Extending Time-of-Flight Optical 3D-Imaging to Extreme Operating Conditions

A dissertation submitted to the
FACULTY OF SCIENCE OF THE UNIVERSITY OF NEUCHATEL

for the degree of
DOCTOR OF SCIENCES

presented by
BERNHARD BÜTTGEN
Dipl.-Ing. Electrical Engineering

born December 9, 1976

accepted on the recommendation of

Prof. Dr. Peter Seitz, examiner
Prof. Dr. Edoardo Charbon, co-examiner
Prof. Dr. Pierre-André Farine, co-examiner
Prof. Dr. Shoji Kawahito, co-examiner

2006

Université
de Neuchâtel

FACULTE DES SCIENCES
Secrétariat-Décanat de la faculté
- Rue Emile-Argand 11
- CP 158
- CH-2009 Neuchâtel

IMPRIMATUR POUR LA THESE

Extending Time-of-Flight Optical 3D-Imaging to Extreme Operating Conditions

Bernhard BUETTGEN

UNIVERSITE DE NEUCHATEL

FACULTE DES SCIENCES

La Faculté des sciences de l'Université de Neuchâtel,
sur le rapport des membres du jury

MM. P. Seitz (directeur de thèse),
P.-A. Farine,
E. Charbon (EPF Lausanne) et
S. Kawahito (Hamamatsu, Japon)

autorise l'impression de la présente thèse.

Neuchâtel, le 11 janvier 2007

Le doyen :
J.-P. Derendinger

UNIVERSITE DE NEUCHATEL
FACULTE DES SCIENCES
Secrétariat-décanat de la faculté
Rue Emile-Argand 11 - CP 158
CH-2009 Neuchâtel

Bibliografische Information der Deutschen Bibliothek

Die Deutsche Bibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über http://dnb.ddb.de abrufbar.

Bernhard Büttgen:

Extending Time-of-Flight Optical 3D-Imaging to Extreme Operating Conditions

www.sierke-verlag.de
ISBN 978-3-933893-92-5

1. Auflage 2007

Contents

Abstract

This work concentrates on the development of compact and economical real time distance measurement cameras, whose measurement principle is based on the time-of-flight method. For distance measurement from a few centimeters to several meters, sensors based on so-called demodulation pixels are used.

Existing time-of-flight camera systems are restricted in operation due to limited distance resolution and a significant susceptibility to interference in relation to high signal dynamics in the scene, strong background incident light intensity or interference by other camera signals. There thus emerges four important areas of development, for which solutions are presented in this work:

1. Enabling multi-camera operation.
2. Extension of pixel intelligence for capturing highly dynamic scenes.
3. Active background light suppression at pixel level.
4. Pixel concepts for measurements with high distance resolution.

Multi-camera functionality is realised with an optical intensity modulation on the basis of pseudo-noise sequences. A detailed analysis of distance measurement with pseudo-noise modulated light signals includes the following aspects: the physically attainable statistical distance measuring accuracy, limited by photon noise; systematic sources of error such as low-pass distortions of the modulation signal; the ability to suppress interfering camera signals.
While the stochastic accuracy in single camera operation is, as expected, around a factor $2\sqrt{2}$ to 4 worse than for a sinusoidally modu1ated light signal, multi-camera operation is realistic with pseudo-noise modulation. In experiments with up to five cameras running simultaneously, distance perturbations under $10\mathrm{cm}$ were measured.

An extension of the signal dynamic range of the demodulation pixels is achieved by the integration of a comparator circuit in the pixel. The circuit facilitates the control of

the integration times of individual pixels.
With the test setup used, a gain in the dynamic range of over 24dB was obtained, although this can be further increased by a skillful adjustment of the integration times.

The pixel concept for background light suppression is based on a column oriented comparator circuit, which supervises the pixel outputs during the readout process. It prevents electrons generated by background light from arriving at the noise sensitive readout circuit.
Measurements show that, on the assumption of typical characteristics of the overall system, an indirect background beam of light, corresponding to a factor of 1.3 times that of natural sunlight, can be suppressed.

