain accuracy. Or, laser-based systems need mechanical parts as the laser captures the whole scene in a point-wise, sequential fashion. As a short overview, the main benefits of ToF imaging with respect to competing depth-sensing methods are listed below [1, 6].

- Depth information is readily available and need not be reconstructed (no feature point detection or depth data registration required).
- Estimation performance does not rely on scene textures.
- No mobile parts within the imager are required.
- Depth information of each pixel is acquired in parallel, allowing high frame rates.
- Form factor is comparatively small.
- No passive illumination of the scene is necessary.

2.3 Disadvantages

In contrast to the promising advantages of ToF imaging outlined in Section 2.2, there exist several drawbacks of this technology as well. They will be shortly discussed in this section. Fortunately, most of the shortcomings can be attributed to the fact that ToF imaging sensors are still subject to intensive development and by no means a fully mature technology. It can be expected that over the years to come, seminal new insights and improvements will be published.

2.3.1 General Shortcomings

Low Resolution

The relatively high complexity of each PMD pixel is a limiting factor regarding spatial resolution of the sensor. During the last years a massive increase of ToF sensor resolution has been observed, but current realizations are still far away even from VGA. Precisely speaking, state-of-the-art imagers comprise about 100.000 pixels, which however is still not comparable to resolutions provided by even cheap RGB cameras.

Measurement Deviations

Raw ToF data itself without correction is potentially unreliable because a multitude of systematic errors exists: Sensor nonlinearities, signal degeneration, process inhomogeneities, temperature drifts, just to mention a few. Besides these systematic deviations that can be identified and (at least partly) compensated, unpredictable errors exist as well, comprising statistical errors (such as noise) but also errors that are more difficult to handle, such as flying pixels, motion artifacts or multipath inferences.

Limited Working Range

The periodicity of the modulation signal inherently limits the working range of the camera to a certain ambiguity range, which was already discussed in the context of (2.2). As a remedy, multiple acquisitions with different modulation frequencies can extend this range. Anyways, the final limit is defined by the power of the light source.

Low Dynamic Range

A big weakness of current ToF systems is their low dynamic range. This is evident from the fact that PMD pixels are linearly sensitive to incident light, whereas the power of the emitted optical signal decreases with the square of the distance [7]. Therefore, one has to deal either with saturation or with high noise due to low amplitudes. One solution to this problem would be the realization of different integration times for each frame, but this comes at the cost of a frame-rate decrease.

Heat

One issue that always arises with electronic devices of comparably high power consumption (such as light-emitting diodes (LEDs)) that are commonly used as illumination source) is heat emission. For highly integrated systems, the illumination unit will be close to the sensor, introducing considerable temperature influences.

2.3.2 Systematic Errors

In this subsection, known systematic deviations that occur in ToF sensing are outlined. It should be noted that the state of investigation differs between these issues, ranging from good understanding to rather vague interpretations. As well, the contents discussed in this chapter just represent outcomes of the literature study carried out in advance; in context of the examined imager, these errors are discussed in Chapter 5.

Wiggling Error (Circular Distance-Error)

One of the best understood but most challenging systematic deviations in ToF imaging is the *Wiggling Error* (also referred to as circular distance-error) whose name stems from its characteristic shape, which is periodic with respect to the target distance. The Wiggling Error originates from the fact that the cross-correlation function contains harmonics of higher order whereas the depth estimation algorithm assumes a plain sinusoidal signal. Its compensation is still difficult though, because the exact shape of the correlation signal is not known. Therefore, the error has to be calibrated over the whole distance range which necessitates extensive calibration setups.

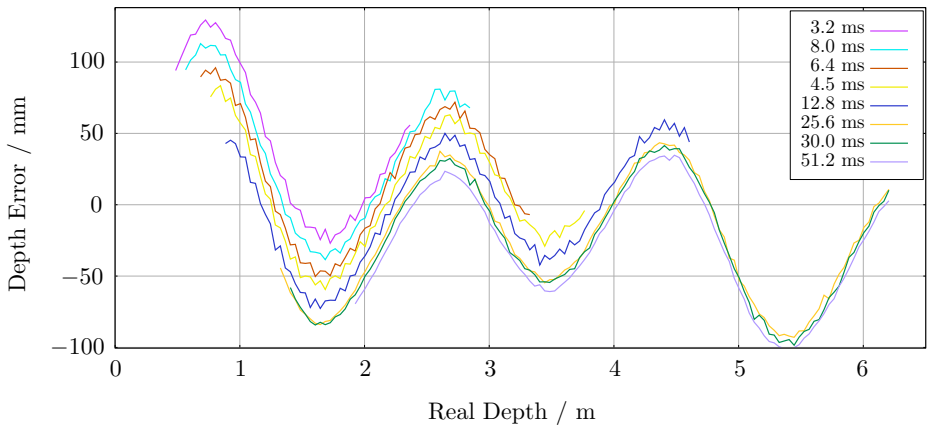


Figure 2.3: Wiggling Error for different integration times (MESA Imaging SwissRanger 3000), evaluated by Rapp [7].

In Fig. 2.3, an exemplary plot of the Wiggling Error, measured for different integration times, is shown. Intensive investigations of commercially available ToF imagers with respect to this kind of error were performed by Müllenhaupt [8] and Rapp [7]. Quantitatively speaking, for modulation frequencies around 20 to 30 MHz, the deviations amount up to 15 cm.

Amplitude-Related Error (also Intensity-Related Error)

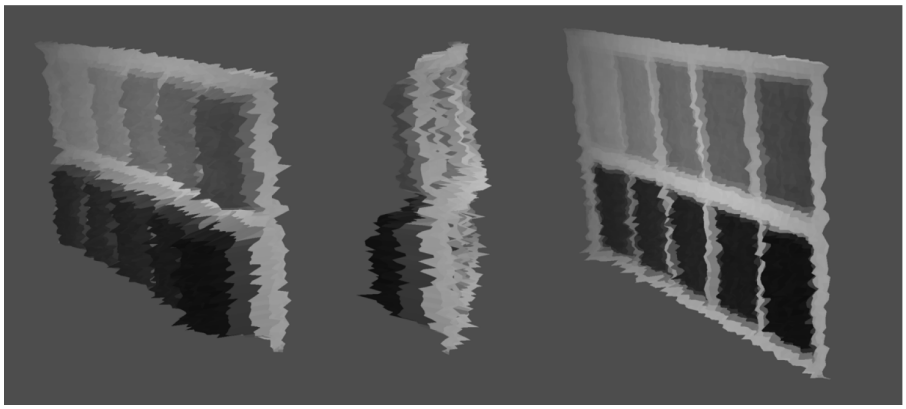


Figure 2.4: Exemplary target that demonstrates the amplitude-related related error, before (left, middle) and after (right) calibration [9].

Another systematic error that is well known but hard to correct is the so-called amplitude-related error (sometimes also referred to as intensity-related error). In short terms, it causes dark surfaces to be measured at different distances as bright ones. The root cause for this error, which is illustrated in Fig. 2.4, is not fully understood yet. In their laboratory, Lindner et al. [9] identified the order of magnitude as several centimeters at a modulation frequency of 20 MHz. As a countermeasure, calibration is suggested in literature.

Integration-Time Related Error

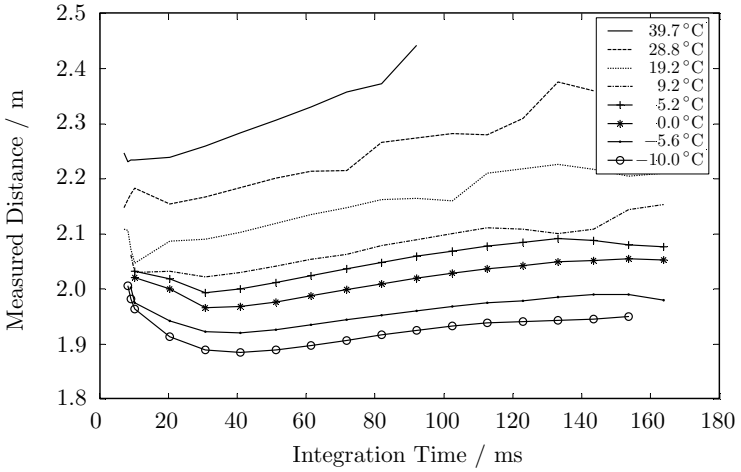


Figure 2.5: Depth deviation with respect to integration time and temperature [10].

Increasing integration time naturally improves the SNR, but also introduces an offset to the measured depth. Fortunately, this error appears to be small and almost constant over the whole distance range, which should simplify compensation significantly. Fig. 2.3 and Fig. 2.5 suggest that deviations in the range of 5 to 10 cm must be considered. However, no analytic dependencies have been derived yet.

Temperature-Related Error

As for all electronic devices, temperatures can cause severe deviations to measurements. Besides the fact that the sensor should warm up long enough before putting it into operation, in steady-state, changes of ambient temperature cause measurement deviations as well. This effect is delineated in Fig. 2.5 which impressively shows that within a working range between $-10\,°C$ and $+40\,°C$, obtained depth values differ up to 50 cm from each other. Fortunately, temperature seems to affect all values in the same way, so it may be compensated in a similar manner as errors caused by different integration times.

Pixel-Related Error (Fixed-Pattern Noise)

For several reasons listed below, each pixel exhibits a characteristic performance that differs from other pixels. This pixel-related error is also referred to as *fixed-pattern noise* (FPN)[1] and generally compensated via pixel-based look-up tables (LUTs). Müllenhaupt [8] showed that FPN can be expected in an order of magnitude of a few centimeters. Obvious root causes for FPN in ToF imaging are

- Path lengths. As ToF bases on measuring propagation delays, deviations in path lengths directly affect measured distances.
- Linear offset/gain. Each pixel potentially offsets or linearly amplifies the measured signal in a characteristic manner. For reasonable phase estimation algorithms, these effects should not influence distance estimation at all.
- Nonlinearites. Besides the linear component above, nonlinear effects may appear as well.

2.3.3 Statistical Errors

Besides deviance that exhibits deterministic dependence on system- or scene parameters, every measurement system has degraded output quality caused by random, non-deterministic influences. In this section, a brief overview regarding these statistical errors is given.

Phase Noise

One major outcome of noise analysis is the fact that the relationship between phase variance and measured intensity of the optical signal is reciprocally proportional and quadratic. This was thoroughly derived by Plaue [11] and Rapp [7]. As a consequence, in a first attempt amplitude values may be used as confidence measures for depth values.

Closely coupled to the finding that phase variance and intensity are inversely proportional was the analytic derivation of the phase distribution. Given the fact that the generation of electron-hole pairs follows a Poisson process, in the limit of accumulating many charges, the resulting phase noise approximately follows a Gaussian distribution. Furthermore, error propagation calculus

[1] Precisely speaking, literature distinguishes between fixed-pattern noise that deals with pixel-dependent deviations in the amplitude domain (such as pixel gain, offset, etc.) and fixed-pattern *phase*-noise (FPPN) that accounts for pixel-dependent phase deviations. However, in the course of this Thesis a differentiation between these two terms was not found necessary, so the more general term FPN will also include FPPN.

gave Frank et al. [12] the insight that finally, the phase follows a normal distribution which is offset by the true value.

Jitter

Although derived from the same oscillator, optical signal and reference signal potentially loose their rigid phase synchronization after passing through different parts of the circuit. This *jitter* can be observed as non-deterministic fluctuations of their relative phases, introducing a random component directly to the phase estimation. As jitter can be modeled as a stationary zero-mean process, it can be reduced via averaging.

Quantization Effects

Frank [12] focused his analyses on quantization effects that should not be underestimated. As will turn out more clearly in Chapter 3, state-of-the-art distance estimation algorithms rely on computing the argument (i.e., the phase) of a complex number. If the amplitude values are quantized, this can lead to severe artifacts for low signals.

2.3.4 Other Sources of Depth Measurement Degradation

There are further problems ToF imaging suffers from. They may be characterized as mostly scene-dependent, which makes their correction very difficult because scene parameters are challenging to model. Therefore, in most cases these errors are compensated via appropriate post-processing filters.

Motion Artifacts

Motion artifacts originate from the fact that image sensors need a certain amount of time until their acquisition of the scene is finished. Concerning ToF sensors, one of these lower limits is the integration time each frame necessitates. If the scene changes during the capturing process, artifacts occur.

Flying Pixels

Due to the relatively low lateral resolution of current ToF sensors, one imaged pixel covers a comparatively big area in 3D space. If this area has homogeneous depth, no problems occur. Troubles arise at boundaries of objects, where background and foreground are potentially mixed into one pixel, causing *flying pixels* [13] which are named after their appearance somewhere between object and background.

Multipath Reflections

Imaging systems are in principle capable of separating desired from undesired information via appropriate optical systems. However, it is still possible that several optical paths from illumination unit to sensor pixel can be drawn. Generally, this is a very natural optical process and not considered as problematic in conventional imaging because it is perceived by humans

in the same manner as by synthetic imaging sensors. The reason why it is unwanted in ToF imaging is because different light rays cover different distances, thus altering depth measurement. It is evident that this is a very challenging issue to overcome [14].

Light Scattering

Internal light scattering occurs when light rays are reflected between sensor plane and optics, within the imager. In its nature comparable to multipath reflections, this effect introduces lateral crosstalk that potentially contaminates distance information as well. Karel [15] performed a detailed analysis and managed to identify light scattering as an additive, linear process. However, he admitted that for proper compensation of the process further studies are necessary.

Sunlight, Incident Light

Sunlight is always a competing light source for active vision systems. This naturally applies for ToF imaging as well: Any infrared light source adds a positive offset to the incoming optical signal, decreasing the remaining dynamic range of the sensor. Consequently, an increase of shot noise is introduced as well. While on-chip solutions already exist that compensate for this unwanted offset attributed to external light sources (*suppression of background illumination* (SBI) [16, 17]), sunlight is still a problem because in outdoor scenarios it exceeds the optical power of the illumination unit by orders of magnitude.

2.4 Further Reading

As this short introduction aimed at communicating basic principles and selected topics which are relevant for this Thesis, it is referred to literature for information that goes beyond that. Amongst others, Foix [6], Kolb [18], and May [13] published excellent introductory material. Lange [3], Schmidt [4], Schneider [2], and Spickermann [5] elaborated technological fundamentals of the PMD more in detail. Frank [12], Müllenhaupt [8], Plaue [11], and Rapp [7] focused in particular on different shortcomings of ToF imaging. In addition to that, several remedies for deviations in ToF imaging were published by Beder [19], Frank [20], Fuchs [21, 22], Kahlmann [10], Lichti [23, 24], and Lindner [9, 25].

— CHAPTER THREE —

Signal Processing Framework

This chapter focuses on establishing a consistent mathematical framework that models the examined ToF imaging system. The major objective was to keep it as simple as possible but as complex as necessary. The signal path is discussed in Section 3.1 and as an extension to Section 2.3, expected real-world difficulties are linked to the respective system components. The overall system output, termed *sensor response function* (SRF), is introduced and elaborated more in-depth in Section 3.2. Finally, in Section 3.3 theoretic foundations of particular issues are established, which is needed for further discussions in the experimental part of this Thesis.

3.1 Signal Path

3.1.1 Overview and Basic Derivation

In Fig. 3.1, the simplified signal path is sketched as a block diagram.

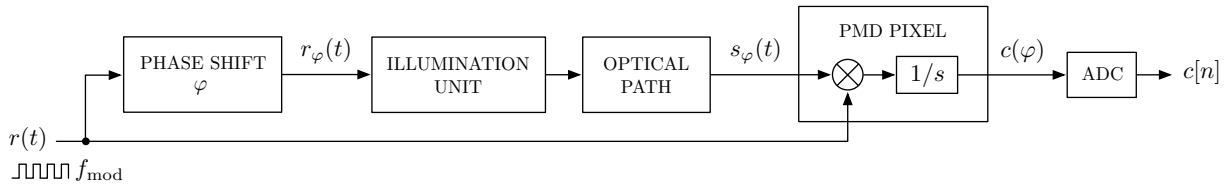


Figure 3.1: Schematic signal path of a ToF imaging system.

With $r(t)$ and $s_\varphi(t)$ fed into the pixel as reference signal and optical signal, a first mathematical approximation of the pixel's behavior is to see it as a correlator of these signals,

$$c(\varphi) = \int_0^\tau s_\varphi(t) r(t) \mathrm{d}t \tag{3.1}$$

where τ is chosen such that integration is performed over an integer multiple of periods. As well, without loss of generality, it is assumed that integration starts at zero. In order to parameterize this function, a *reference phase shift* can be applied to the optical signal[2], denoted by the subscript φ. Sticking to the phase domain in this context maintains independence of the modulation

[2] It should be noted that it is fully equivalent to apply this phase shift to the reference signal instead. Therefore, shifting the optical signal was motivated by the implementation of the examined ToF imager.

frequency, which is an appreciated side effect. As will be referred frequently to $c(\varphi)$, it will be termed *sensor response function* (SRF) throughout this Thesis.

Towards illustrating the principle of distance estimation, the optical path may be modeled as a linear time-invariant (LTI) system with ideal delay, i.e., with an impulse response

$$h_{\text{opt}}(t) = g(d)\delta(t - t_d) \tag{3.2}$$

where $\delta(\cdot)$ refers to the *Dirac delta function*. The function $g(d)$ reflects the signal attenuation according to the target distance. Recalling the *Inverse-square law* from physics, the optical power of a signal emitted by a point light source is reciprocally proportional to the distance, as long as additional absorption effects caused by the medium can be neglected. Since the current signal model features amplitude signals, this knowledge consequently translates into the relation $g(d) \propto 1/d$. Furthermore, t_d refers to the time delay accounting for the object distance. This rather simple model of the optical channel is basically valid, although it does not include multipath reflections that can cause severe depth deviations in real-world scenarios (cf. Section 2.3.3).

The illumination unit may be simplified as

$$h_{\text{ill}}(t) = \delta(t) \tag{3.3}$$

which, in the first place, ignores distortions due to temperature effects and the general low-pass character of LEDs. With the assumptions of (3.2) and (3.3) put together, it turns out after some basic signal processing calculus that

$$s_\varphi(t) = g(d)r_\varphi(t - t_d) \tag{3.4}$$

and, inserted in (3.1),

$$\begin{aligned} c(\varphi) &= g(d) \int_0^\tau r_\varphi(t - t_d)r(t)\mathrm{d}t \\ &= g(d) \int_0^\tau r\left(t - t_d - \tfrac{\varphi}{2\pi f_{\text{mod}}}\right) r(t)\mathrm{d}t \end{aligned} \tag{3.5}$$

where in the second step the reference phase shift was put into the argument. Now it is evident that this expression maximizes if t_d and $\frac{\varphi}{2\pi f_{\text{mod}}}$ cancel out since this yields an integral over $r^2(t)$. Consequently, requiring that

$$t_d \stackrel{!}{=} -\frac{\varphi}{2\pi f_{\text{mod}}} \tag{3.6}$$

the task is to find

$$\varphi_d = -\operatorname{argmax}\{c(\varphi)\} \tag{3.7}$$

from which the distance d can be reconstructed according to (2.2). In the course of this Thesis, the term *argmax-estimation* will directly refer to this process, i.e., to estimating the argument that maximizes the SRF.

Furthermore, it should be recognized that phase-shifting the optical signal is actually equivalent to *sampling* the SRF. For a predefined set of reference phase shifts $\varphi_n,\ n \in \{0, \dots, N-1\}$, this leads to the definition of the parameter- and amplitude-discrete SRF

$$c[n] = Q_B\left(c(\varphi)\right)\big|_{\varphi=\varphi_n} \tag{3.8}$$

after passing through the analog-to-digital converter (ADC). $Q_B(\cdot)$ denotes ideal B-bit quantization from continuous domain to 2^B digital values, which are read out from the chip.

At this stage, the term *N-point sampling* is defined. It refers to N samples that are taken from the SRF at equidistant steps of $2\pi/N$. This process will be represented by the variable φ_n, i.e.,

$$\varphi_n = \frac{2\pi}{N} n \quad \text{with} \quad n \in \{0, \dots, N-1\}\,. \tag{3.9}$$

3.1.2 Real-World Difficulties

The model presented in the preceding section is of course highly idealistic and disregards effects that are considerably challenging in real-world circumstances. On the basis of the discussions in Section 2.3, a closer link between system components and expected problems shall be provided in this subsection.

Phase Shifting Unit: Jitter

Due to the observation that reference phases φ effectively translate into sampling points n (cf. (3.8)), any deviations of real reference phases from desired phases can be seen as sampling point offsets. It is evident that systematic deviations are impossible to reconstruct without further knowledge of the deviation itself. Jitter however, which is supposed to be a stationary zero-mean random process, can be diminished via averaging. This can be shortly proven by a *Taylor series expansion* around the desired sampling point φ_n, where its actual position is offset by a random component Δ_φ:

$$c(\varphi_n + \Delta_\varphi) \approx c(\varphi_n) + \sum_{k=1}^{K} \frac{\Delta_\varphi^k}{k!} \frac{\mathrm{d}c^k(\varphi)}{\mathrm{d}\varphi^k}\Big|_{\varphi=\varphi_n}\,. \tag{3.10}$$

If Δ_φ is small compared to the shape of the transfer function (which is a reasonable assumption), the linear term of the Taylor series is dominant, i.e.,

$$c(\varphi_n + \Delta_\varphi) \approx c(\varphi_n) + \Delta_\varphi \frac{\mathrm{d}c(\varphi)}{\mathrm{d}\varphi}\Big|_{\varphi=\varphi_n} \tag{3.11}$$

where it can be seen that at least in this approximate case, jitter artifacts can be averaged out. Furthermore, it is evident that the influence of jitter that is effectively seen as amplitude noise is directly proportional to the derivative of the function.

Illumination Unit: Low-Pass, Temperature Drifts

The illumination circuit will certainly exhibit performance deviations with respect to temperature and frequency. It can be expected that the LEDs will not be able to perfectly well transform the driving signal into the optical domain over the whole desired frequency range; in short, low-pass behavior will be observed. As well, due to considerable heat emission of LEDs, drifts caused by temperature changes are likely to occur.

Optical Path: Multipath Reflections

The problem of multipath reflections was already introduced in Section 2.3.3. From a signal processing perspective, it can be safely assumed that the optical path exhibits linear behavior, allowing additive modeling of superimposed rays. A logical extension of the impulse response of the optical channel (3.2) would then be

$$h_{\mathrm{opt,MR}}(t) = \sum_i g(d_i)\delta(t - t_{d_i}) \tag{3.12}$$

where the index i refers to the respective reflection path. In particular, d_i denotes the actual distance and t_{d_i} refers to the resulting time delay of this component.

PMD Pixel: Fixed-Pattern Noise

Given the three major scenarios for FPN deduced in Section 2.3.2, i.e., different path lengths, linear gain/offset, and nonlinear distortion, the argmax estimation-rule defined in (3.7) allows some further elaboration of the respective impacts. First of all, as different path lengths introduce additional delays that cannot be distinguished from any other delay (such as within the optical path), this will remain an issue that cannot be mitigated without prior knowledge of the architecture, or calibration. Secondly, any linear gain or offset of the SRF caused by individual pixel characteristics will not influence distance measurement, because the true maximum of the SRF remains at the same position.

Probably the most important feature of argmax-estimation is its invariance to any static nonlinearity (denoted by $f(x)$ for now), as long as this nonlinearity is bijective. In practical terms this means that the nonlinear mapping must be either strictly increasing or decreasing, which is a reasonable requirement. Considering this nonlinearity now as strictly increasing (otherwise the following statements would just be inverted), it is evident that

$$c(\varphi_1) > c(\varphi_2) \iff f(c(\varphi_1)) > f(c(\varphi_2)) \quad \forall \varphi_{1,2} \in [0, 2\pi],\ f(\cdot) \text{ strictly increasing} \tag{3.13}$$

which implicitly confirms the aforementioned assumption.

Analog-to-Digital Converter (ADC): Quantization Effects

Two general issues of the conversion from continuous to discrete values are quantization effects due to finite word lengths and nonlinearities. Quantization effects can be usually modeled as zero-mean, additive, white noise with uniform distribution between $\pm\frac{1}{2}$ LSB and an expected

noise power of $P_N = 1/12$ LSB^2 [26, pp.238–249]. Since the data word length of the investigated imager is 12 Bit, the power of an exemplary full-scale sinusoid (thus with amplitude 2^{11} LSB) is $P_S = (2^{11})^2/2 = 2^{21}$ LSB^2, which leads to an SNR of

$$\mathrm{SNR} = 10\log_{10}\left(\frac{P_S}{P_N}\right) = 10\log_{10}\left(12 \cdot 2^{21}\right) = 74.0\,\mathrm{dB}\,. \tag{3.14}$$

This simple but realistic example shows that regarding quantization noise, no severe degradation of the SRF need be expected as long as the signal amplitude is maintained high enough. Still, for signal amplitudes around 100 LSB an SNR of about 37.0 dB can be expected.

Regarding nonlinearities, the same principles hold for the ADC as for PMD pixels (cf. discussion above): Ideal argmax-estimation is invariant to static nonlinearities that can be modeled as strictly increasing functions.

3.2 The Sensor Response Function (SRF)

As the SRF will be in the spotlight of the experimental part of this Thesis, this chapter is dedicated to some preliminary analyses of its characteristics and further implications on depth estimation.

3.2.1 Basic Assumptions

First of all, as the SRF refers to the correlation function of periodic signals with the same fundamental frequency, it is evident that the SRF is 2π-periodic itself. Furthermore, from the fact that $r(t)$ is of rectangular shape it can be expected that the SRF will have recognizable components of a triangular function, depending on the degeneration of the optical signal. Therefore, as the fundamental wave will dominate the shape of the SRF, it is a straight forward observation that the true maximum of the SRF will be close to the maximum of the fundamental wave (still lacking the precise definition of *close* in this context, however, the phase offset will be small compared to the full period of 2π).

3.2.2 Fourier Series Representation

One reasonable model for representing periodic functions is the *Fourier series* (FS). In its most general form it is the superposition of weighted harmonic complex exponentials:

$$c(\varphi) = \sum_{k=-\infty}^{\infty} C[k]\mathrm{e}^{\jmath k\varphi} \tag{3.15}$$

where $C[k] \in \mathbb{C}$ are the *Fourier coefficients* that contain the complex weight of the k^{th} harmonic. Due to the evident 2π-periodicity of the SRF, its fundamental frequency is simply $\omega_0 = 1$. In turn, this allows direct replacement of the temporal parameter t (as known from the conventional form of the FS) by $\varphi = \omega_0 t = t$. This is motivated by the fact that t is of rather abstract character in this context.