An improvement in distance resolution can only be achieved by higher (de-)modulation frequencies. Existing pixel structures have limited the possible modulation frequency to some tens of megahertz because charge separation was too slow. The pixel concepts presented here solve the problem by the utilisation of fast and therefore efficient charge transfer by lateral drift fields within the semiconductor. Two kinds of pixel are introduced: the first structure uses modulated fields and the second static lateral fields. The theoretically estimated maximum frequencies lie within the range of several hundred megahertz.
Distance measurements confirm the functionality of both pixels. Signal transit times of under 1ns are measured within the pixel with static drift fields, which emphasises the efficiency of the lateral electrical fields in the semiconductor.

This work makes a substantial contribution to the extension of the range of application of 3D time-of-flight technology: cameras with pseudo-noise modulation can be used in an environment with an almost arbitrary number of cameras running simultaneously. The extension of the dynamic range makes the employment of the technology possible in demanding environments with objects having strongly varying distance and reflection characteristics, so that outdoor applications become feasible. The drift field pixel structures permit 3D measurements in real time with sub-millimeter resolution, since particularly high demodulation frequencies are supported.

Kurzfassung

Die vorliegende Arbeit konzentriert sich auf die Entwicklung von kompakten und kostengünstigen Echtzeit-Distanzkameras, deren Messprinzip auf dem Lichtlaufzeitverfahren beruht. Zur Distanzerfassung von wenigen Zentimetern bis einigen Metern werden Sensoren auf Basis von sogenannten Demodulationspixeln eingesetzt.

Ein restriktiver Einsatz bestehender Lichtlaufzeit-Kamerasysteme ist auf eine limitierte Distanzauflösung und eine signifikante Störanfälligkeit gegenüber hoher Signaldynamik in der Szene, starker Hintergrundlichteinstrahlung oder Interferenzen durch andere Kamerasignale zurückzuführen. Folglich ergeben sich vier bedeutsame Entwicklungsschwerpunkte, zu denen in dieser Arbeit Lösungen vorgestellt werden:

1. Ermöglichung von Mehrkamerabetrieb.
2. Erweiterung der Pixelintelligenz zum Erfassen von hoch dynamischen Szenen.
3. Aktive Hintergrundlichtunterdrückung auf Pixelebene.
4. Pixelkonzepte für Messungen mit hoher Distanzauflösung.

Mehrkamerafunktionalität wird mit einer optischen Intensitätsmodulation auf Basis von Pseudo-Noise-Sequenzen realisiert. Eine detaillierte Analyse der Distanzmessung mit Pseudo-Noise modulierten Lichtsignalen schliesst folgende Aspekte ein: die physikalisch erreichbare statistische Distanzmessgenauigkeit, begrenzt durch Photonenrauschen; systematische Fehlerquellen wie zum Beispiel Tiefpassverzerrungen des Modulationssignals; die Fähigkeit zur Unterdrückung interferierender Kamerasignale.
Während die stochastische Genauigkeit im Einzelkamerabetrieb wie erwartet um einen Faktor $2\sqrt{2}$ bis 4 schlechter ist als im Falle eines sinusmodulierten Lichtsignals, ist mit Pseudo-Noise-Modulation sinnvoll ein Multi-Kamera-Betrieb realisierbar. In Versuchen mit bis zu fünf gleichzeitig laufenden Kameras werden Störungen der Distanzen unter $10\mathrm{cm}$ gemessen.

Eine Erweiterung des signalbezogenen Dynamikbereiches der Demodulationspixel erfolgt durch die Integration einer Komparatorschaltung in das Pixel. Die Schaltung ermöglicht eine pixel-individuelle Steuerung der Integrationszeit.
Mit den vorgenommenen Testeinstellungen wird ein Dynamikgewinn von über 24dB erzielt, wobei sich dieser durch eine geschickte Anpassung der Integrationszeiten weiter steigern lässt.

Das Pixelkonzept zur Hintergrundlichtunterdrückung basiert auf einer spaltenweise angeordneten Komparatorschaltung, die die Pixelausgänge während des Auslesevorgangs überwacht. Sie verhindert, dass durch Hintergrundlicht generierte Elektronen in die rauschempfindliche Ausleseschaltung gelangen.
Messungen zeigen, dass unter Annahme von typischen Eigenschaften des Gesamtsystems ein indirekter Hintergrundlichteinfall entsprechend dem $1,3$-fachen der natürlichen Sonnenleistung unterdrückt werden kann.

Eine Verbesserung der Distanzauflösung kann nur über höhere (De-)Modulationsfrequenzen erreicht werden. Bisherige Pixelstrukturen haben die mögliche Modulationsfrequenz wegen der zu langsamen Ladungstrennung auf einige zehn Megahertz begrenzt. Die hier vorgestellten Pixelkonzepte lösen das Problem durch die Ausnutzung von schnellem und daher effizientem Ladungstransport durch laterale Driftfelder innerhalb des Halbleiters. Es werden zwei Pixelarten eingeführt, wobei die erste Struktur modulierte, die zweite statische laterale Felder einsetzt. Die theoretisch abgeschätzten Grenzfrequenzen liegen im Bereich von mehreren hundert Megahertz.
Distanzmessungen bestätigen die Funktionalität beider Pixel. Signallaufzeiten von unter einer 1ns werden innerhalb des Pixels mit statischen Driftfeldern gemessen, was die Effizienz der lateralen elektrischen Felder im Halbleiter hervorhebt.

Die Arbeit liefert einen wesentlichen Beitrag zur Erweiterung der Anwendungsmöglichkeit der 3D-Lichtlaufzeit-Technologie: Kameras mit Pseudo-Noise Modulation können in einer Umgebung mit nahezu beliebiger Anzahl gleichzeitig laufender Kameras eingesetzt werden. Die Erweiterung des Dynamikbereiches ermöglicht den Einsatz der Technologie in anspruchsvollen Umgebungen mit stark variierenden Distanz- und Reflektionseigenschaften der Objekte, so dass Außenbereichsapplikationen realisierbar werden. Die Driftfeld-Pixelstrukturen erlauben 3D-Messungen in Echtzeit mit Submillimeterauflösung, da besonders hohe Demodulationsfrequenzen unterstützt werden.

Symbols

Counter variables

g, l, m, v, μ, ν

General parameters

a, b, C_1, x, y, γ

Greek or fractal mathematical signs

$\mathfrak{c}(t)$	signal over time
$\mathfrak{C}(f)$	Fourier spectrum of a signal $\mathfrak{c}(t)$
$\mathfrak{M}$	matrix describing the optical multi-layer system
$\mathfrak{t}$	optical transmission in terms of the amplitude
α	absorption coefficient
δ	distance error
η	refraction index
ϵ	dielectric constant of a material
ϵ_0	permittivity in vacuum
ϵ_{si}	dielectric constant of silicon
ϵ_{ox}	dielectric constant of silicon-oxide
ϵ_-, ϵ_+	lower and upper bound
λ	wavelength
λ_0	lower boundary wavelength of an optical bandpass
λ_c	cut-off wavelength (optical)
λ_{sig}	wavelength of the optical modulation signal
μ_0	charge mobility in the linear region of electric fields
σ	standard deviation
φ	phase
$\Delta\varphi$	phase difference
φ_Σ	phase of the complex pointer, which is the sum of several pointers
Φ	correlation value/function

Φ_a	autocorrelation
$\Phi_{a,lp}$	low-pass filtered autocorrelation function
$\Phi_{a,mpp}$	autocorrelation function with the multi-path propagated signal
$\Phi_{a,s}$	autocorrelation value only based on the signal component (no background light is taken into account)
Φ_c	cross-correlation
$\Phi_{c,\Sigma}$	sum of cross-correlation values
Φ_{Euler}	Euler's totient
Φ_Δ	discriminator function
Ψ	Potential
Ψ_{max}	maximum channel potential
Ψ_{ms}	work function difference between metal and semiconductor
Ψ_s	surface potential
ρ	optical reflection coefficient
ρ_q	charge density
τ	time-of-flight
τ_{mpp}	time-delay of the multi-path propagated signal component
τ_r	recombination lifetime
τ_c	mean free time
Θ_{avg}	expectation of the ratio between the cross-correlation values and the peak value of the autocorrelation
Θ_{prob}	occurance probability of the maximum cross-correlation value among a set of m-sequences
Θ_{max}	ratio between the maximum of the cross-correlation value and the maximum autocorrelation peak
ϑ	temperature
Υ	optical intensity