As a motivation, the advantages of this representation shall shortly be mentioned:

- Fourier analysis is a well developed discipline in signal processing and inherently reveals the spectral components of the signal.
- The coefficients can be computed from finite, equidistant sample sets via the *discrete Fourier transform* (DFT)[3], which is the *minimum-variance unbiased* estimator [27, pp.88–90]. Coefficients determined by the DFT will be denoted by $\hat{C}[k]$.
- Although estimated from a finite sample set via the DFT, Fourier coefficients allow the SRF to be *interpolated*. It is evident that this is an important feature for argmax-estimation.

The fact that $\hat{C}[k]$ are determined from a finite sample set has critical implications from sampling theory. First of all, as resulting from the *Nyquist–Shannon sampling theorem*, it must be taken into account that in order to perfectly reconstruct a continuous signal of certain bandwidth by a finite sample set, the according sampling rate must exceed twice the signal bandwidth. In the current context, this has the consequence that for N-point sampling, only harmonics $k < N/2$ can be resolved. Moreover, harmonics that exceed $N/2$ will disturb lower ones, an effect called *aliasing*. This effect is well understood and documented in literature [26, pp.185–190].

In order to gain better insights into the relation between the true Fourier coefficients $C[k]$ and those computed via the DFT, $\hat{C}[k]$, it is a good idea to start with the definition[4] of the N-point DFT,

$$\hat{C}[k] = \frac{1}{N} \sum_{n=0}^{N-1} c[n] \mathrm{e}^{-\jmath \frac{2\pi}{N} nk} \,. \tag{3.16}$$

Disregarding quantization effects for now, the signal $c[n] = c(2\pi n/N)$ (3.9) may be expressed via its FS (3.15)

$$\begin{aligned} c[n] &= \sum_{k=-\infty}^{\infty} C[k] \mathrm{e}^{\jmath \frac{2\pi}{N} nk} \\ &= \sum_{k=0}^{N-1} \sum_{m=-\infty}^{\infty} C[k+mN] \mathrm{e}^{\jmath \frac{2\pi}{N} n(k+mN)} \\ &= \sum_{k=0}^{N-1} \left(C[k] + \sum_{m=1}^{\infty} (C[k+mN] + C[k-mN]) \right) \mathrm{e}^{\jmath \frac{2\pi}{N} nk} \end{aligned} \tag{3.17}$$

where in the second step the infinite sum over k was split up into a cascade of one finite sum over k and one infinite sum over m. Lastly, the sum over m was rewritten and the additive mN in the argument of the complex exponential was removed as $\mathrm{e}^{\jmath 2\pi mn} = 1 \; \forall m, n \in \mathbb{N}$ has no effect. Finally, (3.17) arrives at a form which is known as *inverse discrete Fourier transform* (IDFT)

[3] To be precise, the estimation of sinusoidal parameters from a finite sample set of *periodic* functions is realized by the discrete Fourier *series* (DFS). But as the more popular DFT can be, loosely speaking, seen as a generalization of the DFS, it will be the tool of choice here.

[4] In order to maintain consistency with (3.15), the scale factor $1/N$ already comes into play here. This lets the current definition differ slightly from the original one, however, it is a difference of only formal nature.

which can be shown to be the inverse of (3.16), thus

$$\hat{C}[k] = C[k] + \underbrace{\sum_{m=1}^{\infty} (C[k+mN] + C[k-mN])}_{\text{Aliasing}} . \tag{3.18}$$

This equation relates the *true* Fourier coefficients $C[k], k \in \mathbb{Z}$ to the *observed* Fourier coefficients $\hat{C}[k], k \in \{0, \dots, N-1\}$. In fact, the aforementioned aliasing components are readily visible in this expression. Consequently, (3.15) transforms into

$$\hat{c}(\varphi) = \sum_{k=0}^{N-1} \hat{C}[k] \mathrm{e}^{\jmath k\varphi} \tag{3.19}$$

which is, in short, a continuous approximate of the true SRF. For real-valued functions (as is the SRF), some constraints apply to the Fourier coefficients, such as

$$\begin{aligned} C[k] &= C^*[-k] \qquad \text{resp.} \\ \hat{C}[k] &= \hat{C}^*[N-k] \end{aligned} \tag{3.20}$$

which is a fundamental outcome of Fourier analysis. The asterisk sign (*) denotes complex conjugation. Basically, these relations can be obtained via forcing the sums in (3.15) and (3.19) to be real-valued. In turn, these equations can be rewritten as

$$c(\varphi) = A_0 + \sum_{k=1}^{\infty} A_k \cos(k\varphi + \varphi_k) \qquad \text{resp.} \tag{3.21}$$

$$\hat{c}(\varphi) = \hat{A}_0 + \sum_{k=1}^{\lfloor N/2 \rfloor} \hat{A}_k \cos(k\varphi + \hat{\varphi}_k) \qquad \forall\, c(\cdot), \hat{c}(\cdot) \in \mathbb{R} \tag{3.22}$$

where the complex-valued coefficients $C[k], \hat{C}[k]$ were split into *real*-valued amplitudes $A_k, \hat{A}_k$ and phases $\varphi_k, \hat{\varphi}_k$ according to

$$A_k = \quad 2|C[k]| , \qquad \varphi_k = \arg(C[k]) \qquad \text{resp.} \tag{3.23}$$

$$\hat{A}_k = \begin{cases} |\hat{C}[k]| & k = N/2 \\ 2|\hat{C}[k]| & \text{else} \end{cases} , \qquad \hat{\varphi}_k = \arg(\hat{C}[k]) . \tag{3.24}$$

Special treatment for even N is needed as only one appearance of $\hat{C}[N/2]$ appears in the sum of (3.19).

3.2.3 Distance Estimation

In the context of ToF, it is of interest to determine the argmax of the SRF $c(\varphi)$ (3.7). In this subsection, some approaches that appear suitable for the given estimation problem will be derived.

Analytic Approach

The most common approach of finding the minimum or maximum of a function which is given analytically, is to set the derivative zero. In the given context, (3.21) results in

$$\frac{\mathrm{d}c(\varphi)}{\mathrm{d}\varphi} = \sum_{k=1}^{\infty} A_k(-\sin(k\varphi + \varphi_k))k \quad \Rightarrow \quad \sum_{k=1}^{\infty} kA_k \sin(k\varphi + \varphi_k) \stackrel{!}{=} 0 \tag{3.25}$$

for which in general no closed-form solution can be obtained. As well, it should be emphasized that there exist two zero-crossings that are generally *not* spaced by π. Therefore, if an algorithm is employed for finding the zero-crossing that refers to the argmax, (3.21) should be used in advance to obtain suitable initial values.

Approximative Solutions

In practice, the problem of computing the infinite sum in (3.25) naturally vanishes as only $\lfloor N/2 \rfloor$ elements are available from N-point sampling. Moreover, as a compromise between complexity and accuracy, only $M \leq \lfloor N/2 \rfloor$ elements of the sum may be evaluated. This will be referred to as *model order* M in this Thesis and denoted by

$$\hat{c}_M(\varphi) = \hat{A}_0 + \sum_{k=1}^{M} \hat{A}_k \cos(k\varphi + \hat{\varphi}_k)\,. \tag{3.26}$$

However, one remaining problem is that the zero-crossing still needs to be computed iteratively, which is not a viable procedure for real-time systems. Fortunately, some simplifications can be made that lead to closed-form solutions. Starting with the assumption that the fundamental wave is dominant over the others (i.e., $A_1 \gg A_2, A_3, ...$), the argument of the cosine at fundamental frequency in (3.26) will be very small for a maximized SRF, i.e. $\varphi + \hat{\varphi}_1 \approx 0$. In other words, the offset between the phase of the fundamental wave and true argmax of the SRF will be small. Thus, considering models of order M, the following substitution in the derivative can be made:

$$\sum_{k=1}^{M} k\hat{A}_k \sin(k\varphi + \hat{\varphi}_k) = \sum_{k=1}^{M} k\hat{A}_k \sin(k\vartheta + \Delta\hat{\varphi}_k) = 0 \tag{3.27}$$

where ϑ and $\Delta\hat{\varphi}_k$ are defined as follows:

$$\begin{aligned} \vartheta &= \varphi + \hat{\varphi}_1 \approx 0 \\ \vartheta - \hat{\varphi}_1 &= \varphi \\ k\vartheta - k\hat{\varphi}_1 &= k\varphi \\ k\vartheta + \underbrace{\hat{\varphi}_k - k\hat{\varphi}_1}_{\Delta\hat{\varphi}_k} &= k\varphi + \hat{\varphi}_k\,. \end{aligned} \tag{3.28}$$

The right part of Eq. (3.28) equals the argument of the k-th sinusoid in Eq. (3.27). With the knowledge that ϑ will be very small, the goal is now to rewrite this equation in order to obtain trigonometric functions that take only ϑ as an argument. In turn, they can be linearized without

much loss of accuracy. Thus, basic trigonometric calculus yields

$$\sum_{k=1}^{M} k\hat{A}_k \left(\sin(k\vartheta)\cos(\Delta\hat{\varphi}_k) + \cos(k\vartheta)\sin(\Delta\hat{\varphi}_k)\right) = 0\,. \tag{3.29}$$

Now, the first items of the trigonometric series expansion should be good approximates, i.e.,

$$\begin{aligned} &\sin(k\vartheta) \approx k\vartheta \\ &\cos(k\vartheta) \approx 1 - \tfrac{1}{2}k^2\vartheta^2 \end{aligned} \tag{3.30}$$

where for now, the quadratic term of the cosine-approximate is maintained. Substituting (3.30) in (3.29) leads to

$$\begin{aligned} &\sum_{k=1}^{M} k\hat{A}_k \left(k\vartheta\cos(\Delta\hat{\varphi}_k) + (1 - \tfrac{1}{2}k^2\vartheta^2)\sin(\Delta\hat{\varphi}_k)\right) = 0 \\ &\sum_{k=1}^{M} \left(k^2\hat{A}_k\vartheta\cos(\Delta\hat{\varphi}_k) + k\hat{A}_k(1 - \tfrac{1}{2}k^2\vartheta^2)\sin(\Delta\hat{\varphi}_k)\right) = 0 \\ &\sum_{k=1}^{M} \left(k^2\hat{A}_k\vartheta\cos(\Delta\hat{\varphi}_k) + k\hat{A}_k\sin(\Delta\hat{\varphi}_k) - \tfrac{1}{2}k^3\hat{A}_k\vartheta^2\sin(\Delta\hat{\varphi}_k)\right) = 0 \\ &\vartheta\sum_{k=1}^{M} k^2\hat{A}_k\cos(\Delta\hat{\varphi}_k) + \sum_{k=1}^{M} k\hat{A}_k\sin(\Delta\hat{\varphi}_k) - \frac{1}{2}\vartheta^2\sum_{k=1}^{M} k^3\hat{A}_k\sin(\Delta\hat{\varphi}_k) = 0 \\ &\vartheta^2 - 2\frac{\sum_{k=1}^{M} k^2\hat{A}_k\cos(\Delta\hat{\varphi}_k)}{\sum_{k=1}^{M} k^3\hat{A}_k\sin(\Delta\hat{\varphi}_k)}\vartheta - 2\frac{\sum_{k=1}^{M} k\hat{A}_k\sin(\Delta\hat{\varphi}_k)}{\sum_{k=1}^{M} k^3\hat{A}_k\sin(\Delta\hat{\varphi}_k)} = 0 \end{aligned} \tag{3.31}$$

or

$$\vartheta^2 + \vartheta p + q = 0 \tag{3.32}$$

with

$$p = -2\frac{\sum_{k=1}^{M} k^2\hat{A}_k\cos(\Delta\hat{\varphi}_k)}{\sum_{k=1}^{M} k^3\hat{A}_k\sin(\Delta\hat{\varphi}_k)}, \qquad q = -2\frac{\sum_{k=1}^{M} k\hat{A}_k\sin(\Delta\hat{\varphi}_k)}{\sum_{k=1}^{M} k^3\hat{A}_k\sin(\Delta\hat{\varphi}_k)} \tag{3.33}$$

which allows analytic computation of the zero-crossing. Admittedly, the overall complexity still remains high, depending on the model order M. As well, the derived coefficients involve complicated arithmetic that seems rather challenging to be computed in real-time environments. Last but not least, solving a quadratic equation analytically is not a tractable approach either and will yield two solutions in general. For these reasons further simplifications appear recommended, which are derived below. Basically, two computing models are obtained, referred to as *linear* and *quadratic* model.

Linear Model. For very small values of ϑ, the additional simplification

$$\cos(k\vartheta) \approx 1 \tag{3.34}$$

is possible (cf. (3.30)). It can be shown that this eliminates the quadratic term in (3.32) and delivers the following, simplified estimation rule:

$$\begin{aligned}\hat{\vartheta}_{\text{lin}} &= -\frac{q}{p} \\ &= -\frac{\sum_{k=1}^{M} k\hat{A}_k \sin(\Delta\hat{\varphi}_k)}{\sum_{k=1}^{M} k^2\hat{A}_k \cos(\Delta\hat{\varphi}_k)} \,.\end{aligned} \tag{3.35}$$

Quadratic Model. Maintaining the quadratic term in (3.32), an estimator of higher precision will logically be obtained. However, is still necessary to overcome the complexity of the analytic zero-crossing computation. Fortunately, with $\hat{\vartheta}_{\text{lin}} = -q/p$ (3.35) a good estimate already exists that lies in the vicinity of the true solution. Computing the first-order Taylor approximation (cf. (3.10)) at this point yields

$$\begin{aligned}f(\vartheta) &= \vartheta^2 + \vartheta p + q \\ &\approx f(\hat{\vartheta}_{\text{lin}}) + \frac{\mathrm{d}f(\vartheta)}{\mathrm{d}\vartheta}\Big|_{\vartheta=\hat{\vartheta}_{\text{lin}}} (\vartheta - \hat{\vartheta}_{\text{lin}}) \\ &= \frac{p^2 - 2q}{p}\vartheta + \frac{qp^2 - q^2}{p^2}\end{aligned} \tag{3.36}$$

where the zero-crossing of this straight line should result in yet a better estimator:

$$\begin{aligned}\hat{\vartheta}_{\text{quad}} &= -\frac{q}{p}\frac{p^2 - q}{p^2 - 2q} \\ &= \hat{\vartheta}_{\text{lin}}\frac{p^2 - q}{p^2 - 2q} \\ &= \hat{\vartheta}_{\text{lin}}\left(1 + \frac{q}{p^2 - 2q}\right) \,.\end{aligned} \tag{3.37}$$

Putting it Together. Given the estimate $\hat{\vartheta}$, obtained either by the linear or quadratic model, the respective phase that approximates the argmax-estimation computes to (cf. (3.28))

$$\hat{\varphi}_d = \hat{\vartheta} + \hat{\varphi}_1 \,. \tag{3.38}$$

State-Of-The-Art Distance Estimation

Conventional ToF imagers dramatically simplify the discussed model in order to maintain real-time capabilities and advantages in terms of hardware design [3]:

- The SRF is sampled 4-fold, i.e., $N = 4$.
- The used model order for argmax-estimation is $M = 1$.

The impliciations within the current framework are then (3.9) (3.26)

$$\varphi_n = \frac{2\pi}{4}n = \left\{0,\ \frac{\pi}{2},\ \pi,\ \frac{3\pi}{2}\right\} \text{ for } n = \{0, \dots, 3\} \tag{3.39}$$

$$\hat{c}_1(\varphi) = \hat{A}_0 + \hat{A}_1 \cos(\varphi + \hat{\varphi}_1) \,. \tag{3.40}$$

The respective estimation algorithm (referred to as *4-phase algorithm* from now on) is simply the phase of the fundamental wave

$$\begin{aligned}
\hat{\varphi}_d &= -\operatorname{argmax}\{\hat{c}_1(\varphi)\} && \text{cf. (3.7)} \\
&= -(-\hat{\varphi}_1) && \text{cf. (3.25)} \\
&= \arg(\hat{C}[1]) && \text{cf. (3.24)} \\
&= \arg\left(\tfrac{1}{4}(c[0] - \jmath c[1] - c[2] + \jmath c[3])\right) && \text{cf. (3.16)} \\
&= \operatorname{atan2}(c[3] - c[1], c[0] - c[2]) && (3.41)
\end{aligned}$$

with $\operatorname{atan2}(\operatorname{Im}(z), \operatorname{Re}(z))$ realizing the 4-quadrant arc tangent, i.e., computing the phase of a complex number z.

If desired, the amplitude $\hat{A}_1$ can be computed as

$$\begin{aligned}
\hat{A}_1 &= 2|\hat{C}[1]| && \text{cf. (3.24)} \\
&= 2|\tfrac{1}{4}(c[0] - \jmath c[1] - c[2] + \jmath c[3])| && \text{cf. (3.16)} \\
&= \frac{1}{2}\sqrt{(c[0] - c[2])^2 + (c[3] - c[1])^2}\,. && (3.42)
\end{aligned}$$

With the given simplifications discussed above, several drawbacks of this distance estimation method can be expected. First of all, 4-fold sampling will cause severe aliasing artifacts, whose impact on depth estimation is known as Wiggling Error. In terms of sampling theory, regarding the first frequency bin $\hat{C}[1]$ these artifacts amount to (3.18)

$$\begin{aligned}
\hat{C}[1] &= C[1] + \sum_{m=1}^{\infty} (C[4m+1] + C^*[4m-1]) \\
&= C[1] + C^*[3] + C[5] + C^*[7] + C[9] + \cdots && (3.43)
\end{aligned}$$

which states that aliasing of every odd harmonic occurs on the first frequency bin. Furthermore, disregarding aliasing effects, the question remains whether the phase of the fundamental wave sufficiently well accords to the true maximum of the SRF. As already discussed, the fundamental wave is expected to be dominant, so it might not be a bad approximation at all. However, literature study revealed that no comprehensive analysis of the SRF has been undertaken so far.

3.3 Selected Topics

In the practical part of this Thesis several discussions will require some theoretic prerequisites, which are provided in this section. Related to Section 5.2.2, an analytic derivation how the Wiggling Error can be simulated is given in Section 3.3.1. Furthermore, since multipath reflections will be subject to detailed analyses in the context of the amplitude-related error (cf. Section 5.3.2), a theoretic framework is outlined in Section 3.3.2.

3.3.1 Simulating the Wiggling Error

The question whether the shape (and thus the spectral composition) of the SRF changes with the target distance or not will be discussed in Section 5.2. However, assuming spectral consistency over distance for now, any circular depth error caused by violations of the sampling theorem can be reproduced by actually *shifting* one single measured SRF, either on a sample-basis or analytically. In other words, given a captured SRF $c[n]$ by N samples,

$$c[((n-m))_N] = c(\varphi_n - \varphi_m) \qquad \text{with} \qquad \varphi_m = \frac{2\pi}{N} m, \quad m \in \mathbb{Z} \tag{3.44}$$

represents a *virtual* relative target shift of φ_m with respect to some original distance the unshifted SRF refers to. The notation $((\cdot))_N$ reflects the *modulo N* operator. These virtual shifts are of course discrete, depending on the number of samples. Analytically, any *continuous* shift φ_m of the SRF can be expressed as (cf. (3.15))

$$\begin{aligned} c(\varphi - \varphi_m) &= \sum_{k=-\infty}^{\infty} C[k] \mathrm{e}^{\jmath k(\varphi - \varphi_m)} \\ &= \sum_{k=-\infty}^{\infty} C_{\varphi_m}[k] \mathrm{e}^{\jmath k \varphi} \qquad \text{with} \qquad C_{\varphi_m}[k] = C[k] \mathrm{e}^{-\jmath k \varphi_m} \,. \end{aligned} \tag{3.45}$$

Equidistant sampling by any number N that produces aliasing artifacts yields (3.18)

$$\hat{C}_{\varphi_m}[k] = C_{\varphi_m}[k] + \sum_{m=1}^{\infty} \left(C_{\varphi_m}[k + mN] + C_{\varphi_m}[k - mN] \right) \tag{3.46}$$

where it becomes clear that the error of the k^{th} harmonic caused by insufficient sampling, $\hat{C}_{\varphi_m}[k] - C_{\varphi_m}[k]$, is a *function* of φ_m. The actual phase error of the k^{th} harmonic can then be expressed as

$$\begin{aligned} &e_{\varphi_m}[k] = \\ &\quad = \arg(\hat{C}_{\varphi_m}[k]) - \arg(C_{\varphi_m}[k]) \\ &\quad = \arg\left(C_{\varphi_m}[k] + \sum_{m=1}^{\infty} \left(C_{\varphi_m}[k + mN] + C_{\varphi_m}[k - mN] \right) \right) - \arg\left(C_{\varphi_m}[k] \right) \end{aligned}$$

$$
\begin{aligned}
&= \arg\left(C[k]\mathrm{e}^{-\jmath k\varphi_m} + \sum_{m=1}^{\infty}(C[k+mN]\mathrm{e}^{-\jmath(k+mN)\varphi_m} + C[k-mN]\mathrm{e}^{-\jmath(k-mN)\varphi_m})\right) \\
&\qquad - \arg\left(C[k]\mathrm{e}^{-\jmath k\varphi_m}\right) \\
&= \arg\left(\left(C[k] + \sum_{m=1}^{\infty}(C[k+mN]\mathrm{e}^{-\jmath mN\varphi_m} + C[k-mN]\mathrm{e}^{+\jmath mN\varphi_m})\right)\mathrm{e}^{-\jmath k\varphi_m}\right) \\
&\qquad - \arg\left(C[k]\mathrm{e}^{-\jmath k\varphi_m}\right) \\
&= \arg\left(C[k] + \sum_{m=1}^{\infty}(C[k+mN]\mathrm{e}^{-\jmath mN\varphi_m} + C[k-mN]\mathrm{e}^{+\jmath mN\varphi_m})\right) - k\varphi_m \\
&\qquad - (\arg(C[k]) - k\varphi_m) \\
&= \arg\left(C[k] + \sum_{m=1}^{\infty}(C[k+mN]\mathrm{e}^{-\jmath mN\varphi_m} + C[k-mN]\mathrm{e}^{+\jmath mN\varphi_m})\right) - \arg(C[k])\,. \qquad (3.47)
\end{aligned}
$$

The analytic expression of the Wiggling Error is then (3.47) evaluated at $k = 1$, i.e.,

$$
e_{\mathrm{W}}(f_{\mathrm{mod}}, d) = e_{\varphi_m}[1] \qquad \text{with} \qquad \varphi_m = \frac{4\pi f_{\mathrm{mod}}}{\mathrm{c}} d\,. \qquad (3.48)
$$

Consequently, if the true Fourier coefficients can be determined for some initial distance d_0, the Wiggling error of the SRF sampled at *any* rate can be determined analytically according to (3.48) resp. (3.47). In turn, these formulas could be used as a starting point for algorithmic Wiggling Error removal, but it must be taken into account that (3.48) takes the *true* distance as an argument, which is not available in practice.

3.3.2 The Multipath-Reflection Scenario

As a basis for understanding how multipath reflections affect phase information carried by the SRF, the signal path discussed in Section 3.1 should be revised in the context of a nonideal optical path (3.12), which results in the following formula for the SRF (3.5):

$$
c(\varphi) = \sum_i g(d_i) \int_0^{\tau} r\left(t - t_{d_i} - \tfrac{\varphi}{2\pi f_{\mathrm{mod}}}\right) r(t)\mathrm{d}t\,. \qquad (3.49)
$$

Similar to the representation of the SRF via its Fourier series, $r(t)$ can be written as

$$
r(t) = \sum_{k=-\infty}^{\infty} R[k]\mathrm{e}^{\jmath 2\pi k f_{\mathrm{mod}} t} \qquad (3.50)
$$

with the Fourier coefficients $R[k]$. Inserting this definition into (3.49) yields

$$
c(\varphi) = \sum_i g(d_i) \int_0^{\tau} \sum_{k=-\infty}^{\infty} R[k]\mathrm{e}^{\jmath 2\pi k f_{\mathrm{mod}}\left(t - t_{d_i} - \frac{\varphi}{2\pi f_{\mathrm{mod}}}\right)} \sum_{k'=-\infty}^{\infty} R[k']\mathrm{e}^{\jmath 2\pi k' f_{\mathrm{mod}} t}\mathrm{d}t \qquad (3.51)
$$

which can be rewritten as

$$c(\varphi) = \sum_i g(d_i) \sum_{k=-\infty}^{\infty} \sum_{k'=-\infty}^{\infty} \int_0^{\tau} R[k]R[k']\mathrm{e}^{\jmath 2\pi(k+k')f_{\mathrm{mod}}t}\mathrm{d}t\,\mathrm{e}^{-\jmath 2\pi k f_{\mathrm{mod}}t_{d_i}}\mathrm{e}^{-\jmath k\varphi}\,. \tag{3.52}$$

It is evident that for $k + k' \neq 0$, the integral is zero as integration is performed over an integer multiple of periods. Consequently, setting $k' = -k$ eliminates the sum over k' and since the integrand reduces to 1, the integral can be substituted with τ, thus

$$c(\varphi) = \tau \sum_i g(d_i) \sum_{k=-\infty}^{\infty} R[k]R[-k]\mathrm{e}^{-\jmath 2\pi k f_{\mathrm{mod}}t_{d_i}}\mathrm{e}^{-\jmath k\varphi}\,. \tag{3.53}$$

Finally, changing k to $-k$ and recapitulating that $r(t)$ is real (thus $R[k] = R^*[-k]$ (3.20)) results in

$$\begin{aligned} c(\varphi) &= \sum_{k=-\infty}^{\infty} \tau \sum_i g(d_i)|R[k]|^2\mathrm{e}^{\jmath 2\pi k f_{\mathrm{mod}}t_{d_i}}\mathrm{e}^{\jmath k\varphi} \\ &= \sum_{k=-\infty}^{\infty} C[k]\mathrm{e}^{\jmath k\varphi} \end{aligned} \tag{3.54}$$

where

$$C[k] = \tau|R[k]|^2 \sum_i g(d_i)\mathrm{e}^{\jmath 2\pi k f_{\mathrm{mod}}t_{d_i}} \tag{3.55}$$

refer to the Fourier coefficients of $c(\varphi)$. It is now evident that multipath reflections transparently affect the spectral components of the SRF – which is due to the assumed linearity of the system no surprise. In particular, (3.55) concisely states that the k^{th} harmonic can be seen as a weighted superposition of complex exponentials (phasors) according to the multipath components. Of most interest is of course still the fundamental wave, thus for better readability and intuition, for the fundamental wave a normalized phasor may be defined as

$$z = \frac{C[1]}{\tau|R[1]|^2} = \sum_i z_i \qquad \text{with} \qquad z_i = g(d_i)\mathrm{e}^{\jmath 2\pi f_{\mathrm{mod}}t_{d_i}} \tag{3.56}$$

which should be rewritten as

$$z = z_s + \sum_r z_r \qquad \text{with} \qquad z_s = g(d)\mathrm{e}^{\jmath 2\pi f_{\mathrm{mod}}t_d} \quad \text{and} \quad z_r = g(d_r)\mathrm{e}^{\jmath 2\pi f_{\mathrm{mod}}t_{d_r}} \tag{3.57}$$

for better distinction between signal phasor z_s and multipath-reflection phasors z_r. Evidently, the argument of the signal phasor carries the true distance information, i.e., $\arg(z_s) = \varphi_d$.

It should be remarked that the current derivation of $C[k]$ is fully consistent with the rest of this Thesis. For example, in the absence of multipath reflections, $z = z_s = g(d)\mathrm{e}^{\jmath 2\pi f_{\mathrm{mod}}t_d}$ with a phase of the fundamental wave equivalent to $2\pi f_{\mathrm{mod}}t_d$. Thus, argmax-estimation of the first-order model (cf. (3.40)) will yield $\hat{\varphi}_d = -\operatorname{argmax}\{\hat{c}_1(\varphi)\} = 2\pi f_{\mathrm{mod}}t_d$ which corresponds to (2.2).

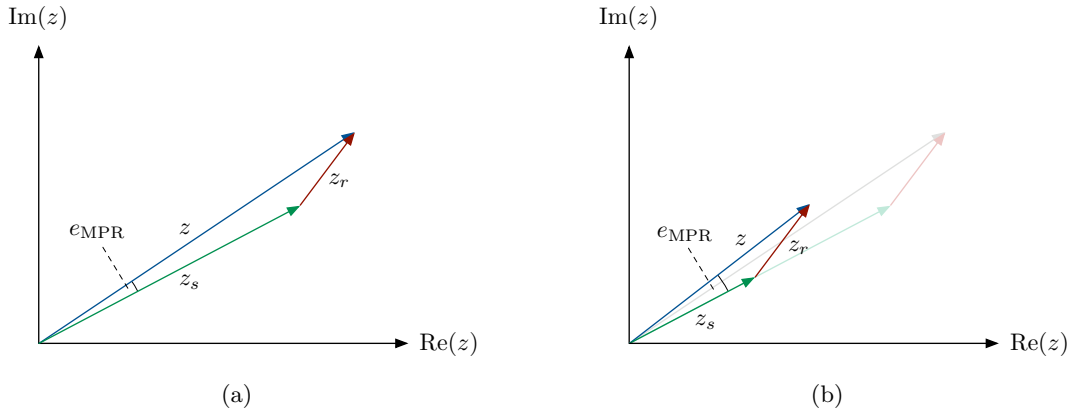


Figure 3.2: Simple example that illustrates the influence of multipath reflections on distance estimation. Denoting the green phasor z_s as the signal phasor that carries the actual distance information and the red phasor z_r as some disturbing phasor, the resulting phase error e_{MPR} evidently increases from (a) to (b) as the signal amplitude decreases.

In Fig. 3.2, a basic illustration is given as an impression how reflective components of the scene influence depth sensing. In particular, it is shown that lower amplitudes of the signal phasor lead to higher deviations. This deviation can be quantified as

$$\begin{aligned} e_{\mathrm{MPR}}(z, d, f_{\mathrm{mod}}) &= \arg(z) - \arg(z_s) \\ &= \hat{\varphi}_d - 4\pi f_{\mathrm{mod}} d/\mathrm{c}\,. \end{aligned} \tag{3.58}$$

— CHAPTER FOUR —

Experiments

This chapter introduces to the practical part of this Thesis. Section 4.1 provides a general motivation that should comprise enough impulses to justify the aimed investigations. Secondly, Section 4.2 is dedicated to the actual experimental setup, whereas in Section 4.3 the parameter space that characterizes each of the measurements is elaborated and discussed in detail. Finally, as a link to Chapter 5, several experiments to be conducted are motivated in Section 4.4.

4.1 Motivation

The very last sentences of Section 3.2 already motivate the practical part of this Thesis. Analyses of the state-of-the-art distance estimation algorithm, which appears to be rather approximate than exact, gives immediate rise to the following three questions:

- In terms of systematic depth deviations, which drawbacks stem from the current distance estimation approach?
- How does the SRF *really* look like, and – more interestingly – under the variation of several parameters, how does it behave?
- Given knowledge about the SRF and a thorough mathematical framework, how can distance estimation possibly be improved?

Unsurprisingly, it was found that direct measurement of the SRF (i.e., capturing it at arbitrary phase shifts) would provide deepest insights into the system, not at least for the following reasons:

- The SRF is of central importance in the distance estimation chain and should per se contain all information that is needed for consistent depth sensing (apart from metadata, e.g., from temperature sensors). Moreover, as no profound empirical analysis has been published yet it is still one of the biggest remaining unknowns in ToF imaging.
- The SRF is completely transparent in the mathematical framework derived in Chapter 3, so any empirical information can be directly applied to the signal model.
- The SRF can be measured without requiring complicated setups that potentially alter the behavior of the system.

4.2 Setup

4.2.1 Overview and Modes of Operation

A very basic sketch of the measurement setup is depicted in Fig. 4.1. Generally, it should be noted that the provided descriptions are kept as abstract as possible but as detailed as necessary to understand the principles. Main components are

- A state-of-the-art ToF imager with specifications given in Tab. 2.1 on page 4. This imager is capable of delivering the signal for the illumination unit, i.e., at reference phase shifts $\varphi_i = \{0, \pi/2, \pi, 3\pi/2\}$ (cf. (3.39)).
- An illumination unit consisting of an 8×8 Osram SFH4259 LED-block, indicated by the diode-symbol.
- The signal generator Agilent Technologies 81150a.
- A PC connected to the imager with MathWorks® Matlab® for data processing.

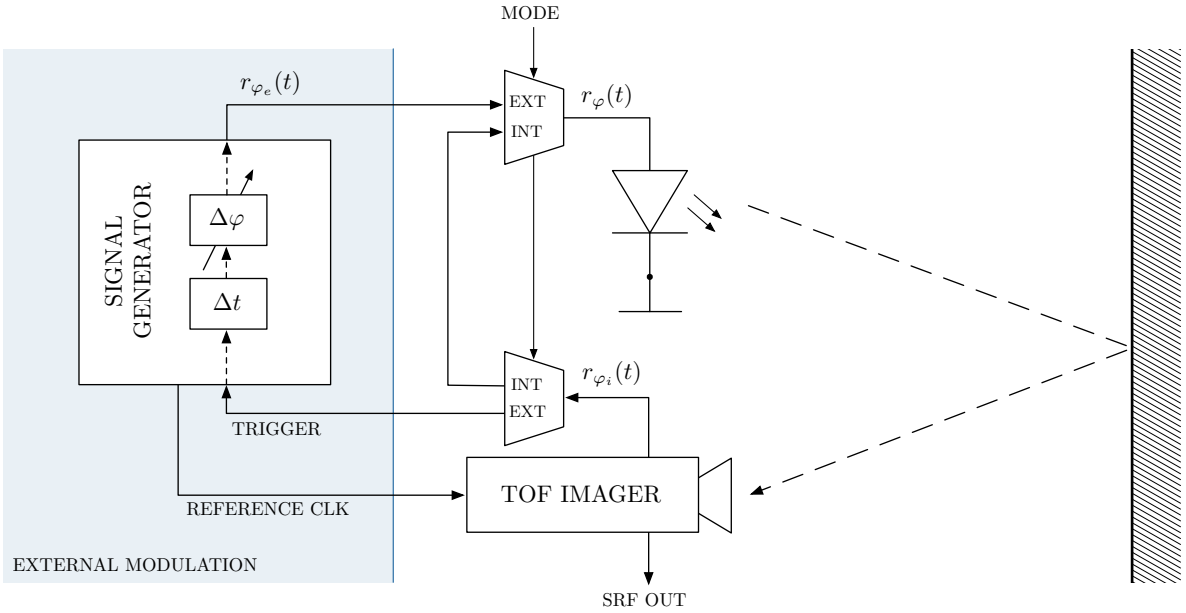


Figure 4.1: Schematic diagram of the measurement setup.

As can be seen in Fig. 4.1 and was mentioned above, the given setup features two modes of operation: *Internal* and *external* modulation, which are selected by the mode-switch.

Internal Measurement/Mode/Modulation

In this mode, the original setup is left as is, directly driving the illumination unit with a modulation signal of four different reference phases at $0, \pi/2, \pi$ and $3\pi/2$. This 4-fold sampled SRF can then directly be read out. At this stage, the term *frame* is defined, which refers to one quadruple of SRF values.

External Measurement/Mode/Modulation

This mode on the other hand aims at custom sampling of the SRF by applying arbitrary phase shifts to the modulation signal as defined by the signal generator. As turned out, the robustest setting was achieved by triggering the signal generator with the modulation signal from chip (which should actually drive the illumination unit). This obviously requires $\varphi_i = 0$ to be set in order to provide a consistent zero-phase trigger signal. As well, it should be mentioned that not only the additional cable lengths introduced further lags to the signal, but also the signal generator internally exhibits a certain trigger delay (denoted by Δt). These effects need to be known in order to set $\Delta\varphi$ accordingly to maintain consistency of $r_{\varphi_e}(t)$ with the rest of the setup. Finally, with the signal generator being directly controlled by the PC, a completely autonomous measurement setup was established.

For completeness, it should be noted that regardless internal or external mode, the ToF imager always delivers frame-based data, i.e., in atomic packets of four SRF values. Logically, in contrast to internal mode, these values refer to the *same* reference phase shift.

4.2.2 Setup Validation

In to order validate the proposed setup and to make the results of internal and external modulation comparable, it is required to align both measurement modes. Basically, three parameters were assumed to model the mismatch between the two modes,

$$\begin{aligned} d &\ldots \text{SRF amplitude offset} && \stackrel{!}{=} 0 \\ k &\ldots \text{SRF amplitude scale} && \stackrel{!}{=} 1 \\ \varphi_s &\ldots \text{SRF phase offset} && \stackrel{!}{=} \text{const}\,, \end{aligned} \tag{4.1}$$

so the following cost function

$$J(d, k, \varphi_s) = \sum_{i,j,n} |k c_{\mathrm{E}}[i, j, \varphi_n - \varphi_s) + d - c_{\mathrm{I}}[i, j, n]|^2 \tag{4.2}$$

was established and minimized. Note that i, j refer to the pixel location as the cost function is globally defined over the whole pixel array. As well, internal and external mode need to be distinguished in this expression, so the respective subscripts I and E were introduced. Maybe a bit misleading is the phase argument $\varphi_n - \varphi_s$ of the external SRF, which accounts for the continuous nature of φ_s. Actually φ_n and n (both discrete) refer to exactly the same sampling points.

In practice, optimization with respect to this continuous parameter φ_s can be obtained by employing the model discussed in the context of the Wiggling-Error simulation (3.45). It should be noted that internal and external modulation were almost perfectly well aligned, as discussed in Section 5.1.1.

4.3 Parameter Space

In order to answer questions regarding the behavior of the SRF, it must be acquired under varying parameters. Therefore, this section is dedicated to a discussion of potentially relevant parameters.

4.3.1 Measurement Parameters

Besides others, the measurement parameters of main interest are

f_{mod} ... Modulation frequency,

τ ... Integration time,

N_E ... Number of samples,

K ... Number of realizations.

The number of samples N_E is equivalent to the earlier introduced term N_E-point sampling. Consequently, the respective phase shifts accord to the definition of (3.9). Sometimes, if suitable in the respective context, this number will also be referred to as sampling *rate*. It should be remarked that for a given SRF captured by N_E samples, any equidistant subset of N samples can be extracted which then in turn refers to N-point sampling, i.e., $N_E/N \in \mathbb{N}$. However, this is only a practical detail that quickens the measurement process and not of theoretic importance.

In order to determine stochastic properties such as mean and variance, each measurement must be repeated a certain amount of times. The parameter K accounts for this number of repetitions, which can be interpreted as a number of *realizations* of an underlying random process.

More of analyzing purposes and therefore not explicitly mentioned is the parameter M, which was introduced as the model order M of the continuously approximated SRF (3.26).

4.3.2 Scene Parameters

Besides the system-dependent variables discussed above, a multitude of parameters is defined by the scene itself. In particular relevant are

d ... Target distance,

r ... Target reflectivity.

Of course, the environment is abstracted dramatically at this stage, and even more problematic, it is in general not possible to treat these parameters independently as they cause effects that are not orthogonal. For example, larger distances weaken the energy of the reflected signal, which is basically the same effect that is caused by darker surfaces. However, target distance and target reflectivity are considered as the main sources of scene-dependent systematic errors (cf. Wiggling Error and amplitude-related error, respectively), so it is indispensable to investigate them and elaborate their influences.

4.3.3 Summary

The parameters discussed in this section can be clustered into those which require distinct measurements of the SRF (such as $f_{\mathrm{mod}}, \tau, N_E, K, d$) and those which are contained implicitly within one measurement (N, M, r). With the relation $N_E/N \in \mathbb{N}$, N_E can be set constant to a value that allows the requested subsets N. As well, K can be set accordingly to a desired level of confidence for mean and variance to be computed. Consequently, this forms a three-dimensional parameter space in $(f_{\mathrm{mod}}, \tau, d)$. However, fully sampling this parameter cube is not tractable as the necessary amount of measurements would exceed any reasonable number – not to mention the huge amount of data to be structured and analyzed. As an additional difficulty, integration time and target distance are not independent since larger distances require longer integration times if a sufficiently high SNR should be maintained. Consequently, it was decided to capture a representative subset of the parameter cube, in particular by the following experiments:

1. *Moving-target experiment* (cf. Section 4.4.2). The parameter plane $(f_{\mathrm{mod}}, \tau_0, d)$ is evaluated, i.e., for varying modulation frequency and target distance. The integration time τ_0 is set constant for each measurement series.

2. *Steady-target experiment* (cf. Section 4.4.3). For constant target distance d_0, modulation frequency and integration time are varied. This refers to the parameter plane $(f_{\mathrm{mod}}, \tau, d_0)$.

4.4 Suggested Experiments

4.4.1 Preliminary Experiments

Before starting with the actual measurements, some preparatory steps are necessary. A short discussion of these steps is provided in this subsection.

Setup Validation

At the first stage, it is mandatory to align internal and external mode in order to validate the whole measurement setup. This task was discussed in Section 4.2.2.

Reproducibility

Secondly, it is a fundamental objective to determine the order of magnitude within which the experiments are reproducible. Several issues are expected to degrade obtained data, such as

- Fixed-Pattern Noise. These pixel-based deviations are regarded as consistent over variation of the discussed parameters. In turn, they may be calibrated out in advance if several pixels should be compared or smoothness is a demand. However, FPN is not problematic if single pixels or lateral means are the quantities of interest.

- Jitter. An additional source for jitter will be the signal generator, so it is important to investigate potential problems that come with it and to estimate the additional amount of averaging necessary in order to reduce these impacts.

- Component Warm-Up. Long-term drifts caused by component warm-up may cause severe deviations at startup, but will get small when some steady state is reached. If necessary, certain *premodulation sequences* prior to actual depth acquisition may be implemented.
- Environmental Influences. As the measurements are not carried out in a temperature-stabilized dark room, further problems may arise from the environment, such as room temperature, daylight, multipath reflections, etc. Besides darkening the room as much as possible, reflective objects that are potentially disturbing should be removed.

Fourier Analysis

Before applying reconstruction models of arbitrary order, it is an obligatory step to determine up to which order this reconstruction is reasonable. In other words: Based on the modulation frequency, spectral components that actually carry information should be identified.

4.4.2 Moving-Target Experiment

As already introduced in Section 4.3.3, the moving-target experiment realizes a cross-section of the discussed parameter cube, i.e., for varying target distance d and modulation frequency f_{mod}, but for constant integration time τ_0. In practical terms, this experiment mainly focuses on elaborating the Wiggling Error, thus investigating the influence on distance itself to distance estimation. The target is chosen to be white and planar. Besides reproducing the known issues, from the acquired data the following fundamental questions should be answered:

- Depending on the modulation frequency, which model order M is necessary to give accurate distance estimates?
- Given the optimal model order, which distance estimation concepts can be developed that are superior to the 4-phase algorithm?

4.4.3 Steady-Target Experiment

With the role of the distance parameter d on depth estimation hopefully understood in principle after the moving-target experiment it is a necessary simplification to remove this degree of freedom in a second step, for which reason this experiment is called *steady*-target experiment. Capturing now a target with different reflectivities under the variance of modulation frequency and integration time, answers to the following questions may be found:

- How do the findings from Section 4.4.2 – in particular regarding model orders – extend to the parameters target reflectivity r and integration time τ?
- Is it possible to identify the root causes of integration-time related and amplitude-related error?
- Which possible countermeasures against these errors can be developed?

— CHAPTER FIVE —

Results – Status Quo

With several experiments motivated in Section 4.4, this chapter aims at presenting and discussing the respective results with possible answers to the asked questions. Additionally, the findings of this chapter should provide a basis of improved distance estimation algorithms that are subject of Chapter 6. For best clarity, the structure of this chapter is kept analog to that of Section 4.4, with the preliminary experiments discussed in Section 5.1, the moving-target and steady-target experiment discussed in Section 5.2 and Section 5.3, respectively.

5.1 Preliminary Experiments

5.1.1 Setup Validation

With an exemplary plot of the SRF given in Fig. 5.1, it can be seen that internal and external mode match quite well. The dots refer to internal modulation at reference phases $n\pi/2, n = \{0, \ldots, 3\}$, whereas the lines refer to the quasi-continuous SRF captured by means of external modulation. The matching algorithm minimizing (4.2) almost exactly hit the desired values over all captured scenes, i.e.,

$$\begin{aligned} d &= 0 \ \pm \ 1\,\mathrm{LSB} \\ k &= 1 \ \pm \ 0.1\% \\ \varphi_s &= \mathrm{const}\,, \end{aligned} \tag{5.1}$$

where the given uncertainty values just refer to the order of magnitude of the fluctuations. The constant phase offset φ_s between the two modes of operation was compensated by an according delay to $\Delta\varphi$ (cf. Fig. 4.1) in order to finally establish a congruent setup.

5.1.2 Reproducibility

During the practical part of this Thesis, it turned out that actually making the measured data reproducible was far more difficult than initially expected. Related to the expected problems stated in Section 4.4, exactly these effects caused the setup to exhibit unpredictable behavior.

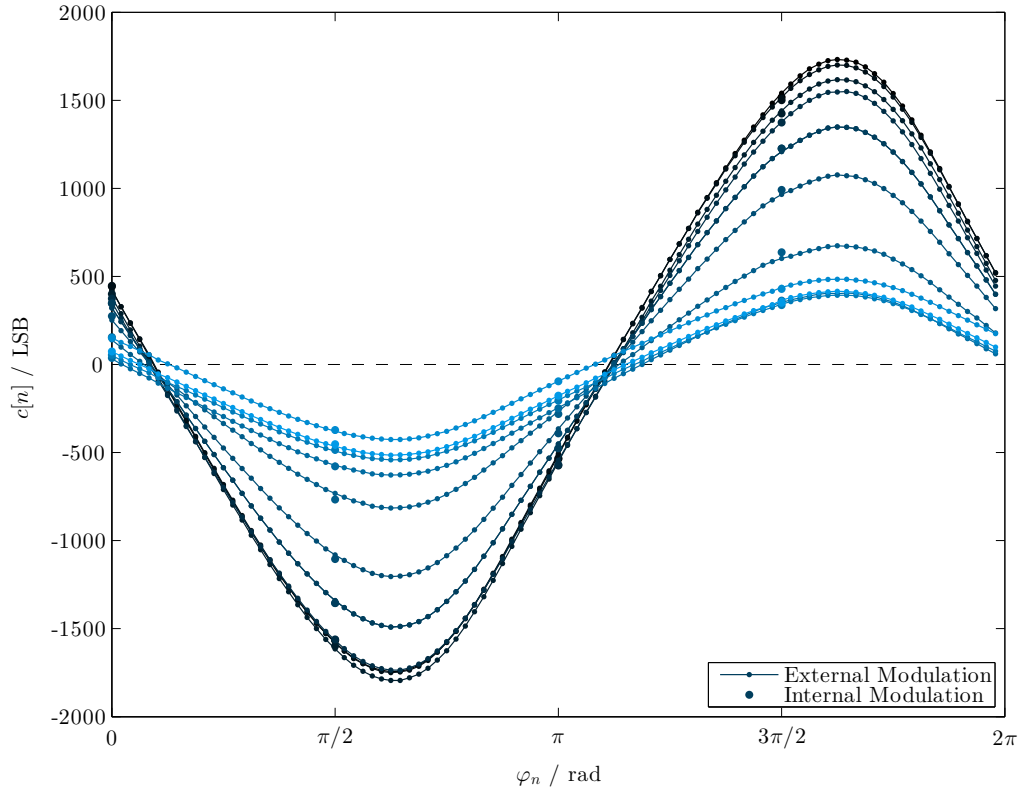


Figure 5.1: SRF plot of the alignment between internal and external modulation. An exemplary subset of 11 neighboring pixels that capture a target of varying reflectivity is shown. For this scene, the matching results over the whole array were $k = 1.0014$ and $d = -2.42$ LSB with a remaining root-mean-square error (RMSE) of 17.24 LSB. $[f_{\mathrm{mod}} = 20\ \mathrm{MHz}, \tau = 800\,\mu\mathrm{s}, N_E = 96, K = 32]$

Jitter

Although expected in general, the impact of additional jitter caused by the external modulation path was surprisingly high. Although on average no systematic deviations occurred (for which reason matching internal and external modulation was not problematic), the fact that jitter is a *timing* issue and therefore transparently influences the quantity to be measured severely deteriorated the reproducibility of obtained results. This can be seen in Fig. 5.2, where the estimation error over several succeeding frames is shown. The blue and green curves are the individual pixel estimates of external and internal modulation, respectively. The overall array mean is colored black. Internal and external measurement were performed under equal circumstances and depth values for both modes were computed according to the 4-phase algorithm.

Fig. 5.2(a) shows results of one of the first iterations of the measurement setup, where the signal generator was triggered by an internal reset signal. What this rather suboptimal setting impressively proves, is that jitter is a completely global effect on all pixels. This is not surprising, as jitter refers to temporal inconsistencies between modulation signal and reference signal, but not between reference signal-paths between the pixels. Directly using the internal modulation signal at reference phase zero as a trigger (cf. Fig. 4.1) increased the performance significantly. However, the external path through the signal generator still introduced additional jitter. In order to overcome this issues and to achieve a good compromise that approximates internal

modulation sufficiently well also in statistical terms, an implicit 4-fold averaging of the values provided by one frame was implemented. The result is depicted in Fig. 5.2(b).

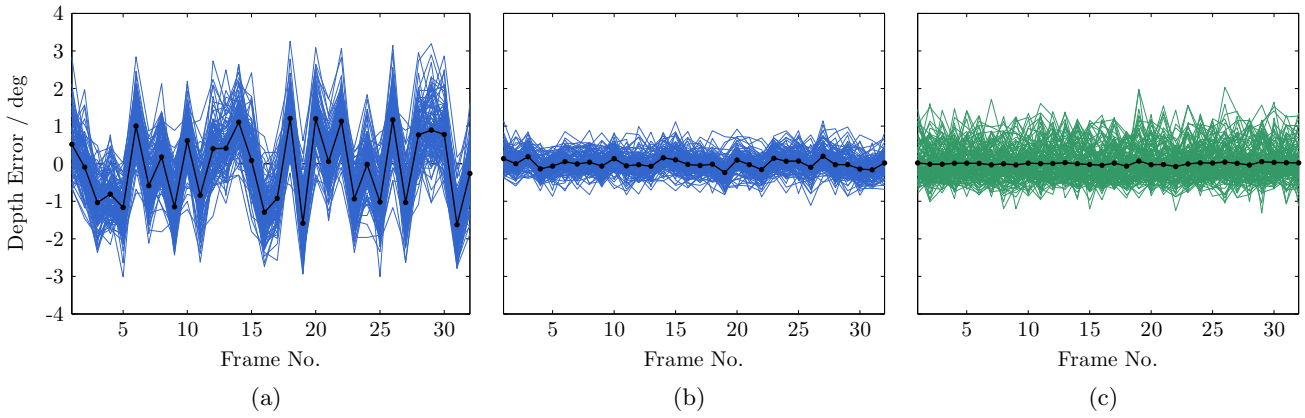


Figure 5.2: Comparison of jitter impact on different measurement setups: (a) external modulation with jittering trigger signal, (b) external modulation with solid trigger signal (current implementation), and (c) internal modulation. Individual pixel values are printed in color, lateral mean in black. [$f_{\mathrm{mod}} = 20$ MHz, $\tau = 800\,\mu$s, $N_E = 4$, $K = 1$]

Comparing Fig. 5.2(b) with Fig. 5.2(c), it can be seen that the pixel noise of external modulation is already smaller than that of internal modulation, whereas external mode still jitters more. For this concrete example, in external mode the RMSE over all frames amounts to 0.29° (individual pixels) and 0.11° (array mean). For internal modulation, the respective values are 0.46° (individual pixels) and 0.032° (array mean).

Component Warm-Up

In the given scenario, the illumination unit and the imaging chip are considerably influenced by self-induced temperature changes, especially at startup. Regarding the illumination unit, it was observed that higher temperatures cause the optical signal level to drop significantly (up to 12% within one minute of continuous drive). However, a far more important issue is that the switching performance of the circuit gets worse, appearing as a temperature-related drift in the measured depth (up to 0.5° at 20 MHz). As the acquisition of an SRF in external mode takes up to several minutes, these drifts result in inconsistent measurement data.

Towards avoiding this issue, a premodulation sequence (of equal temporal behavior) was inserted prior to each measurement in order to reach some steady state before starting the real acquisition. In this sequence, the overall mean depth of the given scene was evaluated and if convergence was detected, the actual measurement procedure was started.