Roman mathematical signs

A	measured amplitude in number of electrons per tap
A_{sig}	mean number of electrons generated by the optical signal component per tap
A_g	charge packet of integration node with number g
$\hat{A}_g$	noisy charge packet of integration node with number g
A_{image}	sensor area
A_{pix}	total pixel area
A_{sense}	sensitive pixel area
B	measured offset in number of electrons

BC	number of electrons generated by the signals of concurrently running cameras
BD	number of electrons generated by the dark current
BG	number of electrons generated by background light
BS	background-to-signal light ratio
b_w	width of the first maximum of the diffraction pattern behind an optical system
c	speed of light in vacuum
c_d	demodulation contrast
c_m	modulation contrast
c_p	non-ideality factor of the pixel
c_{us}	speed of ultrasonic waves in air
D	diameter of an optical aperture, diffusivity
DR_{se}	dynamic range of the sensor
DR_{sc}	dynamic range of the scene
dR	resistance of a part of a resistor line
d_{ox}	oxide thickness
E	electric field
E_0	electric field where the saturation velocity sets in
E_{air}	electric field vector in the air region
E_{pix}	energy per pixel
E_{ph}	photon energy
E_{sense}	energy per sensitive pixel area
$\hat{F}$	complex frequency
FF	fill factor
FW	full well
$F/\#$	F-number
f	frequency
f_c	chip frequency
f_{ce}	center frequency
f_{cut}	cut-off frequency
f_m	modulation frequency
f_s	sampling frequency
Δf	frequency difference
G	generation efficiency
G_{tr}	gain of transit time
G_{mct}	gain of the storage capacity of the MCT pixel
GP	generator polynomial
g_p	polynomial coefficient
H_{air}	magnetic field vector in the air region
$H(x)$	frequency of occurence
$h(t)$	impulse response function

h	Planck's constant
i, I	photo-current
K_B	Boltzmann's constant
k	optical attenuation factor
k_{lens}	optical losses of the lens
k_{mpp}	attenuation factor of the multi-path propagated signals
L_{diff}	diffusion length
l_f	focal length
$M(t)$	modulation signal
$M_b(t)$	bipolar modulation signal
$M_u(t)$	unipolar modulation signal
M_p	number of primitive polynomials
m_{eff}	effective mass
m_p	register length of linear feedback shift registers
N_A	concentration of acceptor ions
N_c	number of cameras
N_{diff}	number of diffusion regions
N_{dop}	doping concentration
N_{drift}	number of drift regions
N_D	concentration of donor ions
N_{el}	number of electrons
N_{low}	low doping concentration
N_{ph}	number of photons
N_{ps}	number of pseudo noise electrons
N_s	number of samples
N_t	number of pixel taps
n	density of free electrons
n_i, n_a	integrated and added noise sources
n_p	PN sequence length
P	optical power
P_e	emitted optical signal power
P_A	mean optical signal power
$P_{A,\Sigma}$	amplitude of the optical power signal after the addition of several sinusoidally modulated signals
P_B	background light power signal
P_{image}	optical power on sensor
P_{ls}	optical power of light source
P_{object}	optical power on object
P_{pix}	optical power on pixel
p_r, P_r	received optical power signal
p	prime factor
QE	quantum efficiency
Q	charge
Q_{ox}	field oxide charge