In Fig. 5.3, a typical premodulation sequence is depicted. With the red curve being a 9-fold moving-average of the actual depth estimates, the convergence towards a certain depth value can clearly be seen. Several figures of merit (FoM) were employed to determine when the depth values have converged. It turned out that robustest performance was achieved by computing the variance of the latest 20 samples and setting the convergence threshold to -50 to -60 dB. The fact that the convergence procedure is not always of simple exponential nature but potentially a superposition of several underlying processes justifies a criterion based on variance. This

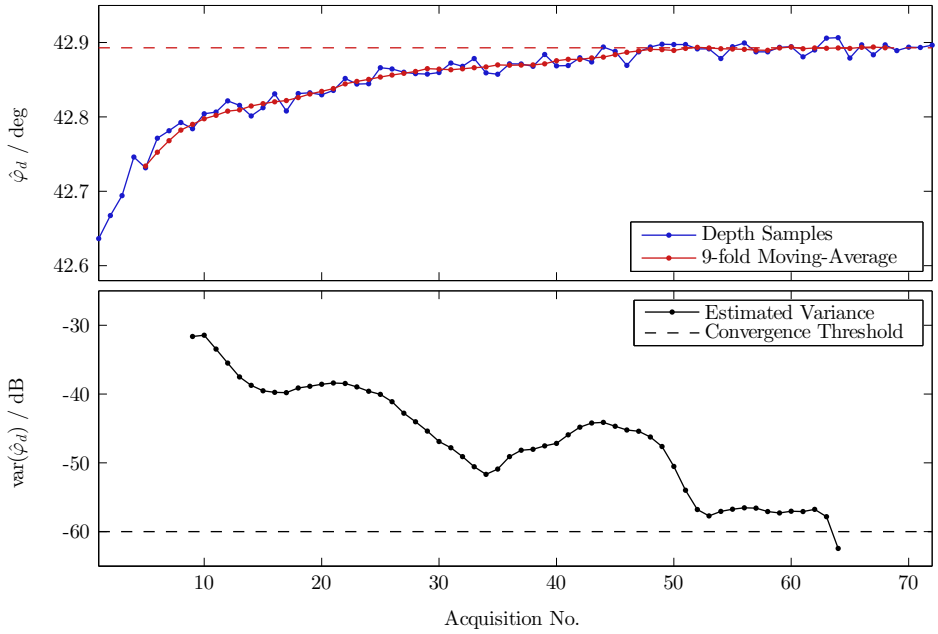


Figure 5.3: Typical premodulation sequence that shows depth convergence towards a certain value as the device warms up. [$f_{\mathrm{mod}} = 20$ MHz, $\tau = 800\,\mu$s, $N_E =$ n/a, $K = 32$]

basically assures that not only the mean needs to reach a steady value but also the fluctuations around this mean need to vanish. Despite this, a simple line-fitting approach or first-order exponential approximation that seemed reasonable in the first place do not account for this additional complexity.

Environmental Influences

The countermeasures against jitter and warm-up effects discussed above already produced reliable results in short-term scenarios. However, no consistency over hours and days was obtained, as

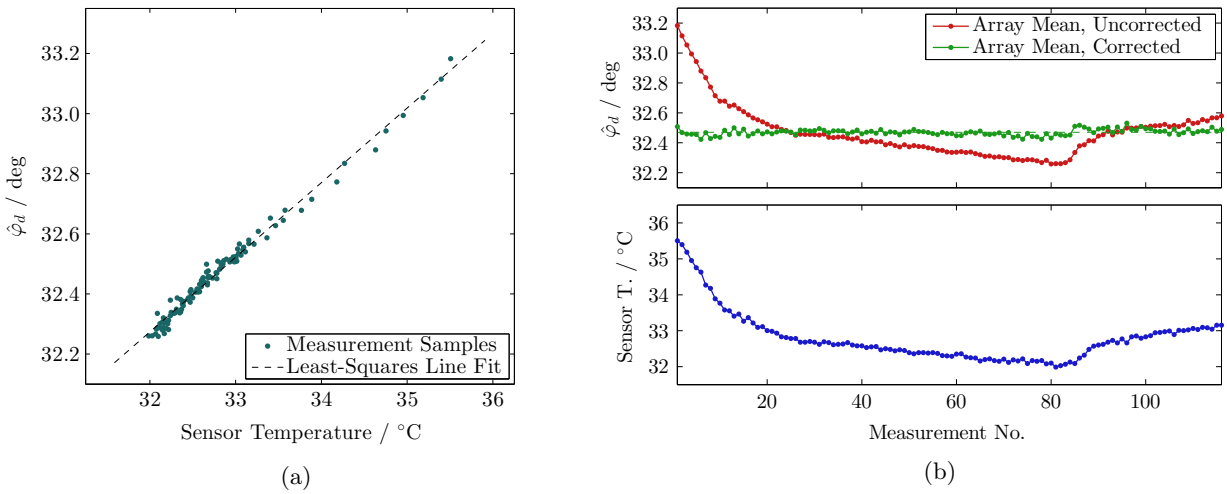


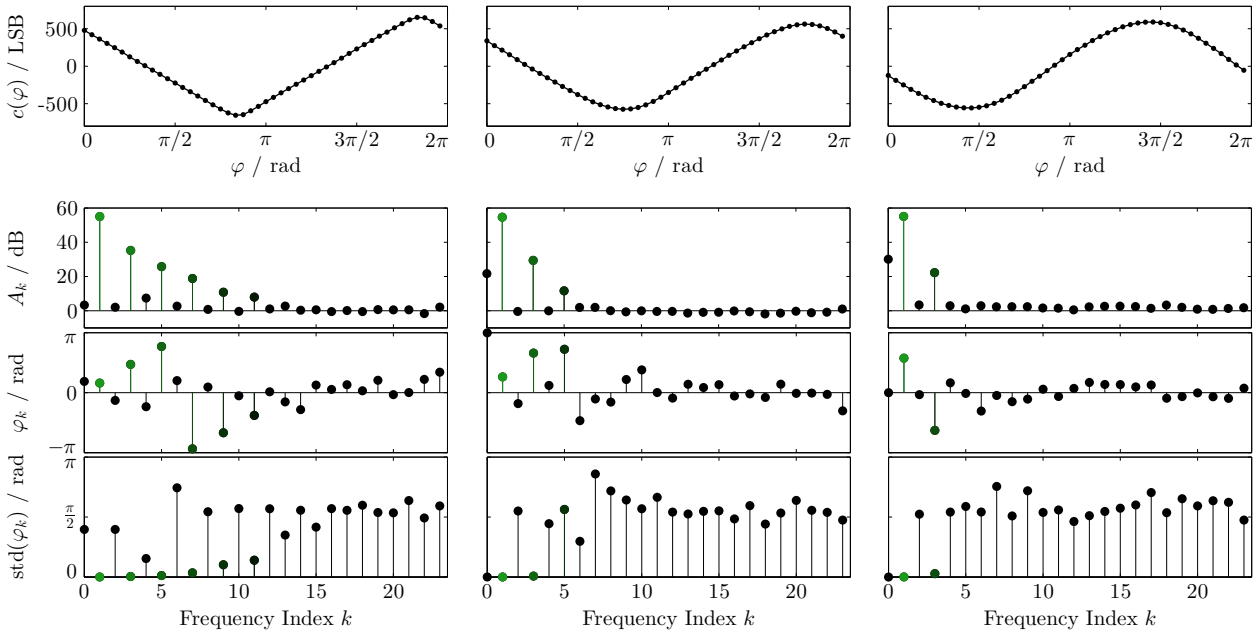
Figure 5.4: Delineation of (a) the linear dependency between sensor temperature and measured depth with a proportionality factor of $k = 0.249$ between the two quantities, and (b) the acquired data before and after linear correction applied. [$f_{\mathrm{mod}} = 20$ MHz, $\tau = 1000\,\mu$s, $N_E = 48$, $K = 32$]

would have been desired. In further long-term measurements it was observed that the measured depth appears to drift with the *ambient* temperature as well. Moreover, it was shown that these drifts can be compensated by employing data from on-chip temperature sensors, at least for measurements at room temperature around 20 °C. Although a linear correction model solved this issue sufficiently well, further investigations are certainly necessary to draw more general conclusions from observed dependencies. However, in the present scenario with applied corrections, a final standard deviation of the array mean in the range of 0.02° was achieved, as shown in Fig. 5.4 (corrected curve).

Fourier Analysis

For modulation frequencies of 5, 20, and 35 MHz, several Fourier analyses under different circumstances were conducted, such as with varying target reflectivity, target distance and integration time. Qualitatively, they appeared to be consistent, so for each examined modulation frequency the results of a representative pixel are given in Fig. 5.5.

From the magnitude spectra (second row), it can be seen that the dominance of odd harmonics impressively well coincides with the findings from theory that triangular signals indeed only consist of odd harmonics. Moreover, the low-pass characteristic of the illumination unit can clearly be seen[5]. This implies that most – if not all – information is contained in odd harmonics,



(a) 5 MHz (b) 20 MHz (c) 35 MHz

Figure 5.5: Fourier spectra for exemplary modulation frequencies. The coefficients supposed to carry information are indicated with green color. Basis for computing mean and standard deviation were 32 samples. [$f_{\mathrm{mod}} = 5/20/35$ MHz, $\tau = 650/700/1000\,\mu$s, $N_E = 48$, $K = 32$]

[5] The interpretation that the observed low-pass behavior of the system in fact stems from the illumination unit was supported by additional measurements of the optical signal itself, which already exhibited the same characteristic. Furthermore, relying on statements of the analog design team, the produced chip is supposed to handle signals of the given frequency range with ease.

an observation that is in particular supported by the phase spectrum (third and fourth row). There, from the standard deviation (STD) it is evident that significant information is only provided by odd spectral components, with decaying reliability as the order increases.

Thus, based on these findings the following fundamental conclusions can be drawn:

- The SRF actually contains all of its exploitable information in its *odd* harmonics. This suggests that any reconstruction model should rely on these harmonics (e.g., model order 3 comprising first and third harmonic, model order 5 comprising first, third and fifth harmonic, etc.). This is not only a simplification but should as well suppress unwanted distortion that may appear in even harmonics.

- The lack of spectral distortion (i.e., the absence of even harmonics) is a strong indicator that the PMD pixel approximates its ideal linear behavior very well. It should be remarked that in the special case of polynomial nonlinearities, the amount of spectral distortion that occurs in even harmonics relates to the DC offset of the SRF. This stems from the fact that polynomial nonlinearities can be represented as convolutions in the frequency domain. It is evident that if a signal contains only odd harmonics, no energy can be leaked to even harmonics as the respective convolution sums amount to zero. However in the case of DC offset, these sums can be nonzero.

- The highest harmonic containing information was observed to be around the 11^{th} at 5 MHz, the 5^{th} at 20 MHz, and the 3^{rd} at 35 MHz. For aliasing-free determination of the frequency components, at least 48-fold sampling was used throughout this Thesis.

- As expected, pixel-dependent effects were observed, such as constant offset (cf. A_0 in Fig. 5.5) and linear gain.

5.2 Moving-Target Experiment

At this stage, the results of the moving-target experiment are presented. Several questions are needed to be answered and tasks to be carried out, such as determining the required model order M needed for consistent depth sensing (cf. Section 5.2.1), further investigations of the Wiggling Error (cf. Section 5.2.2), and as well the evaluation of the approximative estimation algorithms derived in Section 3.2.3.

5.2.1 Finding the Right Model Order

As discussed, the moving-target experiment should give new insights which model order is actually necessary to reconstruct reliable depth information. Since it is known from the previous section that the SRF does actually not consist of the sole fundamental wave but carries information in its odd harmonics, models of several orders were employed (cf. (3.26)), whose argmax-estimation

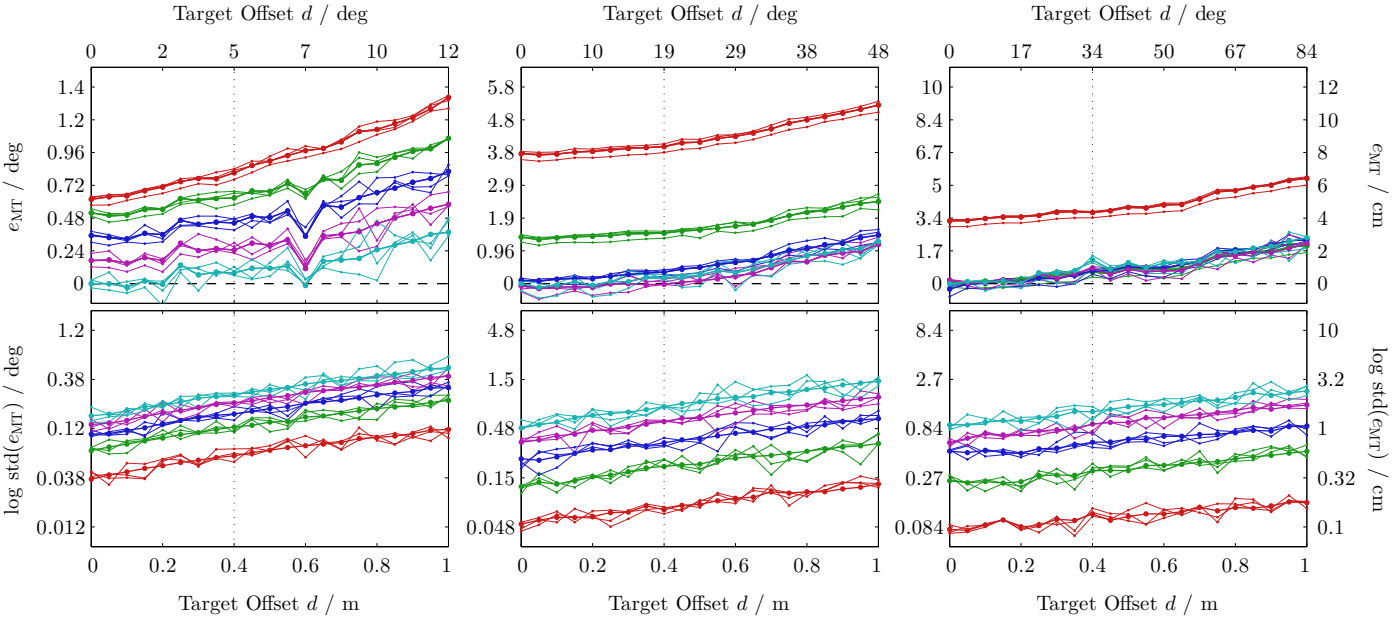
errors

$$\begin{aligned} e_{\mathrm{MT}}(f_{\mathrm{mod}}, M, d) &= \hat{d}_M - d \\ &= -\operatorname{argmax}\{c_M(\varphi)\} \cdot \frac{\mathrm{c}}{4\pi f_{\mathrm{mod}}} - d \end{aligned} \tag{5.2}$$

and the respective standard deviation are depicted in Fig. 5.6. Three individual pixels (thin lines) are given for each model in order to show their accordance to the overall array mean (region of interest (RoI): 11 × 11 pixels, thick lines). The vertical dotted lines indicates the correspondence to the steady-target experiment (cf. Section 5.3.1), in particular to Fig. 5.9. Argmax-estimation was implemented by means of an iterative search for the respective zero-crossing of the model derivative (cf. (3.25)).

As an arbitrary reference, the estimation error was set to zero for maximum model order and smallest distance. Furthermore, the distance parameter d is given as an offset *relative* to the initial target distance of 60 cm. These two generalizations reflect the assumption that at the current stage of investigation, absolute quantities are of limited interest and should not detract from more general system properties.

It should be noted that in the course of this experiment, the reciprocal proportionality between signal amplitude and distance known from theory (i.e., $g(d) \propto 1/d$ (3.2)) was experimentally shown.



(a) 5 MHz (b) 20 MHz (c) 35 MHz

Figure 5.6: Error and standard deviation of the argmax-estimates with respect to target offset d. For each modulation frequency, an appropriate integration time was needed, such as 650 µs at 5 MHz, 700 µs at 20 MHz, and 1000 µs at 35 MHz. The centimeter-scale is constant whereas the degree-scale varies according to the modulation frequency. Angular and metric units relate according to (2.2). $[f_{\mathrm{mod}} = 5/20/35\ \mathrm{MHz}, \tau = 650/700/1000\,\mu\mathrm{s}, N_E = 48, K = 32]$

The apparent characteristics of the captured data depicted in Fig. 5.6 may be summarized as follows:

- The probably most essential observation is that, invariant to modulation frequency and target offset, the estimation errors are *consistently* offset to each other. This indicates that the optical path indeed approximates the ideal delay quite well (as modeled in (3.2)). Consequently, the shape of the SRF does not change. In other words, the consistency of the depth estimates over all model orders and distances directly implies that, in turn, the Fourier coefficients maintain their rigid relationships and thus the shape of the represented signal is preserved. Concluding, higher-order estimation appears *not* to deliver more precise distance values than first-order estimation does.

- For higher modulation frequencies, higher-order estimates converge against the true maximum of the SRF, whereas at 5 MHz, this convergence is not reached because the highest model order is still not capable of resolving all harmonics present in the SRF. However, for even higher model orders this convergence can be certainly assumed as well.

- In quantitative terms, the offset between estimates of different model order and the true maximum of the SRF highly depends on the modulation frequency. Different modulation frequencies cause the illumination unit to shape the signal according to its innate low-pass characteristic, which potentially alters this offset. If depth estimation is realized via lower-order models, this deviation must be known and calibrated.

- Over all modulation frequencies and model orders, a systematic depth drift with increasing distance is visible. The fact that, at a certain modulation frequency, *all* depth estimates drift in the same manner is a strong indicator that the SRF is actually shifted *in its entirety*. In turn, any static nonlinearities that may affect distance estimation at lower amplitudes (as for larger distances) *cannot* be responsible for this drift (cf. (3.13)). Investigations of the definite root cause for this drift are not essential at this stage but will come necessary at later stages.

- Regarding standard deviation, lower-order estimates provide robuster information than their higher-order pendants. This is not surprising, as higher-order models rely on more parameters (estimated via the DFT), of which each introduces additional uncertainty. In particular, this increased sensitivity of higher-order models is visible at the outlier at 5 MHz and 0.6 m relative target offset.

- In terms of *angular* reliability of the depth estimates, lower modulation frequencies perform slightly better than higher ones. The most likely source for this frequency-dependent degradation is jitter. Jitter is considered to introduce temporal fluctuations that are in general independent of the modulation frequency. However the impact on *amplitude* values directly relates to the modulation frequency, which is an implicit outcome of (3.10). Thus, for higher modulation frequencies, noise as a result of jitter increases. However, the only small performance decrease towards higher modulation frequencies is an indicator that jitter causes only a minor contribution to the overall statistical error.

- For lower-order models and higher modulation frequencies it is observable that the individual pixels have a constant offset to each other, which obviously refers to FPN. Thus, the moving-target experiment inherently proves the consistency of FPN over distance and amplitude.

5.2.2 Wiggling Error

One main outcome of the preceding analysis of the moving-target experiment is that – limiting the scope to target distance and modulation frequency for now – the phase of the fundamental wave indeed appears to contain enough information needed for depth sensing. However, it is the violation of the sampling theorem that causes the major part of distance-dependent depth errors in the state-of-the-art 4-phase algorithm. Thus, for better understanding and besides just reproducing the Wiggling Error, the main objective of this subsection is to determine how well this error can actually be *simulated* and how well these analytic findings fit into real-world measurements. In particular, simulating the Wiggling Error is important for two reasons:

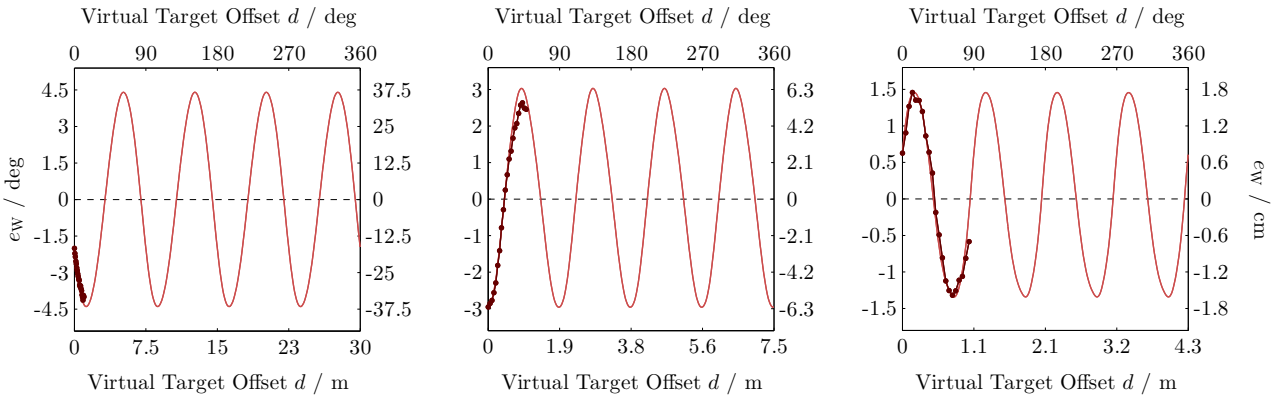
- Due to limited space in the laboratory, measurements over the whole ambiguity range of the respective setting (e.g., 30 m at 5 MHz modulation frequency) are not realizable. Moreover, even if the necessary space was available, one thorough measurement would still not be possible without increasing the integration time considerably to compensate for the optical attenuation. This however is undesirable as changes of integration time are known to introduce additional deviations (cf. integration-time related error). In contrast to these natural limitations, the acquisition of representative data is mandatory for any quantitative claims about the impact of Wiggling. So if simulation succeeds and can be shown to represent the measured error, the aforementioned dilemma can be avoided. Fortunately, the Wiggling Error is known to be a zero-mean oscillation with constant amplitude and frequency over the whole uncertainty range. Therefore, it should be an easy task to verify whether measured data that only covers a fraction of the uncertainty range is correctly represented by a simulation over the whole depth range.
- If the Wiggling Error is understood and can be simulated, the analytic model may be employed to provide fictive but reliable performance analyses of any proposed strategy that aims at reducing this error.

Given the fact that the shape of the SRF does indeed not change with the target distance (as observed in the previous section), the analytic description of the simulated Wiggling Error in Section 3.3.1 is valid and can be deployed for the given purposes. In contrast to the definition of the simulated Wiggling Error (3.48), its *measured* pendant may be defined as

$$e_{\mathrm{W}}(d) = \arg(\hat{C}[1]) - \arg(C[1]) \qquad \text{with target at distance } d \tag{5.3}$$

where $\hat{C}[1]$ is determined from 4-fold sampling and $C[1]$ the true Fourier coefficient. The fact that the notation e_{W} both represents simulated and measured Wiggling Error should not lead to confusion as it is – if not clear from the context – explicitly referred to either of them.

The plots in Fig. 5.7 show how well measured and simulated Wiggling Error coincide. Although in the experiment, the optical signal amplitudes got considerably low for larger distances, only



(a) 5 MHz (b) 20 MHz (c) 35 MHz

Figure 5.7: Measured and simulated Wiggling Error for exemplary modulation frequencies. [$f_{\mathrm{mod}} = 5/20/35$ MHz, $\tau = 650/700/1000\,\mu$s, $N_E = 4$, $K = 32$]

small deviations are visible, most noticeable at 20 MHz. However, as the Wiggling Error of 4-fold sampling takes contributions from all odd harmonics (cf. (3.43)), the shown results once more indicate that the spectral composition of the SRF remains consistent over distance, at least for the odd harmonics (which have been found to carry the information).

Mod. Freq.	**Wiggling Error**	
MHz	deg	cm
5	4.41	36.72
20	3.03	6.31
35	1.46	1.73

Table 5.1: Wiggling Error (max. amplitude) for the examined modulation frequencies.

In Tab. 5.1, the Wiggling Error is given for the respective modulation frequencies. Note that this error is usually given via its maximum value throughout this Thesis.

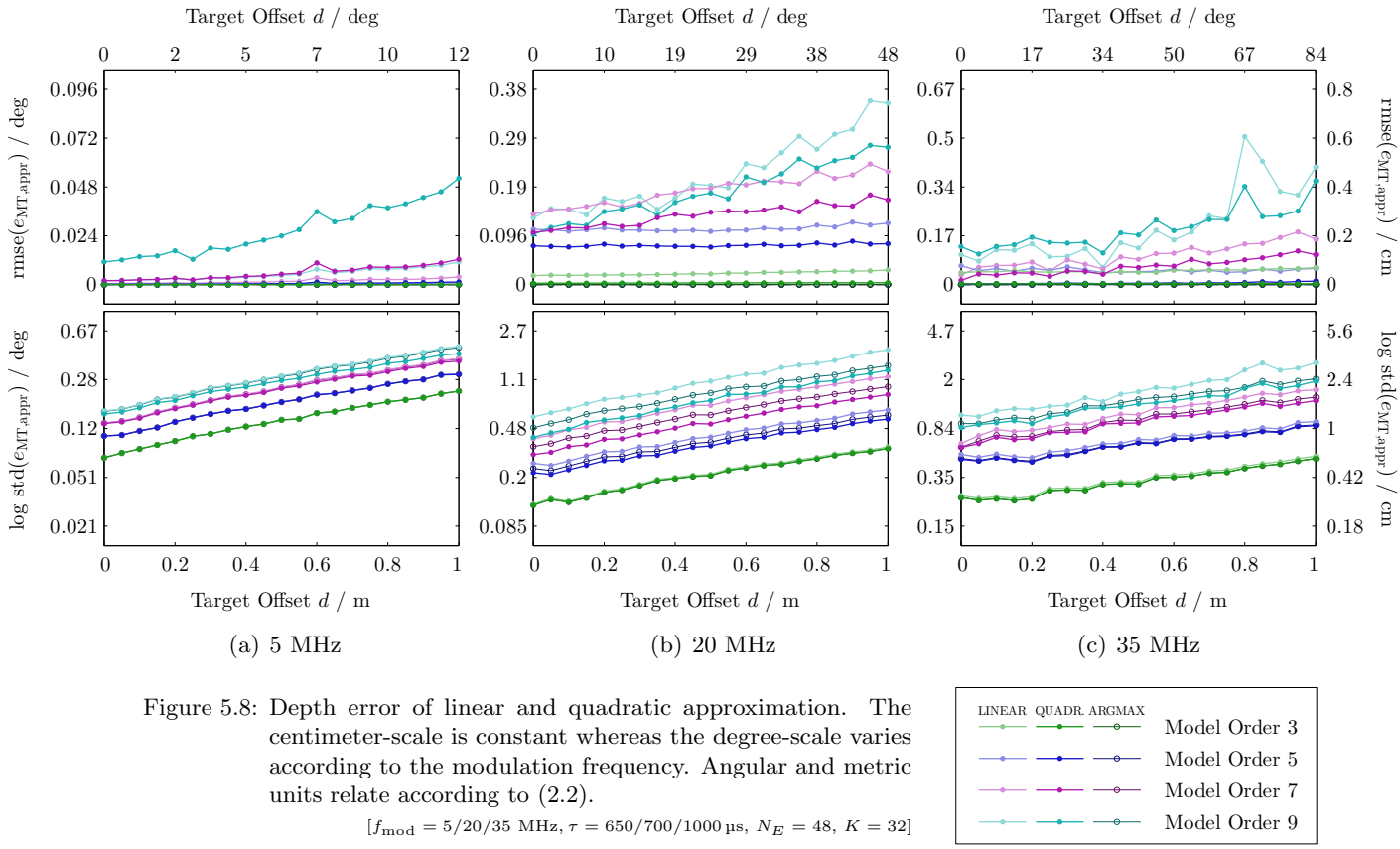
5.2.3 Evaluation of Approximative Approaches

With the real argmax-estimates presented and discussed in Section 5.2.1, it is of interest to assess the performance of approximative solutions (as derived in Section 3.2.3) as well. With $\hat{d}_M$ referring to the distance determined by actual argmax-estimation of model order M (cf. (5.2)), the deviation of linear respective quadratic approximation may be denoted as

$$e_{\mathrm{MT,appr}} = \hat{d}_{\mathrm{appr}} - \hat{d}_M \tag{5.4}$$

with $\hat{d}_{\mathrm{appr}}$ being the approximative argmax-estimate.

In Fig. 5.8, the approximation error is given as an RMSE over the whole array of 11×11 pixels. As a measure of the reliability of the methods in general, the standard deviation of the respective method is provided as well, with a computation basis of 32 samples.



(a) 5 MHz (b) 20 MHz (c) 35 MHz

Figure 5.8: Depth error of linear and quadratic approximation. The centimeter-scale is constant whereas the degree-scale varies according to the modulation frequency. Angular and metric units relate according to (2.2).
[$f_{\mathrm{mod}} = 5/20/35$ MHz, $\tau = 650/700/1000\,\mu$s, $N_E = 48$, $K = 32$]

First of all, the most apparent characteristic is that the proposed approximation methods perform better for lower model-orders. This is no surprise, because as for regular argmax-estimation, higher-order models are parameterized by more coefficients that introduce additional noise. Considering the scale of the y-axis, the algorithms perform astonishingly well and are orders of magnitude below the real deviation of the argmax (cf. Fig. 5.6). The expectation that quadratic approximation is superior to linear approximation due to additional complexity is not valid in general. This can be observed in particular for higher-order models and lower modulation frequencies. The reason for this behavior may be attributed to the computation of the zero-crossing for quadratic approximation. In fact, the zero-crossing of the first-order Taylor approximate could be further away from the true argmax than the linear estimate is. This is supported by the fact that at lower modulation frequencies, the maximum of the fundamental wave is actually very close to the true argmax of the SRF (cf. Fig. 5.6). In turn, linear estimation would be already a very good approximation.

Regarding precision, it can be seen that the approximative solutions feature a similar reliability as the respective argmax-estimates, which were added in the plots for reference. In particular at 20 MHz modulation frequency and for higher-order models, it can be observed that quadratic estimation even outperforms argmax-estimation, albeit its supremacy is admittedly small.

For completeness, the overall RMSE (thus over the target distance range and sensor array) of linear and quadratic approximation versus model order and modulation frequency are provided in Tab. 5.2. The values in the upper table are given in degrees, whereas the lower ones are given in centimeters.

Model Order	Linear			Quadratic		
	5 MHz	20 MHz	35 MHz	5 MHz	20 MHz	35 MHz
3	3.67E-5	0.0222	0.0471	5.77E-6	3.10E-3	2.81E-3
5	3.96E-4	0.110	0.0498	6.36E-4	0.0771	5.84E-3
7	1.96E-3	0.188	0.107	6.63E-3	0.138	0.0675
9	5.82E-3	0.229	0.233	0.0295	0.191	0.203

VALUES IN DEGREES

Model Order	Linear			Quadratic		
	5 MHz	20 MHz	35 MHz	5 MHz	20 MHz	35 MHz
3	3.06E-4	0.0463	0.0561	4.81E-5	6.46E-3	3.34E-3
5	3.30E-3	0.229	0.0592	5.30E-3	0.161	6.96E-3
7	0.0163	0.392	0.128	0.0553	0.288	0.0804
9	0.0490	0.477	0.277	0.246	0.399	0.242

VALUES IN CENTIMETERS

Table 5.2: RMSE of linear and quadratic approximation versus model order and modulation frequency.

5.2.4 Conclusion

The moving-target experiment revealed some fundamental properties of ToF depth-sensing. Given the target distance d and modulation frequency f_{mod} as experiment parameters, it was shown in Section 5.2.1 that for a certain modulation frequency, depth estimation of higher-order models is not superior to lower-order models – contrarily, the latter are much robuster with respect to noise. As a consequence, it appears to be sufficient estimating the phase of the fundamental wave and then correcting for the (constant) offset between estimated phase and true maximum. This is a dramatic simplification of the general distance estimation problem formulated in Section 3.2.3. However, an unresolved issue still remaining is the systematic depth drift of all estimates (i.e., of the whole SRF) with increasing target distance. It was argued that nonlinearities cannot responsible for this drift.

In Section 5.2.2, the Wiggling Error was examined and two possibilities to simulate this error were derived. The first one bases on a sample-wise shift of one externally measured SRF, the other one on a pure analytic approach. In Fig. 5.7 it was shown that simulation perfectly well accords to actual measurements. As expected, the Wiggling Error got smaller with higher modulation frequencies due to the decrease of disturbing harmonics (cf. Tab. 5.1).

The performance of approximative methods (derived in Section 3.2.3) was assessed in Section 5.2.3. It turned out that both methods, linear and quadratic, surprisingly well approach the argmax-estimate, with decreasing accuracy of higher-order models. For example, linear approximation of the third-order argmax shows an RMSE in the range of 10^{-2} degrees throughout all examined frequencies, or even below. Quadratic approximation decreases the RMSE by another order of magnitude to about 10^{-3}. No increased sensitivity to noise was observed.

5.3 Steady-Target Experiment

Motivated by the results of the moving-target experiment from the previous chapter and the questions from Section 4.4.3, the steady-target experiment with varying integration time τ, modulation frequency f_{mod} and target reflectivity r, but constant target distance $d_0 = 60\,\mathrm{cm}$ was conducted. In order to systematically find answers to the questions posed in Section 4.4.3, in a first step the general influence of integration time on ToF depth-sensing was examined (cf. Section 5.3.1). Secondly, Section 5.3.2 elaborates the role of varying target reflectivities on distance measurement (i.e., the amplitude-related error). Section 5.3.3 concludes the discussion about model orders with a comprehensive evaluation of the offset between first-order estimation and true argmax of the SRF.

5.3.1 Integration-Time Related Error

Similar to Section 5.2, for the steady-target experiment the estimation error of model order M is defined as

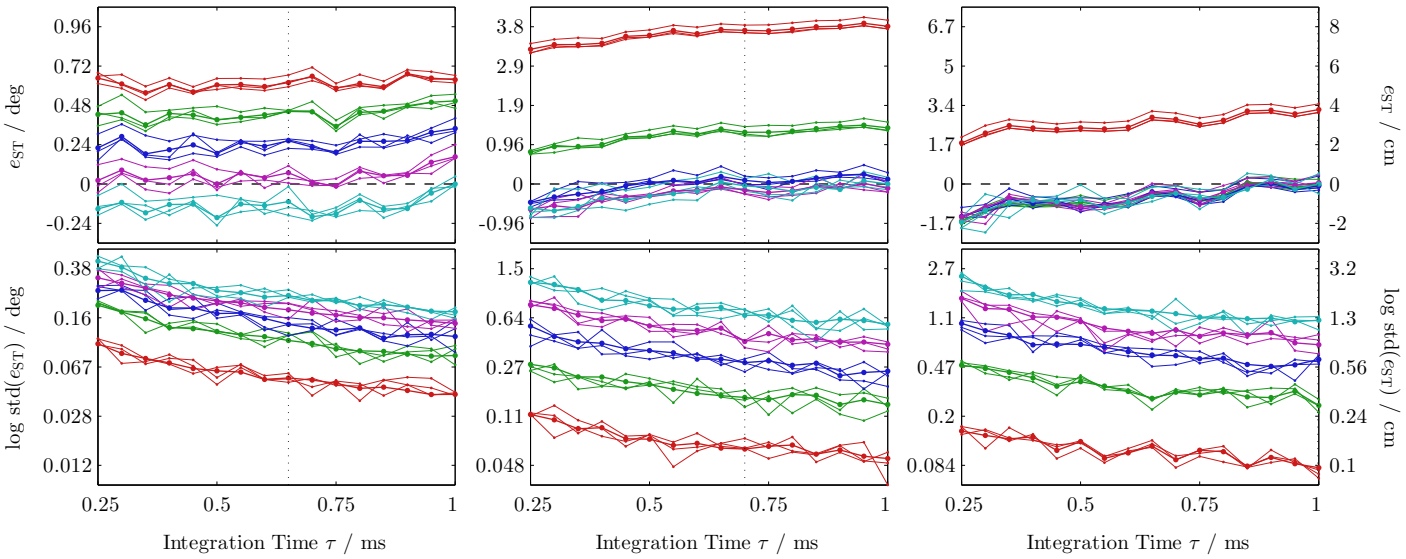
$$\begin{aligned} e_{\mathrm{ST}}(f_{\mathrm{mod}}, \tau, M) &= \hat{d}_M - d_0 \\ &= -\operatorname{argmax}\{c_M(\varphi)\} \cdot \frac{\mathrm{c}}{4\pi f_{\mathrm{mod}}} - d_0\,. \end{aligned} \tag{5.5}$$

In Fig. 5.9, this error and the according standard deviation is given over integration time and for exemplary modulation frequencies. Without loss of generality and for better comparability, the curves were offset so that the error of the highest model order and longest integration time was assumed to be zero – for which reason the shown data should be regarded as relative. Again, individual pixels (thin lines) are given towards better illustration that their behavior accords to the array mean (thick lines, RoI: 10 bright pixels).

It is evident that most of the observations discussed in Section 5.2.1 are still applicable to the current experiment. Even the data are directly comparable, with the vertical dotted lines indicating the respective cross-sections of the two figures within the parameter space. Thus, additionally to the observations already discussed in Section 5.2.1, the following points are worth a more detailed discussion:

- Apparently, different integration times appear not to change the offset between depth estimates that stem from different model orders. Nevertheless, most notably above 20 MHz, an integration-time related drift of the whole SRF can be observed. Related to the issue that temperature severely influences distance measurements (discussed in Section 5.1.2), the most reasonable explanation for this deviation is the fact that the time the illumination unit is driven is proportional to the integration time. In order to save power, it is turned off if the optical signal is not used. Consequently, the increased usage of the illumination unit increases its temperature and delays the optical signal in a deterministic manner.

- It was shown in further experiments that changes of the integration time equally affect all depth data, regardless of amplitude and distance. This is another hint that the erroneous delay is already present at time of emission into the scene.



(a) 5 MHz (b) 20 MHz (c) 35 MHz

Figure 5.9: Error of the argmax-estimates versus integration time, given for several model orders at exemplary modulation frequencies. The centimeter-scale is again kept constant whereas the degree-scale varies according to the modulation frequency. Angular and metric units relate according to (2.2).
[$f_{\mathrm{mod}} = 5/20/35$ MHz, $\tau = 250 \ldots 1000\,\mu$s, $N_E = 48$, $K = 32$]

- Although the established measurement setup features a high degree of reproducibility for a constant set of parameters, varying the integration time introduced slight, non-deterministic artifacts. For example, in the present experiment issues such as the depth drop at $f_{\mathrm{mod}} = 35$ MHz and $\tau = 0.75$ ms or the varying offset between different model orders at 5 MHz and long integration times could not be reproduced. However, since the integration-time related error can be attributed to issues concerning temperature, this is quite plausible. As temperature is a physical parameter of rather inertial nature, the current temperature of the illumination unit is not only a function of the current integration time but also of past integration times and other influences. Towards a full clarification of the relationship between integration time, illumination-unit temperature and resulting depth deviation, a temperature sensor may be integrated into the illumination unit.

- The fact that the offset between individual pixels is constant for lower-order models and higher modulation frequencies (thus for estimates with low standard deviation) proves that FPN is invariant to integration time as well.

5.3.2 Amplitude-Related Error

This section is dedicated to a fundamental discussion of the amplitude-related error. General issues are discussed in the first part, whereas an in-depth analysis of the definite root causes is carried out in the second part. Interpretations of the observed phenomena are provided in the last part.

Overview

Additionally to the introduction in Section 2.3.2, from a signal processing perspective two hypotheses seem plausible to explain depth deviations as a result of different object reflectivities:

Nonlinearities. The amplitudes of the SRF can be seen as being processed by several system components before they are actually read out as digital values. As already argued in Section 3.1.2, nonlinearities can be expected to be introduced in particular by the PMD pixel itself and the ADC. However, based on the analysis in the context of (3.13) these effects should only be visible in lower-order models since argmax-estimation is invariant to static nonlinearites. In other words, nonlinearities introduce spectral distortions that potentially cause shifts in lower harmonics whereas the overall argmax of the function remains the same.

Multipath Reflections. Given the derivation in Section 3.3.2 as a theoretic prerequisite, the reason why additional reflections cause depth deviations dependent on the signal amplitude is evident (cf. Fig. 3.2). However *identifying* the root cause of deviations that apparently correlate with the signal amplitudes is a different issue. Towards finding appropriate methods that classify any amplitude-related error as being the result of multipath reflections, two major implications from the model discussed in Section 3.3.2 may be exploited:

- As shown by (3.57), any additional reflection path incident on the sensor affects all spectral components of the SRF. This does *not* necessarily mean that the argmax is shifted as

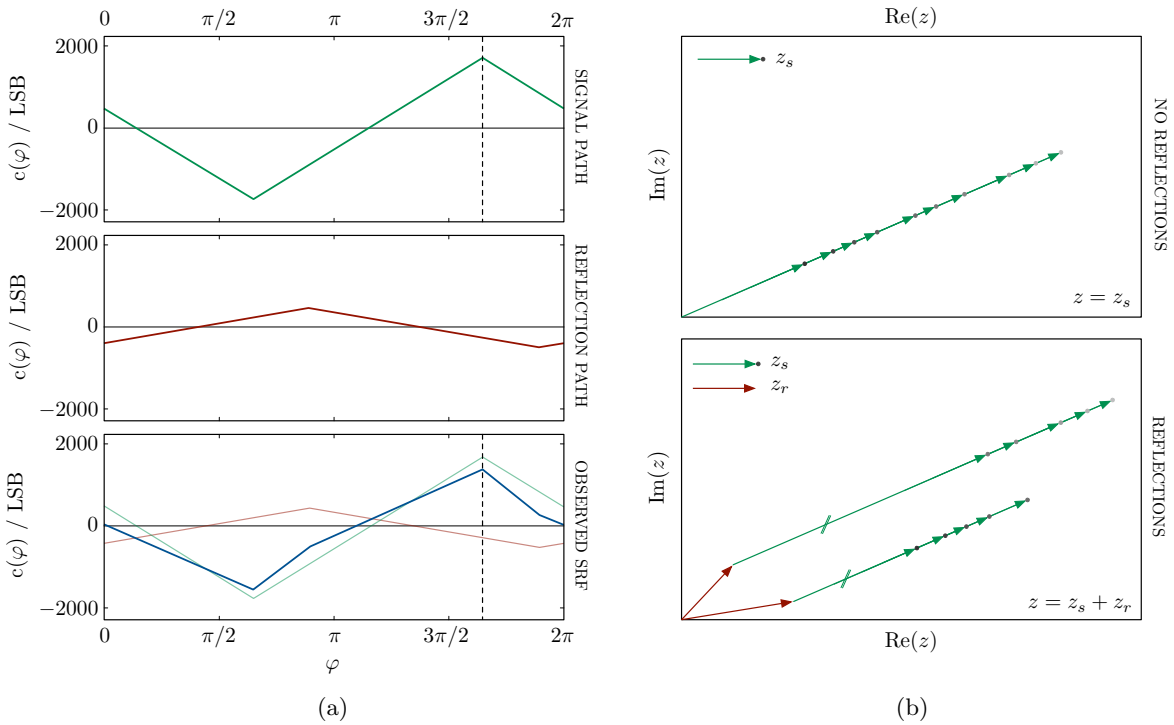


Figure 5.10: Two examples that illustrate the implications of multipath reflections: (a) a counterexample that proves that the argmax of the SRF is not necessarily shifted, and (b) the fact that pixels which refer to the same distance are mapped onto parallel lines in case of different reflections.

well, which can be proven by a quick example. It should be recognized that in case of reflections, not only the spectral composition but as well the SRF itself can be interpreted as a superposition of multipath components, which is evident from (3.49). Given just one reflection and a triangularly shaped SRF, it can be shown that independent from the phase of the reflection, this does not offset the argmax of the SRF unless the amplitude of the reflected signal exceeds the amplitude of the desired signal. An illustration is shown in Fig. 5.10(a). However it is possible (if not very likely for higher modulation frequencies) that the argmax of the SRF shifts as well. As the case of a shifting argmax is excluded by nonlinearities, this can be used as a discrimination between these two hypotheses: If shifts of the true argmax can be observed, the most probable reason for this deviation are multipath reflections. Of course, the reverse implication is not true.

- Referring to one single spectral component of the SRF, it is evident from (3.57) that if only the signal phasor changes its amplitude (as in the case of varying reflectivity), the observed phasor resides on a straight line in the complex plane. Hence, if additional reflections are present potential offsets of this line are introduced, but the orientation of the line is maintained. Consequently, given a measurement of several pixels that refer to points of varying reflectivity but constant distance, these points will be mapped onto parallel lines in the complex plane, where each line represents a distinct set of reflections. Fig. 5.10(b) illustrates this principle.

Measuring the amplitude-related error turned out to be more difficult as expected. Using a target with different shadings (0%, 25%, 50%, 75%, 100% black) and all possible transitions between these shadings, deviations that evidently originate from different intensities were not visible in the first place (cf. depth map in Fig. 5.11 (top) where, for better illustration, the data is overlain with the intensity image of the sensor).

For two reasons, a quantification of the error is impracticable without any further measures. First, FPN causes lateral depth deviations that are in the same order of magnitude as the amplitude-related error. Second, due to the fact that without any conversion ToF sensors deliver radial but not orthogonal distances, planar surfaces appear curved which adds additional difficulty in evaluating the impact of amplitude-based deviations.

As a remedy, it was found that capturing the error in a differential manner provided the most reliable and intuitive basis for further analyses. In other words, this involved a two-stage capturing process: One reference measurement of a white target at the same distance (cf. Fig. 5.11 (middle)) was taken and its resulting depth values were then subtracted from the original data. In Fig. 5.11 (bottom), this differential depth map is provided. It can be seen, that indeed a correlation between depth deviation and amplitude can be observed. The amplitude-related error is defined as

$$e_{\mathrm{AR}} = \hat{\varphi}_d - \hat{\varphi}_{d,\mathrm{ref}} \tag{5.6}$$

where $\hat{\varphi}_{d,\mathrm{ref}}$ refers to the estimated depth of the reference measurement.

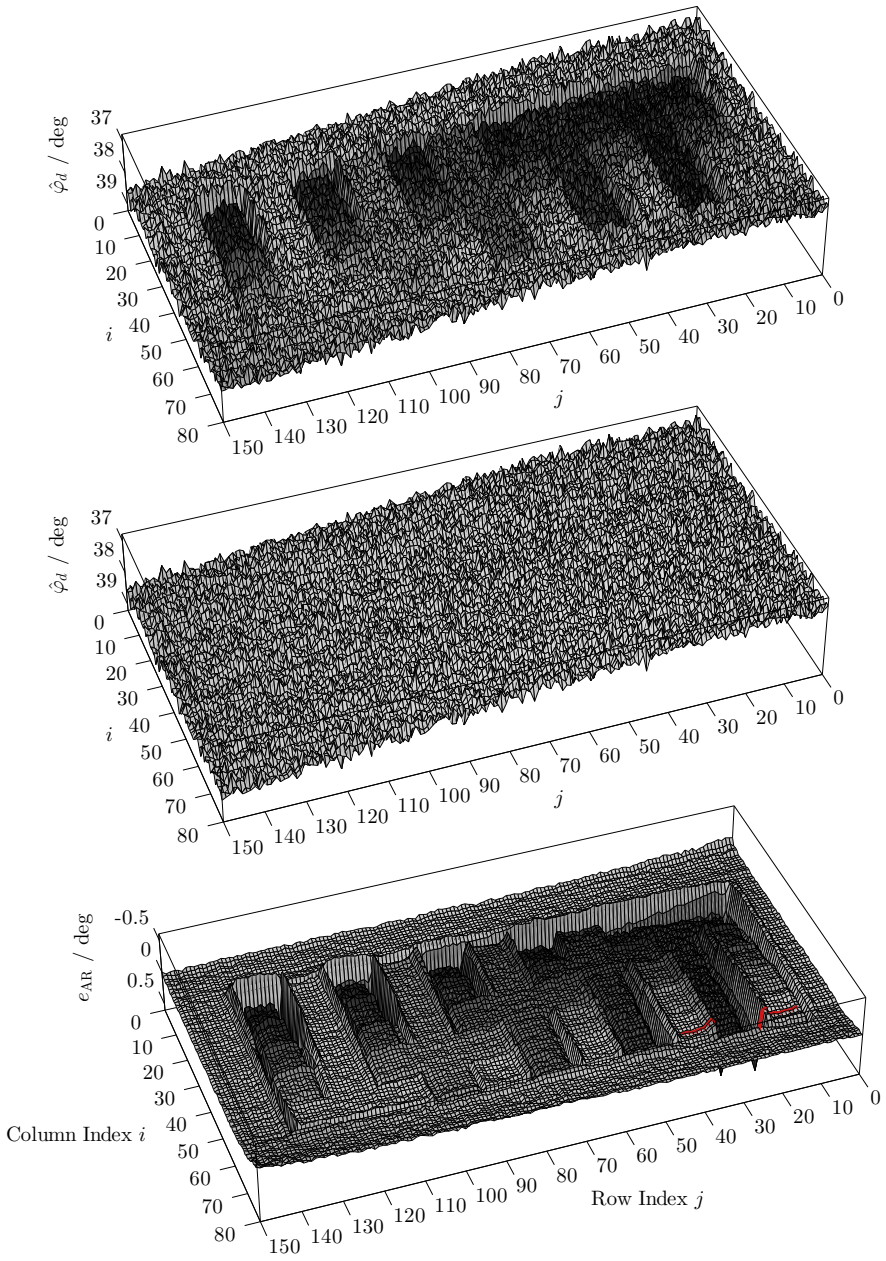


Figure 5.11: Example of the amplitude-related error. As the z-axes were inverted for a more realistic view, the shown data can actually be interpreted as the front view of the target. The indices i and j indicate the pixel location; φ_d refers to first-order depth estimation.

[$f_{\mathrm{mod}} = 20$ MHz, $\tau = 1000\,\mu$s, $N_E = 120$, $K = 32$]

Detailed Analysis

Whereas Fig. 5.11 gives only a rough impression of how the amplitude-related error behaves for the first-order estimate at 20 MHz and an integration time of 1000 µs, it is of high interest to assess its behavior with respect to a variety of modulation frequencies and integration times, as suggested by the steady-target experiment. For reasons of limited space and time, a pixel subset that covers harsh intensity jumps was selected from the whole target image as a representative example for further analyses, as indicated by the red line in Fig. 5.11 (bottom).

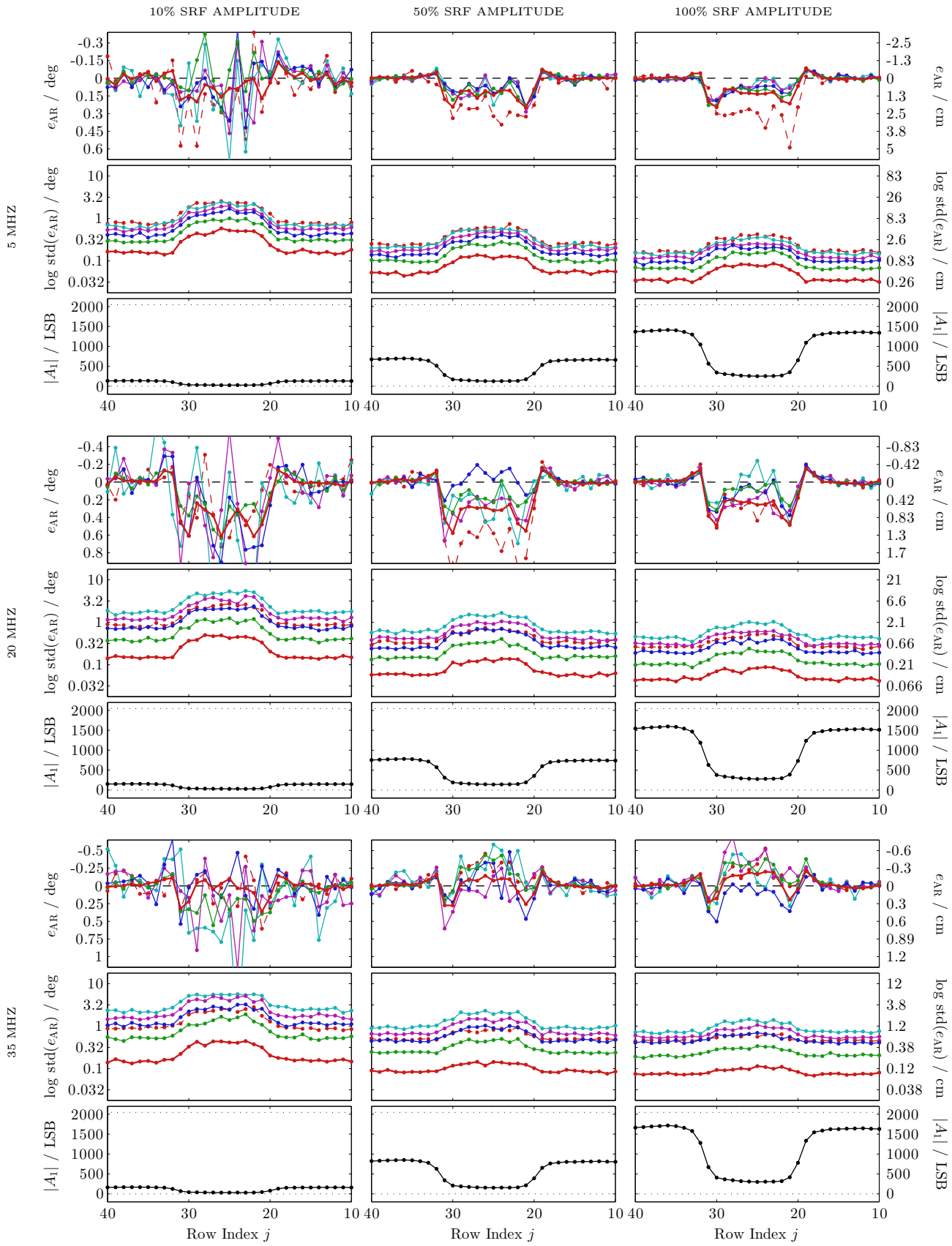


Figure 5.12: Comprehensive evaluation of the amplitude-related error with respect to several modulation frequencies and integration times. For comparison, depth estimates of the regular 4-phase algorithm are included in the plot. Computation basis for mean and standard deviation were 64 realizations.

$[f_{\mathrm{mod}} = 5/20/35\ \mathrm{MHz}, \tau = 110 \ldots 1900\,\mu\mathrm{s}, N_E = 120, K = 64]$

Mod. Freq.	**Integration Time**		
MHz	µs		
	Left Col. (10%)	Center Col. (50%)	Right Col. (100%)
5	110	550	1100
20	140	700	1400
35	190	950	1900

Table 5.3: Integration times used in the experiment presented in Fig. 5.12.