q	elementary charge
R	distance to the object
R_0	non-ambiguity range
R_{max}	maximum distance
R_{meas}	measured distance
R_t	true distance
R_{ph}	optical responsivity
$r(t)$	reference modulation signal applied to the pixel's toggle gates
r_Σ	radius of the circle of erroneous phase
S	switch control signal
s	path length
$ss(t)$	spread spectrum function
t	time parameter
t_{diff}	diffusion time
t_{drift}	drift time
t_{tr}	transit time
T	integration time
T_c	chip duration
T_m	modulation period
V	voltage
V_{bi}	built-in voltage
V_D	supply voltage
V_{eff}	effective voltage
V_{FB}	flat-band voltage
V_g	gate voltage
V_{th}	threshold voltage
v	velocity
v_{th}	thermal velocity
w	depletion width
X	unit discrete time-delay
x_{pn}	depth of the pn-junction in the semiconductor bulk
x_s	location of the maximum potential in the semiconductor bulk

Abbreviations

ACF	autocorrelation function
AD	analogue-to-digital
APS	active-pixel sensor
B-CCD	buried channel CCD
BL	background light
CCD	charge-coupled-device
CDMA	code-division multiple access
CMOS	complementary metal oxide semiconductor
CS	cross section

CSEM	Centre Suisse d'Electronique et de Microtechnique
CW	continuous wave
DFD	drift-field demodulation
DFT	Discrete Fourier Transform
DS - CDMA	direct sequence code-division multiple access
DSP	digital signal processing
ESPI	electronic speckle pattern interferometry
FDMA	frequency-division multiple access
FH - CDMA	frequency hopping code-division multiple access
gc	gate contact
IG	integration gate
LFSR	linear feedback shift register
LP	low-pass
MCI	multi-camera interference
MCT	minimum charge transfer
MPP	multi-path propagation
MTF	modulation transfer function
m - sequence	maximal-length sequence
OUTG	out gate
PG	photo-gate
PGL	photo-gate left
PGM	photo-gate middle
PGR	photo-gate right
PLL	phase-locked loop
PN	pseudo noise
PWI	pixel-wise integration
QE	quantum efficiency
ROI	region-of-interest
SAR	synthetic aperture radar
S-CCD	surface channel CCD
SDMA	space-division multiple access
SEM	scanning electron microscope
SOC	system-on-chip
SPAD	Single photon avalanche diode
SRH	Shockley-Read-Hall
std. dev.	standard deviation
TDMA	time-division multiple access
TG	transfer gate
TH - CDMA	time hopping code-division multiple access
TOF	time-of-flight
WDMA	wavelength-division multiple access
VCO	voltage-controlled oscillator
2D	two-dimensional
3D	three-dimensional

Chapter 1

Introduction

Real-time three-dimensional (3D) imaging has become very popular during the last few years. As often seen in nature, information about the third dimension is very helpful for simplifying or even enabling a lot of tasks in the fields of navigation, observation and security. In these examples different methods for acquiring the distance information are exploited. The following examples point out the usefulness and the power of the 3D-imaging systems found in nature:

- The jumping spider has a total of eight eyes, where two eyes are located on each side of the back head and four eyes on the forehead. This composition provides an all-around view. By stereoscopic superposition of the images, three-dimensional perception is achieved. Figure 1.1 shows a picture of the spider with the two big eyes in the front.

- Bats explore their surrounding environment by measuring the reflections of ultrasonic signals they emit. Distance information to objects is obtained by the time-of-flight measurement. Bats emit ultrasonic signals and measure the reflected signal deformation and delay. Based on these parameters they obtain information about the distances to the objects and their shape. In fact, bats are capable of handling multiple reflections, for example produced by a complex environment such as trees. [116]

- Dolphins explore their environment by ultrasonic measurement as well. Figure 1.2 shows their special sensing instruments including the melon, nasal sacs and specialized skull structures. Reflected pulses from the swim bladders of other animals are detected. In this way, dolphins are able to explore their environment in a three-dimensional sense, observe potential enemies or recognize any prey. [67, 116]

Figure 1.1: Jumping spider has a total of eight eyes.