With the amplitude-related error being evaluated at 5, 20 and 35 MHz at respective integration times that produce comparable SRF amplitudes (cf. Tab. 5.3), the obtained results are depicted in Fig. 5.12. Depth estimates of several model orders should deliver insights whether changing intensities cause the true argmax to deviate – which motivates the theory of optical effects as a root cause.

It should be remarked that given the SRF at its maximum amplitude (i.e., with its peak near 2048 LSB) in the right column, the middle column refers to 50% and the left column refers to 10% amplitude. Note that the amplitude values are given with respect to the fundamental wave, $|A_1|$, which will not approach 2048 LSB as the true maximum of the SRF does. Furthermore, for better comparability between 2D and 3D plots, the depth axis is reversely displayed in either modes. Since lower depth values refer to closer objects, these figures can be interpreted as an upside-down view onto the target.

The plots in Fig. 5.12 exhibit some significant characteristics that deserve a thorough discussion:

- As expected, the standard deviation of depth estimates rises with decreasing SRF amplitude. This causes higher-order estimates to be considerably unreliable in areas of lower reflectivity, especially for shorter integration times where the SRF amplitude falls below 200 LSB even in bright areas.

- Regardless the fact that the SNR is particularly low for higher-order estimates, even at longer integration times they do not appear to deliver systematically better depth estimates. Based on the discussion above, this is a clear suggestion that optical effects may be responsible for the seen deviations.

- A very characteristic behavior of the amplitude-related error – which was already observed in Fig. 5.11 – is its jumping behavior at intensity changes. Especially visible at 20 MHz and longest integration time but also over the whole set of plots, it is evident that these jumps affect estimates of all orders. As can be seen in the respective lower row, i.e., the SRF amplitudes that show rather smooth intensity transitions, this behavior clearly opposes any correlation to amplitude values. Naturally, nonlinearities depending on SRF amplitudes cannot be responsible for this effect.

- Referring to first-order depth estimates, for a given modulation frequency it can be seen that – besides varying SNR – the error maintains its characteristic over all integration

times. This additionally repudiates any reasonable mapping between depth error and SRF amplitude.

- Regardless the jumping-edge behavior discussed above, within the zone of low reflectivity the observed error actually changes in an unrecognized manner with the modulation frequency (about 0.15° at 5 MHz, 0.2° at 20 MHz, −0.15° at 35 MHz). Apart from that, also the *shape* of the error changes, with the jumping-edge effect being rather latent at 5 MHz but dominant at 35 MHz.

- Depth estimates obtained by the 4-phase algorithms do not appear to systematically perform worse. The additional deviations may be attributed to the considerably increased noise (as only 4 samples are taken instead of 120), or – if the SRF is indeed shifted as a whole – to Wiggling artifacts.

Summarizing these observations, no evidence for nonlinearities can be found. It is a matter of further investigation however to find explanations for the jumping-edge behavior at changing intensities. A preliminary step towards better understanding would be the isolation of these border-effects. Since the target used in this experiment comprises rather small areas of different reflectivity, it can't be excluded that Fig. 5.12 shows the overlapping of several effects that stem from intensity changes either from the left or right border. For this reason, the whole experiment was conducted again with a target of a single black rectangle on white background, but of larger area. As was seen that integration time does not change the characteristic of the error, for best SNR the integration time was set to its respective maximum (cf. Tab. 5.3).

The results of the repeated steady-target experiment are depicted in Fig. 5.13, which delivers valuable insights into the behavior of depth deviation within homogeneous areas of low reflectivies:

- Surprisingly, at 5 MHz first-order estimation appears to deliver by far the most accurate depth data - superior to any higher-order estimation. This reflects the fact that the argmax of the SRF is erroneously drifting, whereas the fundamental wave remains consistently at its place. Most likely, this behavior can be explained by the fact that low reflectivities decrease the SNR dramatically which leads to considerable deviations for higher-order estimates.

- For higher modulation frequencies, depth estimates are systematically offset within the dark area. Regardless of the model order, shifts towards the camera can be perceived. This stays in contrast to Fig. 5.11 and Fig. 5.12.

- It is evident from the current experiment that the amplitude-related deviations observed at earlier stages can be attributed to jumping-edge effects. Again, the fact that all model orders are affected is a strong hint towards multipath reflections as a root cause.

The analyses of the two-fold conducted steady-target experiment with emphasis on the amplitude-related error have shown that the hypothesis of multipath reflections being responsible for amplitude-dependent depth deviations is by far the more probable of the two theories. In neither of the two experiments any evidence for problematic nonlinearities was found, at least within the

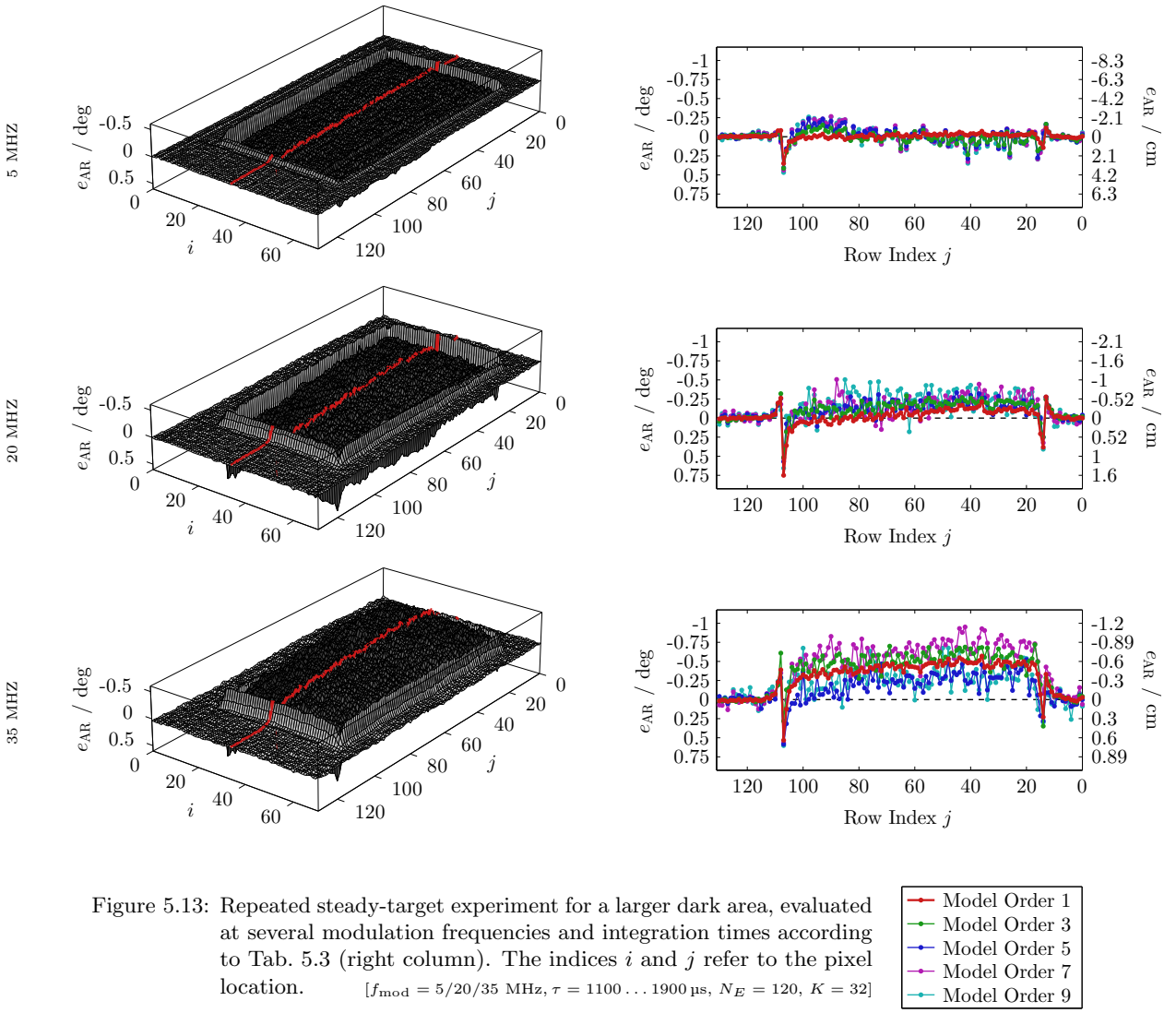


Figure 5.13: Repeated steady-target experiment for a larger dark area, evaluated at several modulation frequencies and integration times according to Tab. 5.3 (right column). The indices i and j refer to the pixel location. $[f_{\mathrm{mod}} = 5/20/35\ \mathrm{MHz}, \tau = 1100 \ldots 1900\,\mu\mathrm{s}, N_E = 120, K = 32]$

investigated order of magnitude. Therefore, a final inspection at 35 MHz (cf. Fig. 5.13 (bottom)) is carried out for additional clarity whether the model of multipath reflections established in Section 3.3.2 is indeed valid or needs to be reworked.

It is pointed out in Fig. 5.10(b) that, due to the fact that the analytic model of multipath reflections assumes a superposition of phasors in the complex plane, pixels that refer to constant distances but varying reflectivities are imaged onto parallel lines in the complex plane. As the steady-target experiment refers to this particular scenario, it is an appreciated opportunity to experimentally prove the theoretic concept. Fig. 5.14 shows that indeed, the captured target is imaged onto parallel lines. Especially in the zone of transitions between black and white, depicted in VIEW 2, this parallelism is manifest – which in turn proves that also the jumping-edge behavior can be fully attributed to multipath reflections.

It is clear that although only one particular example is evaluated, the obtained proof of the multipath-reflection concept generalizes to all modulation frequencies and integration times since the characteristic of the error is consistent over these parameters. In fact, no reasonable explanation exists why the presence of reflections should change for different camera settings.

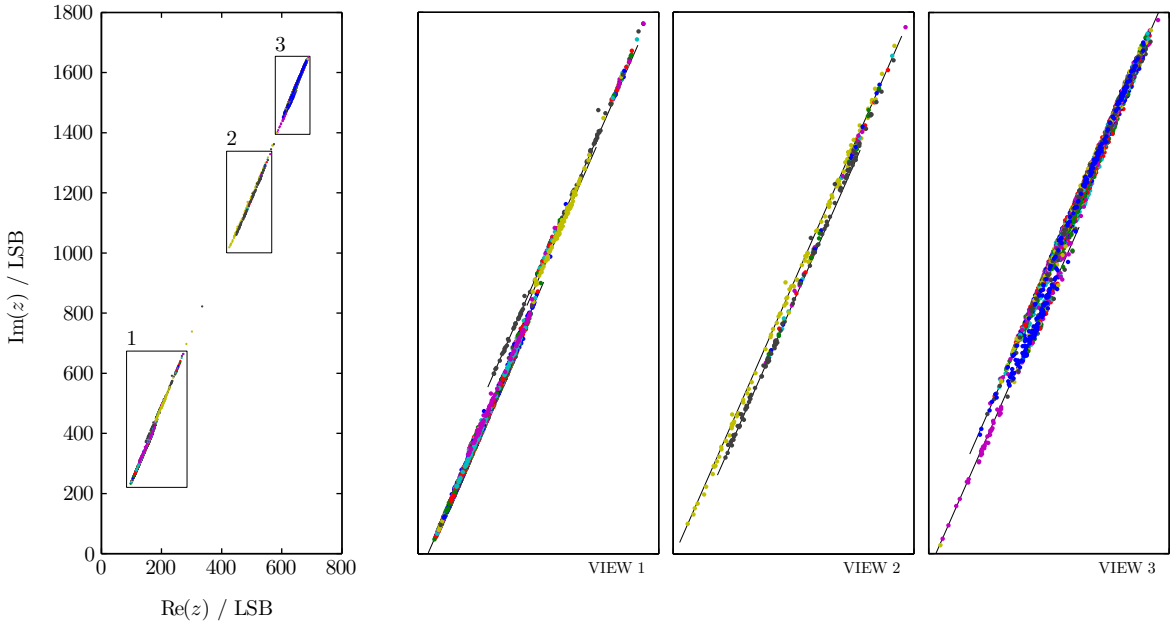


Figure 5.14: Experimental proof of the analytic multipath-reflection model, evaluated by means of the steady-target experiment at 35 MHz (cf. Fig. 5.13 (bottom)). VIEW 1, 2 and 3 refer to the respective sections in the left plot. Parallel point sets are emphasized by underlain lines. Clusters of the same color refer to the same columns. Note that the shown data is FPN-corrected.

Interpretation of the Jumping-Edge Behavior

Over all integration times and modulation frequencies, a characteristic discontinuity of depth values at intensity changes was observed. In the context of Fig. 5.14, this jumping-edge behavior was attributed to multipath reflections. The fact that this phenomenon was observed in several measurements and independent of the scene suggests that some internal reflections may be the reason for this deviation. Indeed, the generic model derived in Section 3.3.2 does not distinguish whether the superposition of light rays is an internal or external process. Thus, a very plausible reason of the jumping-edge would be the internal reflection between sensor and camera body or lens. In fact, this is a known issue termed light scattering, as introduced in Section 2.3.4.

Based on the illustration in Fig. 5.15, a specialization of the generic multipath-reflection model (3.57) can be established. Considering just one pixel row for now, the multipath phasor $z_r[i,j]$ may be written as

$$z_r[i,j] = w(\alpha_{ij})\rho_s\rho_b g \left(d_0 + \frac{2\Delta d}{\cos(\alpha_{ij})}\right) \mathrm{e}^{\jmath 2\pi f_{\mathrm{mod}}\left(t_{d_0}+\frac{2t_{\Delta d}}{\cos(\alpha_{ij})}\right)} \tag{5.7}$$

with, according to Fig. 5.15, the following parameters:

- i,j Pixel indices. The parameterization $z_r[i,j]$ refers to the influence of the j^{th} pixel on the i^{th} pixel.
- d_p Distance between neighboring pixels (constant).
- d_0 Distance to the target (constant). Relates to the according delay t_{d_0} via the speed of light.

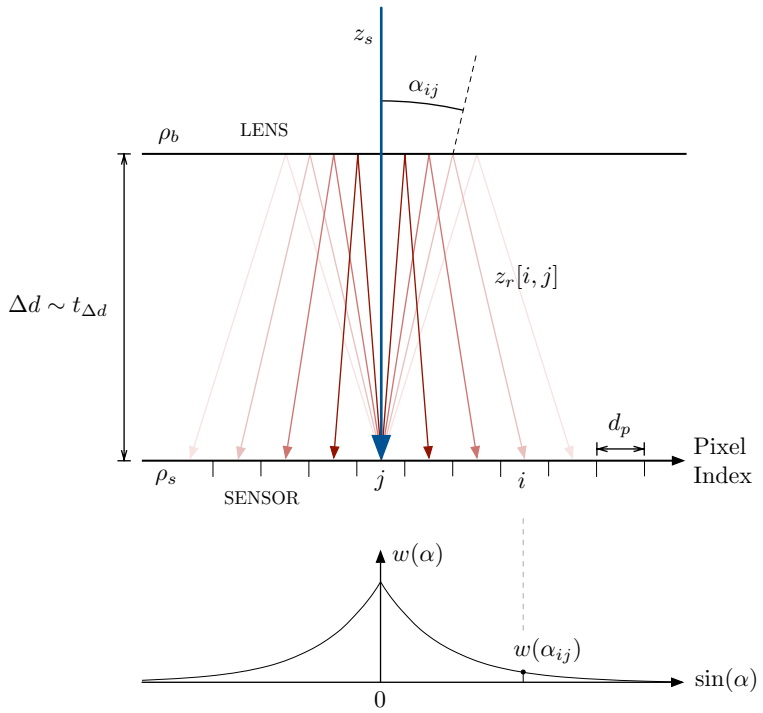


Figure 5.15: Illustration of the jumping-edge behavior modeled by the multipath-reflection framework. The symbols are explained in the text.

Δd Distance to opponent reflector (constant), with $t_{\Delta d}$ being the respective time delay.

α_{ij} Reflection angle from pixel j to pixel i. Naturally, for this one-dimensional case this angle computes to

$$\alpha_{ij} = \operatorname{atan}\left(\frac{|j-i|d_\mathrm{p}}{2\Delta d}\right) . \tag{5.8}$$

$w(\cdot)$ Weighting function that accounts for the amplitude decay with increasing reflection angle. In the current implementation, exponential decay was assumed as it delivered best results.

ρ_s Reflection coefficient of the sensor (constant) with values between zero and one.

ρ_b Reflection coefficient of the body / lens (constant) with values between zero and one.

$g(\cdot)$ Weighting function that accounts for the reciprocal proportionality between amplitude and distance (cf. (3.2)). Note that $2\Delta d/\cos(\alpha_{ij})$ accounts for the additional distance introduced by the reflection path (the same applies for the additional delay).

Finally, the observed phasor $z[i]$ is computed via

$$\begin{aligned} z[i] &= z_s[i] + \sum_j z_r[i,j] \\ &= g(d_0)\mathrm{e}^{\jmath 2\pi f_\mathrm{mod} t_{d_0}} + \sum_j w(\alpha_{ij})\rho_s\rho_b g\left(d_0 + \tfrac{2\Delta d}{\cos(\alpha_{ij})}\right)\mathrm{e}^{\jmath 2\pi f_\mathrm{mod}\left(t_{d_0}+\frac{2t_{\Delta d}}{\cos(\alpha_{ij})}\right)} . \end{aligned} \tag{5.9}$$

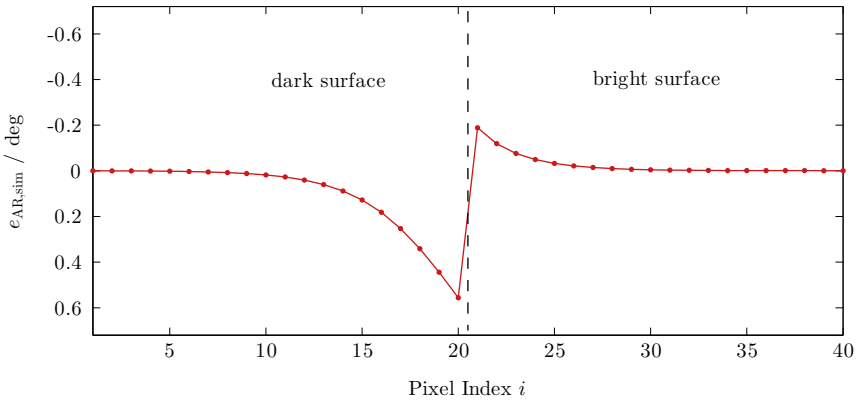


Figure 5.16: Simulation results of the presented jumping-edge model based on multipath reflections. A modulation frequency of 20 MHz was assumed, and the original target distance was set to $d_0 = 1.0\,\mathrm{m}$.

For large d_0, $g(\cdot)$ may be replaced by $|z_s[j]|$ as the additional term in the argument is small with respect to d_0. As well, it should be noted that this model transparently extends to the two-dimensional case.

In Fig. 5.16, an exemplary simulation of the jumping-edge behavior by means of this model is given. As far as possible, real data such as lateral pixel distance or distance between silicon and lens was used. As well, measurement results were applied, for example amplitude values. Therefore, the model should approximate real measurements (such as those given in Fig. 5.13) which it does indeed. Quantitatively speaking, the simulated error does not only accord to the measured one, also the characteristic that the error is larger at the dark side of the transition is surprisingly well covered by the model. This can be seen as another experimental proof of the multipath-reflection model. Note that the overall curve was offset by about 0.8°, which was removed for better comparability. This offset is not surprising, as all the reflections (which are also present in homogeneous areas) introduce a constant, positive bias to the observed depth.

5.3.3 First-Order Model Offset

In all previous experiments, no evidence was found that higher-order models provide any information that makes them superior to first-order estimates – not for varying target distance and integration time, nor for different target reflectivities. However, considerable offsets between estimates of different order can be expected when altering the modulation frequency. This is already indicated in Fig. 5.6 and Fig. 5.9 and should be elaborated more in detail, for which reason this subsection provides a comprehensive evaluation of the depth offset between first-order estimation and the true argmax of the SRF. This offset is defined as

$$\begin{aligned} e_{\mathrm{OS}}(f_{\mathrm{mod}}, \tau) &= \hat{d}_1 - \hat{d}_{M_{\max}} \\ &= \left(\operatorname{argmax}\{c_{M_{\max}}(\varphi)\} - \operatorname{argmax}\{c_1(\varphi)\}\right) \frac{\mathrm{c}}{4\pi f_{\mathrm{mod}}} \end{aligned} \tag{5.10}$$

where $\hat{d}_1$ and $\hat{d}_{M_{\max}}$ refer to first-order and maximum-order depth estimates, respectively. For the current experiments, the maximum order was set to $M_{\max} = 9$.

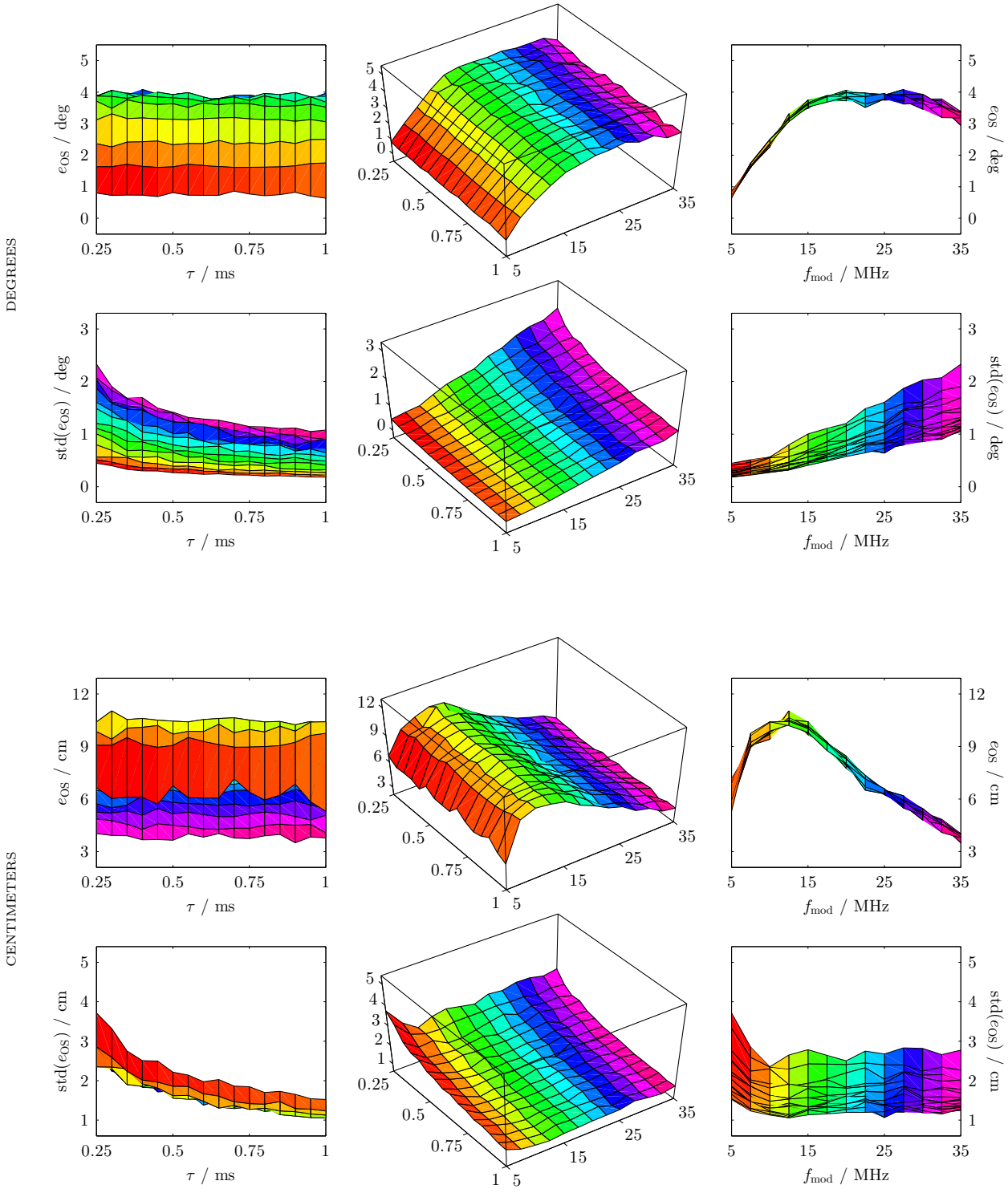


Figure 5.17: Offset in degrees (top) and centimeters (bottom) between first-order estimate and true argmax of the SRF over varying integration times and modulation frequencies.

[$f_{\text{mod}} = 5 \ldots 35$ MHz, $\tau = 250 \ldots 1000\,\mu$s, $N_E = 48$, $K = 32$]

In Fig. 5.17, the first-order offset is given over a range of $\tau = \{0.25, 0.30, \ldots, 1.00\}$ ms and $f_{\text{mod}} = \{5, 7.5, \ldots, 35\}$ MHz, thus for 16×13 measurements. As it was observed that this offset is consistent over different pixels, 11 bright pixels (with 32 realizations each) were merged together

for additional smoothness. Most intuitively delineated as a 3D-surface in the center column, the left and right columns provide the according projections. In the same manner, the standard deviation is provided as well which should give additional insights into the reliability of the data.

The most apparent characteristic of the figures is that the first-order model offset is indeed completely invariant to changes in integration time, which is most evidently proven by the respective upper right subfigures. Furthermore, The frequency-dependent offset is considerable. It amounts from 1° at 5 MHz to about 4° at 20 MHz. Due to the changing relationship between angular and metric units with varying frequency, the largest deviation in centimeters can be expected at about 12.5 MHz (10 cm).

As an important note, it should be remarked that the discussed data is still relative and should only serve as a motivation towards first-order estimation. In practice, for consistent depth sensing further information about the illumination unit will be needed, such as its behavior over temperature and modulation frequency. What has been shown only, is the fact that this information cannot be extracted from the SRF.

5.3.4 Conclusion

In the course of the steady-target experiment, several new insights into the nature of systematic errors were gained. Regarding the integration-time related error, it was shown in Section 5.3.1, that higher-order estimates are not superior to lower-order estimates. Nevertheless, an intergration-time based shift of the entire SRF was observed. Due to the fact that this shift is already present in the optical signal at time of arrival at the sensor, the illumination unit is most likely the reason for this deviation. Actually, this is very plausible as longer integration times cause the illumination unit to warm-up, which is known to introduce signal delays.

The major part of this chapter was dedicated to the amplitude-related error in Section 5.3.2. It was shown that actually no direct dependency between SRF amplitude and depth deviation exists. In fact, darker areas do not per se appear at different distances as bright areas, which is shown in Fig. 5.13 (top). The only characteristic deviation observed over all measurements was the discontinuity of depth values at intensity changes. However, this effect could not be linked to absolute values of the SRF either but modeled by multipath reflections. Thus, due to the fact that a thorough proof of multipath reflections being responsible for amplitude-related deviations was established, *calibrating* this issue with respect to amplitude values is obsolete. Instead, algorithmic models that rely on the models presented may be developed.

The analyses of the steady-target experiment were concluded in Section 5.3.3, where an extensive evaluation of the offset between first-order estimation and true argmax is provided. The given plots evidently show that the only parameter that affects this offset is the modulation frequency, which is no surprise. Therefore, for consistent depth sensing via first-order estimation, this frequency-dependent offset must be known. Furthermore, as the illumination unit itself can be expected to introduce frequency-dependent delays into the optical signal (thus shifting the argmax), this absolute deviation must be determined in advance as well.

— CHAPTER SIX —

Proposals for Improved Distance Estimation

This chapter presents two concepts that were developed towards distance estimation that is superior to state-of-the-art solutions. Estimation models that operate with different sampling strategies are deduced and evaluated in Section 6.1. Additionally, a simple enhancement technique termed *adaptive super-resolution* (ASR) algorithm, which is applicable to any sampling rate and effectively reduces aliasing distortions is derived and assessed in Section 6.2.

6.1 Alternative Sampling Strategies for First-Order Models

The results from Chapter 5, in particular those from the moving-target experiment in Section 5.2, have shown that in general it is a valid simplification to reduce the general distance estimation problem to the phase estimation of the fundamental wave. Indeed, this is what state-of-the-art ToF imagers actually do, however, the delivered depth data suffers from artifacts due to violations of the sampling theorem. The resulting Wiggling Error was examined in Section 5.2.2 and shown to deviate by a maximum of several degrees from the true phase of the fundamental wave (cf. Tab. 5.1). Theoretical findings (3.43) have furthermore shown that 4-fold sampling causes aliasing of all odd harmonics onto the first frequency bin of the DFT, which is exactly the bin evaluated by the 4-phase algorithm. In combination with the results from Fourier analysis in Section 5.1.2, i.e., that the SRF spreads considerable information into its odd harmonics, from a signal-processing viewpoint the 4-fold sampling can be regarded as the worst-case sampling strategy applicable to signals of this characteristic – although admittedly convenient in terms of hardware design due to its symmetry. Nevertheless, this motivates the search for alternative sampling strategies that are superior to 4-fold sampling.

6.1.1 Generic Derivation

Given an N-fold sampled SRF by $c[n], n \in \{0, \dots, N-1\}$ where sampling point n refers to a reference phase shift of $2\pi n/N$, the traditional 4-phase algorithm (3.41) naturally extends via

the DFT to (3.16)

$$\begin{aligned}\hat{\varphi}_d &= \arg(\hat{C}[1]) \\ &= \arg\left(\frac{1}{N}\sum_{n=0}^{N-1} c[n]\mathrm{e}^{-\jmath\frac{2\pi}{N}n}\right) \\ &= \operatorname{atan2}\left(-\sum_{n=0}^{N-1} c[n]\sin(\tfrac{2\pi}{N}n)\ ,\ \sum_{n=0}^{N-1} c[n]\cos(\tfrac{2\pi}{N}n)\right)\end{aligned} \tag{6.1}$$

and the formula of the respective amplitude (3.42) to

$$\begin{aligned}\hat{A}_1 &= 2|\hat{C}[1]| \\ &= 2\left|\frac{1}{N}\sum_{n=0}^{N-1} c[n]\mathrm{e}^{-\jmath\frac{2\pi}{N}n}\right| \\ &= \frac{2}{N}\sqrt{\left(\sum_{n=0}^{N-1} c[n]\cos(\tfrac{2\pi}{N}n)\right)^2 + \left(\sum_{n=0}^{N-1} c[n]\sin(\tfrac{2\pi}{N}n)\right)^2}\ .\end{aligned} \tag{6.2}$$

It should be remarked that these equations define a certain class of estimators which is parameterized *only* by the sampling rate N. As a consequence, all members of this class (such as is the 4-phase algorithm and any others extracted below) can only differ by some performance measure that also depends on N. In other words, all performance differences are either an issue of aliasing (systematic) or subject to noise (stochastic), which both depend on N. *All other properties are equal*, such as the robustness against DC offset or linear gain applied to the SRF. However, nonlinear distortion causes spectral artifacts, and in turn, these will affect some class members more than others due to aliasing.

Concerning notation, at this stage it may be confusing to which sampling rate $\hat{\varphi}_d$ and $\hat{A}_1$ actually refer to. However, this should be clear from the context in most cases, so in order maintain a concise notation, reference to N is only provided if found necessary.

6.1.2 Examples

Although the formulas given in (6.1) and (6.2) apply for the general case, in real-time systems the effective number of samples will be strictly limited. Thus, as exemplary sampling rates that should have a chance for realization in practice, 3-fold and 5-fold sampling are discussed in this subsection.

3 Samples

Taking three samples at equidistant steps refers to capturing $c[n]$ at reference phase shifts $\varphi_n = \{0, 2\pi/3, 4\pi/3\}$. From (6.1) and (6.2), the estimation formulas are defined as follows:

$$\begin{aligned}\hat{\varphi}_d &= \operatorname{atan2}\left(-(c[1]\sin(\tfrac{2\pi}{3}) + c[2]\sin(\tfrac{4\pi}{3}))\ ,\ c[0] + c[1]\cos(\tfrac{2\pi}{3}) + c[2]\cos(\tfrac{4\pi}{3})\right)\\ &= \operatorname{atan2}\left(-(c[1]\tfrac{\sqrt{3}}{2} + c[2](-\tfrac{\sqrt{3}}{2}))\ ,\ c[0] + c[1](-\tfrac{1}{2}) + c[2](-\tfrac{1}{2})\right)\\ &= \operatorname{atan2}\left(\sqrt{3}(c[2]-c[1])\ ,\ 2c[0] - c[1] - c[2]\right)\end{aligned} \tag{6.3}$$

and

$$\begin{aligned}\hat{A}_1 &= \frac{2}{3}\sqrt{\left(c[0] - \tfrac{1}{2}(c[1]+c[2])\right)^2 + \tfrac{3}{4}\left(c[2]-c[1]\right)^2}\\ &= \frac{2}{3}\sqrt{c[0]^2 + c[1]^2 + c[2]^2 - c[0]c[1] - c[0]c[2] - c[1]c[2]}\,.\end{aligned} \tag{6.4}$$

Regarding aliasing, 3-fold sampling causes the following harmonics to be mapped back onto the first frequency bin (3.18) (3.20)

$$\begin{aligned}\hat{C}[1] &= C[1] + \sum_{m=1}^{\infty}\left(C[3m+1] + C^*[3m-1]\right)\\ &= C[1] + C^*[2] + C[4] + C^*[5] + C[7] + \cdots\end{aligned} \tag{6.5}$$

where it is evident that the third harmonic does *not* appear on the first frequency bin but instead, the second and fourth harmonics do. Generally speaking, for aliasing frequencies below the seventh order, the trade-off between 3-fold sampling and 4-fold sampling is second plus fourth harmonic versus third harmonic. As is known from previous analyses and the discussions about the spectrum of the SRF (cf. Section 5.1.2), this trade-off strongly indicates significant advantages of 3-fold sampling over 4-fold sampling in terms of aliasing artifacts. In Section 6.1.3, this advantage will be confirmed by according simulations. Furthermore, tests with real data are provided in Section 6.2.4.

5 Samples

Similar to above, taking 5 samples at equidistant steps of $\varphi_n = \{0, 2\pi/5, 4\pi/5, 6\pi/5, 8\pi/5\}$ leads to the following estimation formulas:

$$\begin{aligned}\hat{\varphi}_d &= \operatorname{atan2}\big(-(c[1]\sin(\tfrac{2\pi}{5}) + c[2]\sin(\tfrac{4\pi}{5}) + c[3]\sin(\tfrac{6\pi}{5}) + c[4]\sin(\tfrac{8\pi}{5}))\ ,\\ &\qquad c[0] + c[1]\cos(\tfrac{2\pi}{5}) + c[2]\cos(\tfrac{4\pi}{5}) + c[3]\cos(\tfrac{6\pi}{5}) + c[4]\cos(\tfrac{8\pi}{5})\big)\\ &= \operatorname{atan2}\big((c[4]-c[1])\sin(\tfrac{2\pi}{5}) + (c[3]-c[2])\sin(\tfrac{4\pi}{5})\ ,\\ &\qquad c[0] + (c[4]+c[1])\cos(\tfrac{2\pi}{5}) + (c[3]+c[2])\cos(\tfrac{4\pi}{5})\big)\end{aligned} \tag{6.6}$$

and

$$\begin{aligned}\hat{A}_1 = \frac{2}{5}\sqrt{\left(c[0] + (c[4]+c[1])\cos(\tfrac{2\pi}{5}) + (c[3]+c[2])\cos(\tfrac{4\pi}{5})\right)^2 +}\\ \overline{+\left((c[4]-c[1])\sin(\tfrac{2\pi}{5}) + (c[3]-c[2])\sin(\tfrac{4\pi}{5})\right)^2}\,.\end{aligned} \tag{6.7}$$

The obtained expressions already seem to be challenging enough for implementation in real-time frameworks although the constant trigonometric coefficients can be computed in advance and stored in a LUT. Anyway, going for 6 samples or even more seems not reasonable in this context.

The harmonics reflected back to the first frequency bin are in case of 5-fold sampling

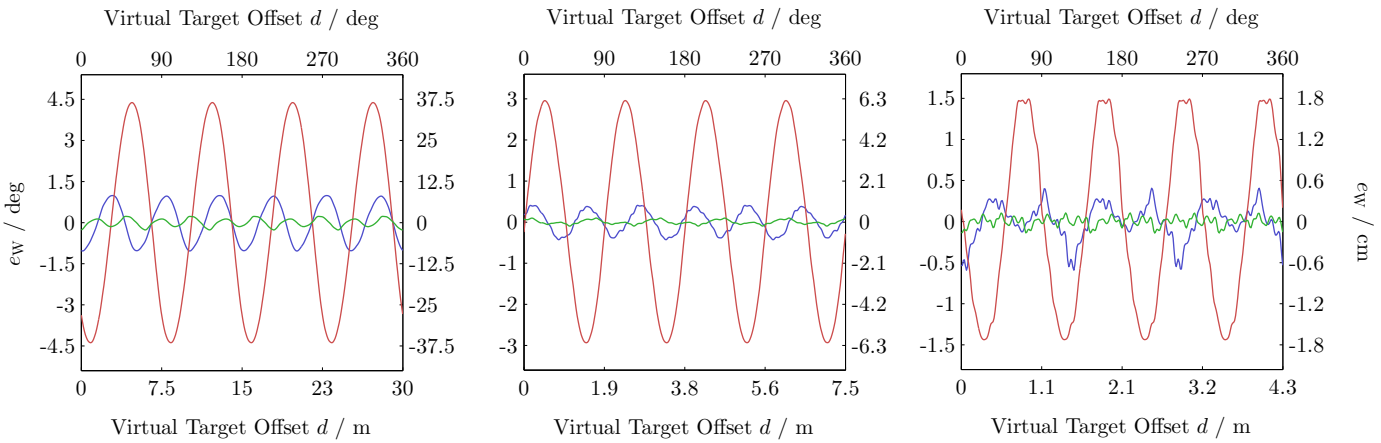
$$\begin{aligned}\hat{C}[1] &= C[1] + \sum_{m=1}^{\infty} \left(C[5m+1] + C^*[5m-1]\right) \\ &= C[1] + C^*[4] + C[6] + C^*[9] + C[11] + \ \cdots . \end{aligned} \tag{6.8}$$

It is no surprise that 5-fold sampling provides a better spectral resolution than 3 or 4 samples, but with knowledge about the SRF spectrum, this sampling rate appears to be particularly advantageous: The lowermost harmonic mapped back which is considered as somehow problematic is the ninth. In other words, compared original 4-fold sampling, this one sample more makes estimation robust against the third, fifth and seventh harmonic. Consequently, significant improvements can be expected for this sampling strategy as well, as further discussed in the following section.

6.1.3 Evaluation

In Fig. 6.1, the simulated Wiggling Error (cf. definition in (3.48)) for the sampling rates discussed above is depicted, respectively at modulation frequencies of 5, 20 and 35 MHz. The improvement to be expected from different sampling rates is significant.

From actual numbers of the respective error given in Tab. 6.1, it is evident that not only 3-fold sampling decreases the error by a factor of 2.5 up to about 7, but 5-fold sampling makes the Wiggling Error almost vanish at modulation frequencies around 20 MHz. Moreover, it should be mentioned that due to the higher amount of oscillations over the whole distance range, *on average* the error will decrease even more.



(a) 5 MHz (b) 20 MHz (c) 35 MHz

Figure 6.1: Evaluation of the proposed sampling strategies by means of a Wiggling-Error simulation, given for exemplary modulation frequencies.
[$f_{\mathrm{mod}} = 5/20/35$ MHz, $\tau = 1500\,\mu$s, $N_E = 120$, $K = 32$]

Sampling Rate	**Wiggling Error**								
	5 MHz			20 MHz			35 MHz		
	deg	cm	factor	deg	cm	factor	deg	cm	factor
3-fold	1.035	8.627	×**4.23**	0.430	0.896	×**6.88**	0.592	0.705	×**2.52**
4-fold	4.384	36.530	×1.00	2.958	6.163	×1.00	1.493	1.777	×1.00
5-fold	0.267	2.225	×**16.42**	0.100	0.209	×**29.54**	0.140	0.166	×**10.68**

Table 6.1: Wiggling Error (max. amplitude) of different sampling rates and modulation frequencies. The improvement factor is given with respect to 4-fold sampling at the same modulation frequency.

6.2 Adaptive Super-Resolution (ASR) Algorithm

The alternative sampling strategies presented in Section 6.1 have already shown significant improvements over the traditional 4-phase algorithm. However, although 3-fold sampling astonishingly well reduces the Wiggling Error by taking one sample *less*, still systematic deviations are present for all modulation frequencies. On the other hand, 5-fold sampling indeed almost completely eliminates the Wiggling Error, but at slightly higher cost compared to 4-fold sampling.

The question is now whether other ways can be found that increase the spectral resolution without simply taking more samples. This leads to the idea that it should be possible to combine consecutive measurements in order to achieve the desired resolution increase. Similar to above, in Section 6.2.1 the ASR algorithm is deduced in a very general fashion, whereas in Section 6.2.3 concrete application examples are presented. Practical aspects of the algorithm are discussed in Section 6.2.2 and performance assessment is carried out in Section 6.2.4.

6.2.1 Derivation

The basic idea of the ASR algorithm is to apply a constant offset $\Delta\varphi$ to the reference phase shifts from frame to frame. Indicating this offset by the according subscript (cf. (3.45)), the DFT coefficients of two succeeding frames may be written as (3.16) (3.18)

$$\hat{C}[k] = \sum_{n=0}^{N-1} c[n]\mathrm{e}^{-\jmath\frac{2\pi}{N}kn} = C[k] + \sum_{m=1}^{\infty}(C[k+mN] + C[k-mN]) \tag{6.9}$$

$$\hat{C}_{\Delta\varphi}[k] = \sum_{n=0}^{N-1} c_{\Delta\varphi}[n]\mathrm{e}^{-\jmath\frac{2\pi}{N}kn} = C_{\Delta\varphi}[k] + \sum_{m=1}^{\infty}(C_{\Delta\varphi}[k+mN] + C_{\Delta\varphi}[k-mN])\,. \tag{6.10}$$

The phase offset applied to the second frame can now be written explicitly as

$$\hat{C}_{\Delta\varphi}[k] = C[k]\mathrm{e}^{-\jmath\Delta\varphi k} + \sum_{m=1}^{\infty}\left(C[k+mN]\mathrm{e}^{-\jmath\Delta\varphi(k+Nm)} + C[k-mN]\mathrm{e}^{-\jmath\Delta\varphi(k-Nm)}\right)\,. \tag{6.11}$$

In order to compensate the complex rotation applied to $C[k]$, a compensation phase shift of $\Delta\varphi k$ to the whole expression can be introduced, which results in

$$\hat{C}_{\Delta\varphi}[k]\mathrm{e}^{\jmath\Delta\varphi k} = C[k] + \sum_{m=1}^{\infty}\left(C[k+mN]\mathrm{e}^{-\jmath\Delta\varphi Nm} + C[k-mN]\mathrm{e}^{+\jmath\Delta\varphi Nm}\right) . \tag{6.12}$$

The goal is now to assemble the two succeeding frames and choose $\Delta\varphi$ in a manner such that some aliasing components possibly cancel out. Combining the frames in an additive fashion yields

$$\begin{aligned}\frac{1}{2}&\left(\hat{C}[k] + \hat{C}_{\Delta\varphi}[k]\mathrm{e}^{\jmath\Delta\varphi k}\right) = \\ &= C[k] + \frac{1}{2}\sum_{m=1}^{\infty}\left(C[k+mN]\left(1+\mathrm{e}^{-\jmath\Delta\varphi Nm}\right) + C[k-mN]\left(1+\mathrm{e}^{+\jmath\Delta\varphi Nm}\right)\right)\end{aligned} \tag{6.13}$$

for which it is evident that choosing $\Delta\varphi = \pi/N$ leads to

$$1+\mathrm{e}^{\pm\jmath\pi m} = 1+(-1)^m = \begin{cases} 2 & m \text{ even} \\ 0 & m \text{ odd.}\end{cases} \tag{6.14}$$

This causes every second term of the sum in (6.13) to cancel out, i.e.,

$$\frac{1}{2}\left(\hat{C}[k] + \hat{C}_{\pi/N}[k]\mathrm{e}^{\jmath k\pi/N}\right) = C[k] + \sum_{m=1}^{\infty}\left(C[k+2mN] + C[k-2mN]\right) . \tag{6.15}$$

This is in fact equal to $2N$-fold sampling, motivating the term *super-resolution*. The phase of the fundamental wave computes then to

$$\begin{aligned}\hat{\varphi}_d &= \arg\left(\hat{C}[1] + \hat{C}_{\pi/N}[1]\mathrm{e}^{\jmath\pi/N}\right) \\ &= \arg\left(\sum_{n=0}^{N-1} c[n]\mathrm{e}^{-\jmath\frac{2\pi}{N}n} + \sum_{n=0}^{N-1} c_{\pi/N}[n]\mathrm{e}^{-\jmath\frac{2\pi}{N}n}\mathrm{e}^{\jmath\pi/N}\right)\end{aligned} \tag{6.16}$$

with the following observations being worth discussed:

- The chosen $\Delta\varphi = \pi/N$ actually refers to the *half* of the distance between two sampling points (in time domain traditionally referred to as sampling period). For this reason, it is not surprising that assembling two frames whose sampling points are offset by half the sampling period doubles the spectral resolution. With concrete numbers, this means that 3-fold sampling can be turned into 6-fold sampling, 4-fold sampling into 8-fold sampling, etc.

- Which frame of the two is offset by π/N has actually no effect, as long as the respective compensation shift is applied correctly. Therefore, the ASR algorithm can be implemented such that new depth information is provided *each* single frame, using the current frame and the past one. Furthermore, it should be remarked that this compensation shift $\mathrm{e}^{\jmath\pi/N}$ is constant over all measurements but only depends on the number of samples N.

- Referring to signal processing theory, the underlying idea of the presented derivation is closely related to the principle of *decimation in time* which allows a more efficient computation of the DFT via the *fast Fourier transform* (FFT) [26, p.749].

- The fact that the two frames are assembled in an additive fashion allows easy enabling or disabling the algorithm. In other words, if the second term in (6.16) is simply canceled, a single N-fold sampled frame remains that per se provides fully qualified depth information. For this reason, the algorithm is endowed with the attribute *adaptive*. This can be very useful in cases of rapidly changing scenes in order to avoid motion artifacts due to inconsistent frames. Of course, disabling the algorithm allows Wiggling artifacts to reoccur again, so a reasonable trade-off must be found. It should be emphasized that the ASR algorithm is pixel-based, so any adaptation can be pixel-based as well.

- For odd sampling rates, (6.16) can be furthermore simplified. Instead of choosing $\Delta\varphi$ half the sampling period, it can be set to $\Delta\varphi = \pi$. This refers to a sampling point that is not covered by odd sampling rates but still offsets the sampling points between the frames by half the sampling period, which is the fundamental requirement for the algorithm to work correctly. Thus, applying an offset of π changes the complex compensation shift to the constant (-1), thus

$$\begin{aligned}\hat{\varphi}_d &= \arg\left(\hat{C}[1] - \hat{C}_\pi[1]\right) \\ &= \arg\left(\sum_{n=0}^{N-1} c[n]\mathrm{e}^{-\jmath\frac{2\pi}{N}n} - \sum_{n=0}^{N-1} c_\pi[n]\mathrm{e}^{-\jmath\frac{2\pi}{N}n}\right) .\end{aligned} \tag{6.17}$$

- It can be shown that if the ASR algorithm is applied to odd sampling rates, its major effect (apart from increasing the SNR due to additional sampling points) is the removal of even aliasing harmonics. Consequently, if applied to even sampling rates, only odd aliasing harmonics will vanish. This becomes clear when observing the indices of the components that will cancel out (cf. (6.15)), i.e., $(k + (2m-1)N)$ and $(k - (2m-1)N)$. For $k = 1$ and N odd, these expressions will be even because $(2m-1)N$ (a product of two odd numbers) is odd again, and the sum of two odd numbers is even. In the same manner it can be shown that for $k = 1$ and N even, the respective indices will become odd.

- Regarding amplitude values, they can still be computed via the ordinary formula (6.2) for the respective sampling rate.

6.2.2 Practical Aspects

A basic block diagram how the algorithm could be implemented is given in Fig. 6.2. The additional superscripts $c^{(m)}[n]$ and $\hat{C}^{(m)}[1]$ refer to the frame number m. For simplicity, the signal path between PROCESSING and the argument computation is kept complex-valued but must be split up into two parallel paths that carry real- and imaginary parts. As discussed above, for odd sampling rates the compensation shift reduces to ± 1 if the phase offset applied to φ_n is set to π instead of π/N.

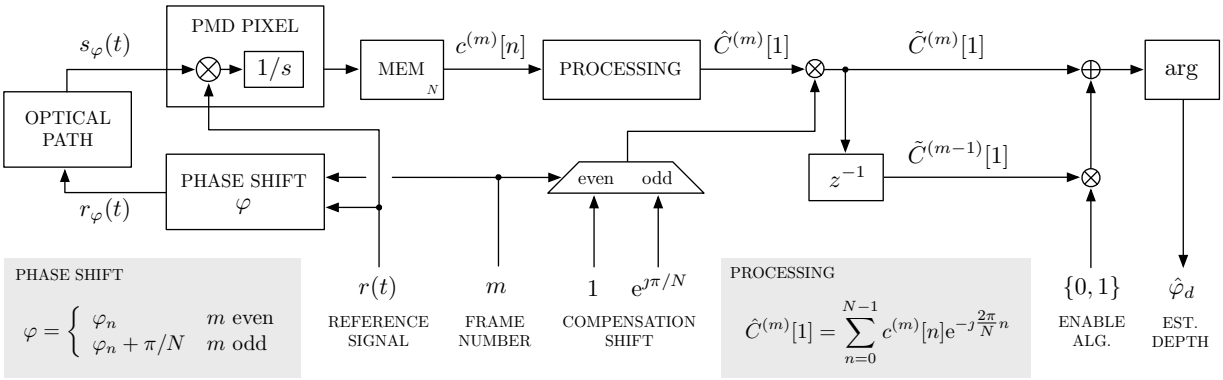


Figure 6.2: Simplified structure of a possible implementation of the ASR algorithm.

Whether the algorithm should be enabled or disabled will directly relate to some FoM that detects changes in the scene. Given $\hat{C}^{(m)}[1]$ and $\hat{C}^{(m-1)}[1]$, a very simple threshold criterion could be based on the changes of real- and imaginary part which does not even involve any additional computation. The phase estimate of the m^{th} frame computes then to

$$\hat{\varphi}_d[m] = \arg\left(\tilde{C}^{(m)}[1] + \tilde{C}^{(m-1)}[1]\right) . \tag{6.18}$$

The additional hardware complexity required by the algorithm can be directly read out from Fig. 6.2. Regarding memory, only one complex value that represents the past frame must be stored. Reference phase shifts of π/N must be implemented. As well, some additional logic is required for selecting the reference phases and the compensation shift according to m. Finally, one additional complex multiplication and addition per pixel are necessary. Generally, it should be noted that the complexity of the algorithm does *not* depend on N.

6.2.3 Examples

Basically, the ASR algorithm can be implemented regardless of the sampling rate as it only operates on the processed (complex-valued) depth estimates. However, it will be most probably attached to imagers that work with few samples, such as 3, 4, or 5 samples. Effectively, the algorithm turns these rates into 6-fold, 8-fold, and 10-fold sampling respectively, so the expected improvements are shortly discussed below.

3 Samples ASR

Given the fact that 3-fold sampling with applied ASR is equivalent to 6-fold sampling, the impact of aliasing can be formulated as

$$\begin{aligned} \hat{C}[1] &= C[1] + \sum_{m=1}^{\infty} \left(C[6m+1] + C^*[6m-1]\right) \\ &= C[1] + C^*[5] + C[7] + C^*[11] + C[13] + \cdots . \end{aligned} \tag{6.19}$$

Compared to 3-fold sampling (6.5), the only difference is that now the first frequency bin is robust against even harmonics of the SRF which was already discussed above. Thus, for the general case this is only a minor improvement as even harmonics are supposed to be very low. However, as even harmonics may occur as a consequence of nonlinear distortion, 3-fold sampling with ASR will be more robust in this aspect.

4 Samples ASR

Huge improvements can be expected in this case, because

$$\begin{aligned}\hat{C}[1] &= C[1] + \sum_{m=1}^{\infty} (C[8m+1] + C^*[8m-1]) \\ &= C[1] + C^*[7] + C[9] + C^*[15] + C[17] + \cdots \end{aligned} \tag{6.20}$$

shows that the two most problematic harmonics, the third and fifth, are no longer mapped back onto the first frequency bin.