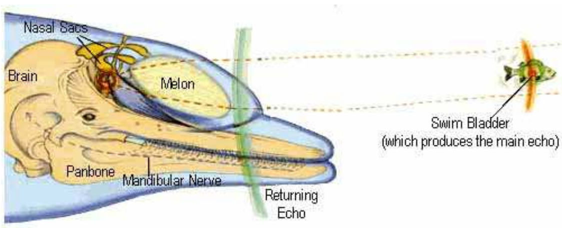


Figure 1.2: Echolocating systems of dolphins. [67]

The performance of the visual systems of the above-mentioned animals is remarkable: The spider processes eight images concurrently. The bat handles ultrasonic signals from multiple reflections and evaluates distances to objects as well as their material properties. Conflicts with the ultrasonic systems of other bats rarely occur. Each bat separates its own ultrasonic signal. And the dolphin's navigation system has to manage the high-speed processing of ultrasonic signals that propagate four and a half times faster in water than in air. Dolphins can orient themselves even in very noisy environments with respect to ultrasonic interferences.

These examples show that a three-dimensional perception of the environment allows for an improved orientation even in environments where the optical contrast in intensity is very low (bats are active during the night and dolphins live in the sea where the light is strongly scattered and absorbed depending on the optical wavelength and the depth below sea level).
This additional information provided by the third dimension is also exploited in technical systems, enabling new applications which are briefly summarized in the next subsection as part of the motivation for the current thesis work. The challenges for building technically similar 3D-vision systems as found in nature are obvious. The computation of 3D-models of the environment by the composition of several two-dimensional (2D) images necessitates fast parallel processing and consequently it is very power consuming. The perception of the environment by actively emitting signals and exploiting the deformations of the reflected signals requires fast signal processing as well. Furthermore, conflicts due to potential interferences with other signals have to be resolved.

1.1 Motivation

The main motivation for developing an optical 3D-camera is to enable different new applications exploiting the additional 3D-information of the scene. Fields of applications are safety systems, biometrical measurements, security, machine vision, machine-user interface (MUI), robotics [18], gaming. [75] A few of them are briefly discussed in the following.

Safety systems can profit from real-time 3D-imaging. As an example, smart airbag systems can be realized. [75] By observing the movements of a car's occupants, the force of an explosion of the airbag in the case of an accident can be accurately adapted to the current situation. Potentially, with the use of a 3D-imaging system outside the car in a so-called pre-crash detection system, the collision could have been avoided completely. Pre-crash detection systems are part of the general development of predictive driver assistance systems that cover three main development steps: adaptive cruise control, predictive safety systems and systems for active collision avoidance. [58] The real-time acquisition of the 3D-information is useful for all development steps.
3D-cameras can be exploited for the implementation of new computer input devices. The detection of the movements for example of the head into any direction could facilitate the life of handicapped persons enormously.
In industry, real-time 3D-cameras could be used for the accurate control of robots. For example, a robot can perform tasks in a dangerous environment while it is controlled via 3D-interaction by a remote person in a safe environment.

These are just a few examples that motivate someone to develop further existing optical 3D-cameras. The improvements of the current technologies target the following items:

- compact system,
- 3D-image acquisition with high frame for real-time applications,
- high lateral resolution,
- high distance resolution,
- operation of one camera in an uncontrolled environment with high background light power,
- operation of one camera in an environment where an arbitrary number of additional cameras could approach but must not distort the 3D-acquisition due to interference effects.

These targets of technology improvement are the subject of this dissertation. The first two items concerning the compact system and high frame rate which inherently implies low processing power, favour the use of the time-of-flight method for 3D-measurement.

While stereoscopic or triangulation methods cannot arbitrarily be miniaturized because they require a minimal baseline between the sensor and the camera or between two cameras, respectively, the time-of-flight principle is an excellent method for exploitation with smart pixel structures. A one-chip solution is imaginable because the time-of-flight distance measurement can be performed in parallel by smart-pixel structures as described herein.

1.2 Outline

This dissertation begins with a first theoretical part treating the time-of-flight (TOF) principle. The focus lies on sinusoidally modulated TOF measurement. In direct relation to the pixel model, two approaches for extracting the distance information are discussed: demodulation by sampling and demodulation by correlation. The limitations of these modulation techniques are summarized. In order to understand how to dimension the systemn, a section on the power budget is included. Finally, the requirement of the TOF system in terms of the dynamic range are highlighted.