5 Samples ASR

Finally, applied to 5-fold sampled data, the ASR reduces the aliasing impact to

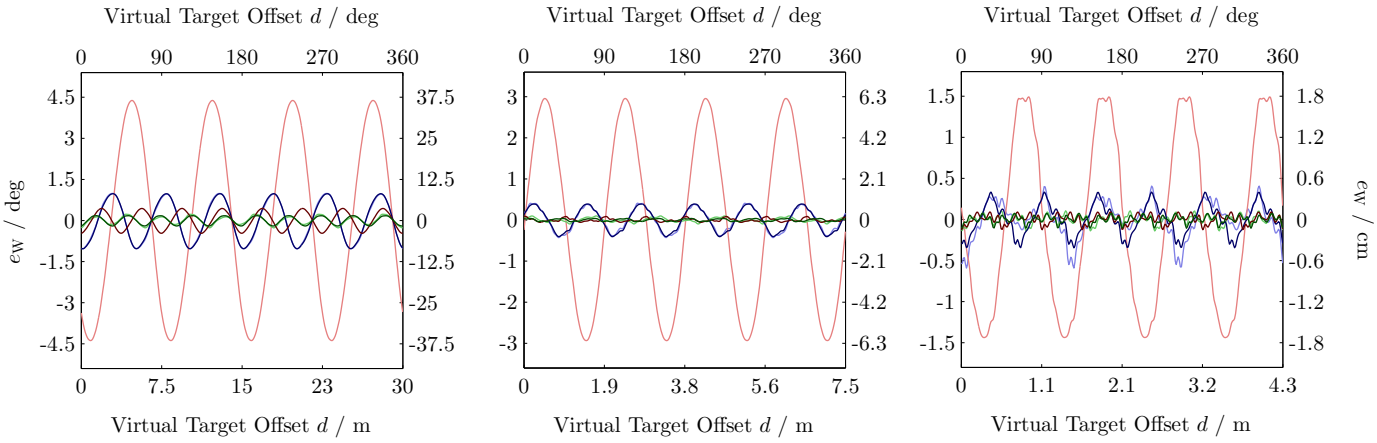
$$\begin{aligned}\hat{C}[1] &= C[1] + \sum_{m=1}^{\infty} (C[10m+1] + C^*[10m-1]) \\ &= C[1] + C^*[9] + C[11] + C^*[19] + C[21] + \cdots . \end{aligned} \tag{6.21}$$

At a first glance this seems very promising as well, but for the same reasons stated in the case of 3-fold sampling, the ASR algorithm only removes even harmonics from aliasing distortion. This may lead to more robust behavior in case of nonlinearities, but will not show much improvement in the general case.

6.2.4 Evaluation

In Fig. 6.3, the results of a performance simulation are depicted. Compared to the alternative sampling rates discussed and evaluated in Section 6.1, the most evident characteristic is that the algorithm improves the performance of 4-fold sampling dramatically, whereas for 3-fold and 5-fold sampling only a small decrease of the maximum Wiggling Error can be seen. These observations are supported by the fact that the ASR algorithm removes even aliasing components when applied to odd sampling rates (and vice versa). The improvement for odd sampling rates may stem from the additional removal of artifacts due to even harmonics or low SNR.

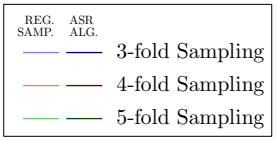
As an extension of Tab. 6.1, in Tab. 6.2 actual numbers are provided that reflect the previous observations. In combination with the given signal structure, the ASR algorithm reveals its main strength when applied to even sampling rates, in particular to 4-fold sampling. Comparing the respective numbers, an improvement factor between about 10 and 35 can be expected. In other words, above 20 MHz the Wiggling Error can be seen as eliminated.



(a) 5 MHz (b) 20 MHz (c) 35 MHz

Figure 6.3: Wiggling Error simulation of the ASR algorithm applied to the proposed sampling rates. Note that for 3-fold and 5-fold sampling the curves overlap, especially at low modulation frequencies where almost no improvement can be expected.

$[f_{\mathrm{mod}} = 5/20/35\ \mathrm{MHz}, \tau = 1500\,\mu\mathrm{s},\ N_E = 120,\ K = 32]$



The fact that the algorithm works best around 20 MHz but not at 35 MHz seems surprising at a first glance, but the reason becomes clear when considering the changing spectral composition of the SRF with increasing modulation frequencies. The overall performance of the algorithm can be related to the ratio of the *total harmonic distortion* (THD) caused by aliasing, before and after its application. Denoted by $\mathrm{THD}_{\mathrm{pre}}$ (pre-application) and $\mathrm{THD}_{\mathrm{post}}$ (post-application), the ratio

$$\frac{\mathrm{THD}_{\mathrm{post}}}{\mathrm{THD}_{\mathrm{pre}}} = \frac{\sum_{m=1}^{\infty} \left(|C[1+2mN]|^2 + |C[1-2mN]|^2\right)}{\sum_{m=1}^{\infty} \left(|C[1+mN]|^2 + |C[1-mN]|^2\right)} \tag{6.22}$$

approaches its minimum if the algorithm works best, which is zero in the ideal case. For example,

Sampling Rate	**Wiggling Error**								
	5 MHz			20 MHz			35 MHz		
	deg	cm	factor	deg	cm	factor	deg	cm	factor
3-fold, regular	1.035	8.627	×1.00	0.430	0.896	×1.00	0.592	0.705	×1.00
4-fold, regular	4.384	36.530	×1.00	2.958	6.163	×1.00	1.493	1.777	×1.00
5-fold, regular	0.267	2.225	×1.00	0.100	0.209	×1.00	0.140	0.166	×1.00
3-fold, ASR alg.	1.030	8.579	×**1.01**	0.406	0.946	×**1.06**	0.347	0.411	×**1.71**
4-fold, ASR alg.	0.467	3.889	×**9.39**	0.087	0.172	×**34.12**	0.123	0.150	×**11.87**
5-fold, ASR alg.	0.199	1.658	×**1.34**	0.041	0.093	×**2.45**	0.107	0.127	×**1.31**

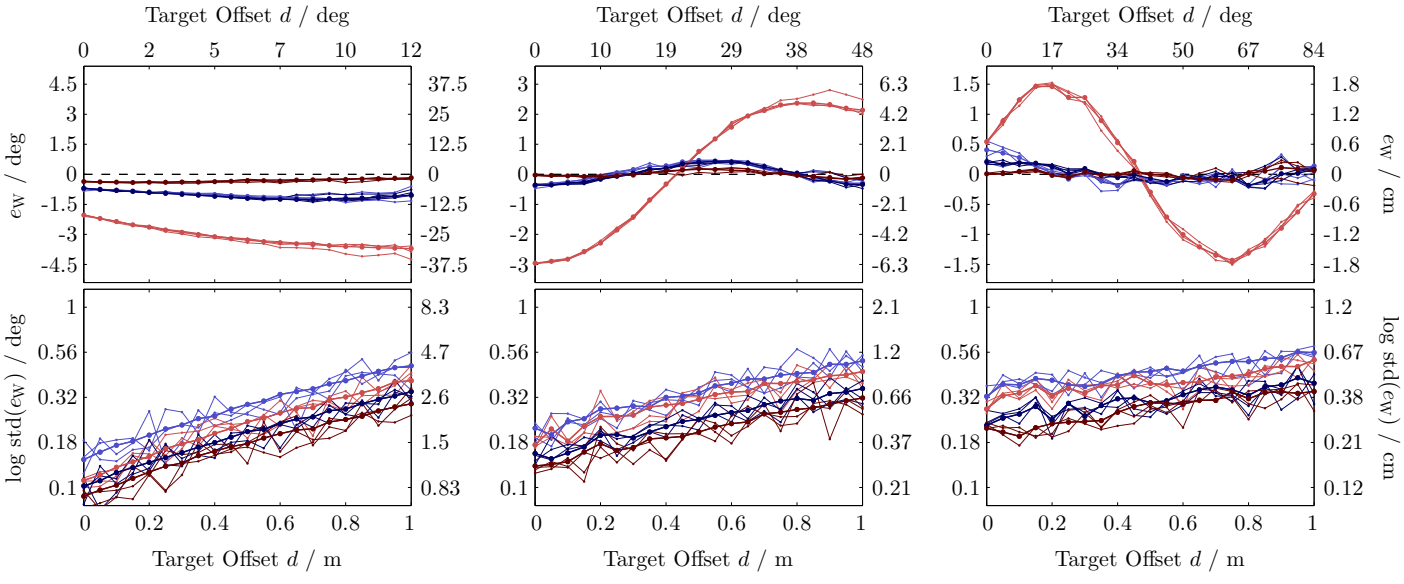
Table 6.2: Wiggling Error (max. amplitude) and respective improvement of the ASR algorithm over regular sampling, evaluated for different sampling rates and modulation frequencies. The given factors directly represent the improvement of the algorithm over regular sampling.

at 4-fold sampling and medium modulation frequencies significant impact due to aliasing still occurs, mostly from third and fifth harmonics. However, the algorithm is capable of removing the impact of these harmonics almost completely, so $\mathrm{THD}_{\mathrm{post}}/\mathrm{THD}_{\mathrm{pre}}$ is very low. For higher modulation frequencies the low-pass characteristic of the illumination unit comes into play and $\mathrm{THD}_{\mathrm{pre}}$ decreases – faster than $\mathrm{THD}_{\mathrm{post}}$ because most of the power decrease occurs in the lower harmonics (i.e., third, fifth) which does not affect $\mathrm{THD}_{\mathrm{post}}$. Consequently, $\mathrm{THD}_{\mathrm{post}}/\mathrm{THD}_{\mathrm{pre}}$ will rise. On the other hand, at lower modulation frequencies both measures increase as more power is spread into the signal harmonics. However, the relative increase of $\mathrm{THD}_{\mathrm{post}}$ is dominant because $\mathrm{THD}_{\mathrm{pre}}$ already contains the power of the third and fifth harmonic. In turn, $\mathrm{THD}_{\mathrm{post}}/\mathrm{THD}_{\mathrm{pre}}$ grows as well.

Real Data

Additionally to the simulated evaluation of the proposals presented in Section 6.1 and Section 6.2, evidence for the prognosticated behavior based on *real* data may be desired. For this reason, the moving-target experiment was repeated for 3-fold sampling and 4-fold sampling.

Fig. 6.4 impressively well proves the correctness of the proposed improvements. Again, thick lines represent the array mean (RoI: 11 × 11 pixel) and thin lines refer to individual pixels. From the standard deviation it can be seen that three samples slightly reduce the reliability of the data. In contrast to that, the ASR algorithm decreases the standard deviation as it relies on more data. Hence, also in terms of precision both approaches, 3-fold sampling and the ASR algorithm are completely reasonable.



(a) 5 MHz (b) 20 MHz (c) 35 MHz

Figure 6.4: Measured Wiggling Error of the moving-target experiment with according standard deviation, evaluated for the proposed methods. Recapitulate the definition of e_{W} from (5.3).
[$f_{\mathrm{mod}} = 5/20/35$ MHz, $\tau = 650/700/1000\,\mu$s, $N_E = 48$, $K = 32$]

— CHAPTER SEVEN —

Conclusions

This final chapter concludes this Thesis. A short summary of obtained results and achieved improvements is given in Section 7.1. The outlook presented in Section 7.2 should motivate further investigations of the topic, stimulated by observations that could not be satisfactorily explained or further elaborated.

7.1 Summary & Contribution of this Work

7.1.1 Fundamental Insights

Spectral Characteristic of the SRF

One of the first but probably most essential observation was the fact that the spectral composition of the SRF actually comprises *only* odd harmonics (Section 5.1.2). Whereas this could have been expected from a signal processing perspective, the consistency of this spectral characteristic over all measurements is astounding. Naturally, this has dramatic implications on the performance of distance estimation algorithms that rely on small samples sets (such as the 4-phase algorithm).

Temperature Influences

A known fact that was strongly supported by measurement results is the impact of temperature on depth sensing. Two kinds of depth drifts due to temperature changes were observed: Deviations as a result of component warm-up and changes of the ambient temperature (Section 5.1.2). In the latter case, a complete deterministic relationship between ambient temperature and depth drift was determined.

Sufficiency of First-Order Depth Estimation

It was thoroughly derived in Section 3.2.3 that distance estimation is equivalent to estimating the argmax of the SRF. However, all measurements throughout showed that just estimating the phase of the fundamental wave is sufficient for consistent depth sensing. Precisely speaking, first-order depth estimation exhibits a deterministic offset to the true argmax of the SRF which varies *only* with the modulation frequency, but not with target distance (cf. Section 5.2), nor with integration time (cf. Section 5.3). A summarizing analysis is provided in Section 5.3.3. Basically, this characteristic indicates a high level of linearity of the optical path, the PMD pixel,

and the ADC. Moreover, lower-order estimation performs better with respect to noise, with the standard deviation being lowest for first-order estimation.

Wiggling Error

The Wiggling Error, which is known to be a result of aliasing distortion, was analytically derived in Section 3.3.1, where a simulation framework was presented as well. In Section 5.2.2, it was shown that the simulated error perfectly well extrapolates measured data. Besides this established possibility to accurately prognosticate the Wiggling Error for any SRF sampled at any rate – with just one aliasing-free spectrum of the SRF needed in advance – it was argued that the determined formula could serve as a basis for algorithmic correction of this error.

Regarding the quantitative impact of the Wiggling Error, results known from literature were basically reproduced: With modulation frequencies between 5 and 35 MHz, the maximum Wiggling Error ranged from 4.41 ° (36.72 cm) to 1.46 ° (1.73 cm) (cf. Tab. 5.1 on page 44).

Integration-Time Related Error

Further understanding of the integration-time related error was developed. The hypothesis of the illumination unit being responsible for integration-time based deviations was confirmed. As argued in Section 5.3.1, the observed shift of the entire SRF indicates that no sensor-based nonlinearity can be responsible. Furthermore, the illumination unit is known to exhibit considerable increases in temperature if driven for a longer period – as in the case of longer integration times. In turn, the switching behavior of the electronics decelerates and the signal is delayed. Whereas this temperature-based delay is no fundamental surprise, the fact that the shape of the SRF is perfectly well maintained actually is. As a result, first-order depth estimation remains consistent. This is in particular evident in Section 5.3.3, where it can be seen that the first-order model offset remains constant with respect to variations of the integration time.

As a consequence, the logical next step towards algorithmic compensation of integration-time related errors is to investigate the correlation between integration time, illumination-unit heat-up and resulting signal delays.

Amplitude-Related Error & Multipath Reflections

One major outcome of this Thesis is the finding that the amplitude-related error per se *does not exist*. This may surprise at a first glance, but it was thoroughly shown in Section 5.3.2 that any deviations related to changes in signal amplitude can be fully related to multipath reflections. It is possible that these reflections occur inside the camera (referred to as *light scattering* in literature) or are a result of the scene composition.

In this Thesis, a mathematical framework of multipath reflections was established (Section 3.3.2) and proven to model real data (Section 5.3.2). The established relation to amplitude-based deviations has the game-changing implication that the amplitude-related error *cannot be calibrated*. Even without the underlying theory of multipath reflections as a root cause, measured data (such

as depicted in Fig. 5.12) do not show any correlation between actual SRF amplitude and depth deviation at all.

The observed jumping-edge behavior was experimentally proven in Section 5.3.2 to be a result of internal reflections since it was perfectly well simulated by the multipath-reflection framework. Due to the fact that the characteristic of internal reflections is uncoupled from the scene, this should allow for generic compensation in an algorithmic manner.

7.1.2 Reduction of Systematic Depth Deviations

Two approaches were discussed that effectively reduce the Wiggling Error, which is considered as the systematic error with highest impact.

Improved Sampling Strategies

In Section 6.1, alternative methods of sampling the SRF were proposed and evaluated. The most promising approaches, 3-fold sampling and 5-fold sampling, were discussed.

3 Samples. Due to the fact that 4-fold sampling is in particular suboptimal if applied to signals that contain only odd harmonics (such as the SRF), taking 3 equidistant samples turned out to outperform the traditional 4-phase algorithm by a factor between 2.5 (at a modulation frequency of 35 MHz) up to about 7 (at 20 MHz). In concrete numbers, the maximum depth deviation resulting from aliasing distortions was reduced from 6.16 cm to 0.90 cm at 20 MHz (cf. Tab. 6.1). It should be emphasized that this improvement was achieved by taking one sample *less*, which has additional advantages in terms of memory, frame rates, etc. Furthermore, the resulting depth estimation formula (6.3) is still very simple to compute, and tests with practical data have not shown any significant increase of the standard deviation (cf. Fig. 6.4).

5 Samples. As expected, taking one sample more than the traditional method as well resulted in a dramatic decrease of the Wiggling Error between the factor 10 (at 35 MHz) and even 30 (at 20 MHz). At 20 MHz, the maximum error effectively reduced from 6.16 cm to 0.21 cm (cf. Tab. 6.1). However, the estimation formula turned out to be slightly more complicated, involving additional multiplications. As well, more memory for the additional sample is needed, and a frame rate decrease must be accepted if this method should be implemented.

Adaptive Super-Resolution (ASR) Algorithm

Independent of the deployed sampling strategy, an algorithm was presented in Section 6.2 that effectively doubles the spectral resolution by combining two consecutive frames. Additional system complexity is low, just needing one complex number to be stored per pixel, and one complex multiplication and addition per pixel. As well, phase shifts of half the original shifts must be implemented (cf. Section 6.2.2).

It was shown that the ASR algorithm performs astonishingly well if applied to even sampling rates, e.g., to 4-fold sampling. In this case, a reduction of the maximum Wiggling error of the factor 9 (at 5 MHz) and 12 (at 35 MHz) up to 34 (at 20 MHz) can be expected (cf. Tab. 6.2).

Above 20 MHz, the Wiggling Error can be regarded as eliminated as its maximum falls below 0.1 cm, which is in the same order of magnitude as other depth deviations.

One particular advantage of the algorithm is its additive nature, thus if two consecutive frames are detected to be inconsistent, it can be simply deactivated. For this reason it is attributed with the term *adaptive*. Further investigations may focus on making the ASR algorithm truly adaptive instead of just enabling or disabling it.

It should be remarked that although the algorithm takes information of two consecutive frames, depth information is delivered each single frame. Therefore, no frame-rate decrease need be expected.

7.1.3 Implications on Calibration

Based on the thorough analysis of systematic errors carried out in the course of this Thesis, calibration procedures suggested in literature are evidently obsolete. Instead, the obtained results indicate that the focus should be moved towards algorithmic error compensation. Most important (and most challenging) example is the effect of multipath reflections which definitely cannot be calibrated. As an overview, for error the implications of the results regarding calibration are shortly discussed.

Wiggling Error: Can be removed entirely above 20 MHz via the ASR algorithm, or at least reduced up to a high degree by 3-fold sampling. Depending on application and modulation frequency, calibration should be obsolete in many cases.

Amplitude-Related Error: Was shown to be a result of multipath reflections. No calibration is possible.

Integration-Time Related Error: Most probably an issue of the illumination unit. If this theory can be proven, temperature data from the illumination unit can be used for algorithmic correction. Calibration that is left in this context refers to the illumination unit, but not to the ToF imager.

Temperature-Related Error: Fully deterministic, linear dependencies between ambient (room) temperature and depth deviation were observed. For an extended temperature range, this dependency will be more complicated. Calibration will be most probably needed, but hopefully not sample-based.

Pixel-Related Error (FPN): No algorithmic compensation is possible. However, due to the fact that FPN was shown to be consistent over all parameters, calibration is simple.

7.2 Outlook

As an extension to the achievements discussed in the previous section, this Thesis concludes with an outlook on issues that could not be evaluated satisfactorily and therefore remain as impulses for further research.

Extending Temperature Range. The results from Section 5.1.2 need to be extended to a wider range of temperatures.

Moving-Target Experiment: Depth Drift. It was observed in Fig. 5.6 that an additional depth drift of the whole SRF occurred at larger target distances. Whereas the definite root cause remains unrecognized, one plausible reason is the introduction of multipath reflections at larger distances. On the one hand, the signal amplitude evidently decreases, which potentially increases reflection-based errors (cf. discussion in the context of Fig. 3.2). On the other hand, the scene composition naturally changes due to the changed position of the target, with more reflecting surface of the table exposed at larger distances. Although the conclusions from the moving-target experiment are still valid, it should be repeated in a environment that is free from unwanted reflections.

Integration Time versus Illumination-Unit Temperature. It was argued in Section 5.3.1 that increased temperatures of the illumination unit are most likely responsible for integration-time related depth deviations. However, this hypothesis needs to be verified with actual temperature data of the illumination unit at hand.

Amplitude-Related Error: Algorithmic Compensation of the Jumping-Edge Behavior. It was observed in Section 5.3.2 that in areas of changing signal amplitudes, depth data appears to be discontinuous. This jumping-edge behavior was shown to be a result of multipath reflections and simulated by the theoretic model. Given this as a starting point, generic correction algorithms may be developed.

Bibliography

[1] V. Castañeda and N. Navab, "Time-of-Flight and Kinect Imaging," http://campar.in.tum.de/twiki/pub/Chair/TeachingSs11Kinect/2011-DSensors_LabCourse_Kinect.pdf, 2011 (last accessed Mar 16 2012).

[2] B. Schneider, "Der Photomischdetektor zur schnellen 3D-Vermessung für Sicherheitssysteme und zur Informationsübertragung im Automobil," Ph.D. dissertation, Department of Electrical Engineering and Computer Science at University of Siegen, 2003.

[3] R. Lange, "3D Time-of-Flight Distance Measurement With Custom Solid-State Image Sensors in CMOS/CCD Technology," Ph.D. dissertation, Department of Electrical Engineering and Computer Science at University of Siegen, 2000.

[4] M. Schmidt, "Analysis, Modeling and Dynamic Optimization of 3D Time-of-Flight Imaging Systems," Ph.D. dissertation, Ruperto-Carola University of Heidelberg, Germany, 2011.

[5] A. Spickermann, "Photodetektoren und Auslesekonzepte für 3D-Time-of-Flight-Bildsensoren in 0,35-um-Standard-CMOS-Technologie," Ph.D. dissertation, Fraunhofer Institute for Microelectronic Circuits and Systems, University of Duisburg-Essen, 2010.

[6] S. Foix, G. Alenya, and C. Torras, "Lock-in Time-of-Flight (ToF) Cameras: A Survey," *Sensors Journal, IEEE*, vol. 11, no. 9, pp. 1917 –1926, sept. 2011.

[7] H. Rapp, "Experimental and Theoretical Investigation of Correlating ToF-Camera Systems," Master's thesis, University of Heidelberg, Germany, 2007.

[8] M. Müllenhaupt, "Evaluating Time-of-Flight Sensors using an Industrial Robot," Master's thesis, Dept. of Computer Engineering and Microelectronics, Berlin University of Technology, 2010.

[9] M. Lindner and A. Kolb, "Calibration of the Intensity-Related Distance Error of the PMD ToF-Camera," in *Intelligent Robots and Computer Vision XXV: Algorithms, Techniques, and Active Vision*, D. P. Casasent, E. L. Hall, and J. Röning, Eds., vol. 6764. SPIE, 2007, p. 67640W.

[10] T. Kahlmann, F. Remondino, and H. Ingersand, "Calibration for increased accuracy of the range imaging camera swissranger," in *In Proceedings of the ISPRS Commission, Symposium Image Engineering and Vision Metrology*, 2006, pp. 1542 –1549.

[11] M. Plaue, "Technical Report: Analysis of the PMD Imaging System," Interdisciplinary Center for Scientific Computing, Univ. of Heidelberg, Tech. Rep., 2006.

[12] M. Frank, M. Plaue, H. Rapp, U. Koethe, B. Jaehne, and F. A. Hamprecht, "Theoretical and Experimental Error Analysis of Continuous-Wave Time-of-Flight Range Cameras," *Optical Engineering*, vol. 48, no. 1, 2009.

[13] S. May, D. Droeschel, D. Holz, S. Fuchs, E. Malis, A. Nüchter, and J. Hertzberg, "Three-Dimensional Mapping with Time-of-Flight Cameras," *J. Field Robotics*, pp. 934–965, 2009.

[14] S. Fuchs, "Multipath Interference Compensation in Time-of-Flight Camera Images," in *Pattern Recognition (ICPR), 2010 20th International Conference on*, aug. 2010, pp. 3583 –3586.

[15] W. Karel, S. Ghuffar, and N. Pfeifer, "Quantifying the Distortion of Distance Observations Caused by Scattering in Time-of-Flight Range Cameras," *International Archives of Photogrammetry Remote Sensing and Spatial Information Sciences*, vol. XXXVIII, pp. 316–321, 2010. [Online]. Available: http://www.isprs.org/proceedings/XXXVIII/part5/papers/228.pdf

[16] T. Möller, H. Kraft, J. Frey, M. Albrecht, and R. Lange, "Robust 3D Measurement with PMD Sensors," in *In: Proceedings of the 1st Range Imaging Research Day at ETH*, 2005.

[17] B. Buettgen and P. Seitz, "Robust Optical Time-of-Flight Range Imaging Based on Smart Pixel Structures," *Circuits and Systems I: Regular Papers, IEEE Transactions on*, vol. 55, no. 6, pp. 1512 –1525, july 2008.

[18] A. Kolb, E. Barth, R. Koch, and R. Larsen, "Time-of-Flight Sensors in Computer Graphics (State-of-the-Art Report)," in *Proceedings of Eurographics 2009 - State of the Art Reports*. The Eurographics Association, 2009, pp. 119–134. [Online]. Available: http://www2.imm.dtu.dk/pubdb/p.php?5801

[19] C. Beder and R. Koch, "Calibration of Focal Length and 3D Pose Based on the Reflectance and Depth Image of a Planar Object," *Int. J. Intell. Syst. Technol. Appl.*, vol. 5, no. 3/4, pp. 285–294, Nov 2008. [Online]. Available: http://dx.doi.org/10.1504/IJISTA.2008.021291

[20] M. Frank, M. Plaue, and F. A. Hamprecht, "Denoising of Continuous-Wave Time-Of-Flight Depth Images using Confidence Measures," *Optical Engineering*, vol. 48, no. 7, July 2009.

[21] S. Fuchs and S. May, "Calibration and Registration for Precise Surface Reconstruction with ToF Cameras," 2007.

[22] S. Fuchs and G. Hirzinger, "Extrinsic and Depth Calibration of ToF-Cameras," in *Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on*, june 2008, pp. 1 –6.

[23] D. D. Lichti, C. Kim, and S. Jamtsho, "An Integrated Bundle Adjustment Approach to Range Camera Geometric Self-Calibration," *ISPRS Journal of Photogrammetry*

and Remote Sensing, vol. 65, no. 4, pp. 360 – 368, 2010. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0924271610000353

[24] D. D. Lichti and C. Kim, "A Comparison of Three Geometric Self-Calibration Methods for Range Cameras," *Remote Sensing*, vol. 3, pp. 1014–1028, 2011. [Online]. Available: http://www.mdpi.com/2072-4292/3/5/1014/pdf

[25] M. Lindner and A. Kolb, "Lateral and Depth Calibration of PMD-Distance Sensors," in *Advances in Visual Computing*, ser. Lecture Notes in Computer Science. Springer Berlin / Heidelberg, 2006, vol. 4292, pp. 524–533.

[26] A. Oppenheim and R. Schafer, *Discrete-Time Signal Processing (3rd ed.)*. Upper Saddle River, NJ, USA: Prentice-Hall, Inc., 2010.

[27] S. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*, ser. Prentice Hall Signal Processing Series. Prentice Hall, 1993, no. Vol.1.