Chapter 3 treats the system's enhancement to multi-camera operation. Theoretical limitations and performances for single-camera as well as multi-camera operations are derived mathematically, where Pseudo-Noise sequences are used as modulation signal. The error analysis covers the multi-camera interference (MCI), the stochastic accuracy based on photon shot noise, the system non-linearities, the impact of background light, multi-path propagation (MPP) and the effect of low-pass filtered modulation signals. The chapter concludes with experimental results.

Chapter 4 summarises the fundamental physical properties of solid-state photo-sensing. The first section answers the question: "How are photons detected with a semiconductor material?" The second section gives an overview of the two major standard process technologies used for solid-state imagers. After that, the principle of charge-coupled devices (CCD) is highlighted, covering the charge storage, charge transport and charge detection and the achievable quantum efficiency (QE) of such devices.

Chapter 5 introduces the demodulation pixels that have been developed for the extension of the time-of-flight 3D-imaging to various operating conditions: The pixel-wise integration (PWI) pixel is used for enhancing the dynamic range of the standard lock-in pixel; the minimum-charge transfer (MCT) pixel still operates under high background light incident and the drift-field demodulation (DFD) pixels are designed for high frequencies. These three concepts are discussed section by section, where each section concludes with a subsection about the experimental verification of the pixel concept.

At the end of the dissertation a brief summary of the achieved results and an outlook for future work are given.

Chapter 2

Optical Time-of-Flight 3D-Imaging

The application of the time-of-flight method in optical 3D-imaging is discussed in this chapter. The basic principle is classified within the group of contactless distance measurements and explained on the basis of sinusoidal demodulation. A power budget for understanding the design parameters of the whole system is described. Challenges in the design of the 3D-TOF system are dynamic range and multi-camera operation.

The considerations are related to optical distance measurements although other types of waves could also be used for the contactless distance measurement, for example ultrasonic waves, microwaves etc. Some well-known examples of ranging applications with ultra-sound signals are boat sonars, material characterization [39, 46] and automatic parking systems [58]. The repetition rate of ultrasound systems is mainly limited by the propagation speed of about $c_{us} \approx 340\mathrm{m/s}$ in air at a temperature of $\vartheta = 293K$, which is slow compared to the speed of light. Generally, ultrasound systems are not dangerous for human beings or animals as long as no cavitation occurs. Microwaves are used for example in radar applications for air and sea boat observations [101] and for 3D-mapping of the earth by synthetic aperture radar (SAR) technology [101]. Advantages are low absorption, for example by water particles in air and fast propagation speed (speed of light) which makes them applicable for long-distance ranging over several kilometers.

But why do we concentrate herein on electromagnetic waves in the optical range for 3D-imaging? The applications mentioned before are mainly applications where speed and compactness are no relevant issues, or just a point-wise range measurement is performed. However, a compact 3D-camera delivering real-time streams of 3D-images needs to exploit small-sized sensors with a high density of integrated distance measuring pixels. Neither ultrasonic waves nor microwaves have the ability to be imaged on small pixel structures having sizes of a few micrometers. This is due to the diffraction phenomenon which determines the minimal lateral resolution of waves when they pass through an aperture [86,92]. Table 2.1 compares the lateral resolutions of ultrasonic waves, microwaves and optical waves when a circular aperture of diameter $D = 2.44\mathrm{cm}$ and a focal length $l_f = 10\mathrm{cm}$ is assumed. The diffraction spot sizes have been computed by the Rayleigh criterion. Only optical waves allow for forming beams in the image plane having diffraction spot sizes b_w of a few micrometers. Consequently, based on smart pixel structures, optical 3D-imaging is a preferred technique.

2.1 Classification of TOF Distance Measurement

We can distinguish between three different methods of contactless distance measurement as indicated by figure 2.1 [93]: triangulation, interferometry and time-of-flight. Unified analysis of all three methods have shown that they are fundamentally limited by the same physical limitations set by the light-inherent noise source and the fundamental theoretical behaviours of all of them can be expressed in a very similar fashion [96]. However, the practical implementations are completely different and make each method suitable for certain applications. An overview of currently commercially available products for the distance measurement based on triangulation, time-of-flight or interferometry, is provided for example in [17].

Triangulation: The geometrical properties of the measurement set-up are exploited for the extraction of the distance. We distinguish between active and passive triangulation methods. As an example, the basic principle of an active triangulation system is shown in figure 2.2 (a). When the baseline b, the focal length of the optical system l_f and the distance x of the spot imaged onto the detector are known, the distance of this

Table 2.1: Comparison between the lateral resolutions of different types of waves based on the diffraction phenomenon [92, 93]. Assumptions are a diameter of the optical aperture of $D = 2.44\mathrm{cm}$ and a focal length of $l_f = 10\mathrm{cm}$.

	ultrasound	microwaves	optical waves
wavelength λ	$3\mathrm{mm}$	$3\mathrm{mm}$	$0.6\mu\mathrm{m}$
frequency f	$110\mathrm{kHz}$	$100\mathrm{GHz}$	$500\mathrm{THz}$
diffraction spot size b_w	$3\mathrm{cm}$	$3\mathrm{cm}$	$6\mu\mathrm{m}$

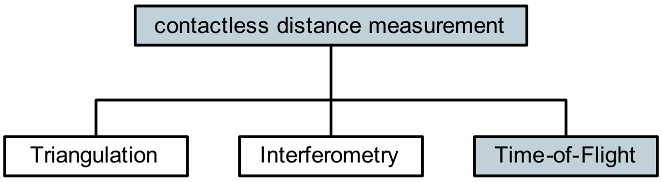


Figure 2.1: Methods of contactless distance measurement. [93]

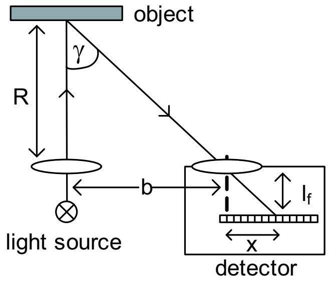
particular triangulation set-up is computed as [96]

$$R = \frac{b}{\tan(\gamma)} = \frac{b \cdot l_f}{x}. \tag{2.1}$$

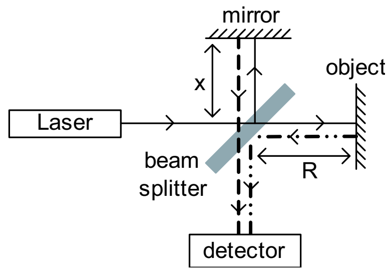
To obtain 3D-images, this set-up has to be enhanced by scanning components. An example of a triangulation range finder with synchronized scanners is given in [66]. The order of required scanning dimensions can be reduced by the lightsheet triangulation technique. Instead of emitting a single lightbeam that is imaged onto a one-dimensional detector array, a lightsheet is emitted and imaged onto a two-dimensional detector array [93]. Non-scanning active 3D-triangulation systems are generally based on the acquisition of brightness images of a scene that is illuminated by structured light. In this way the light can be spatially or temporally encoded. Potentially, several images have to be taken for their combination into one 3D-image. A large number of known projection patterns exists, each with its own advantages and disadvantages [45,93,118].
Passive triangulation methods do not emit an optical signal. They generally require the imaging of the object several times, but they do not necessarily need scanners. Two techniques should be mentioned. The first one is stereoscopic 3D-imaging. Several images are taken from different viewpoints of the camera and combined into one 3D-image [43]. The second method is called depth-of-focus [43]. Several 2D-images are taken with different focussing properties from one point of view. Exploiting the fact that the blurring effect due to defocused optics increases with increasing distance, 3D-images can be computed by this technique.

All of the before-mentioned triangulation techniques suffer from the high computational power or the need for a triangulation baseline and scanning components. These aspects make a set-up of a compact real-time 3D-image camera with high lateral resolution difficult with the usage of triangulation methods.

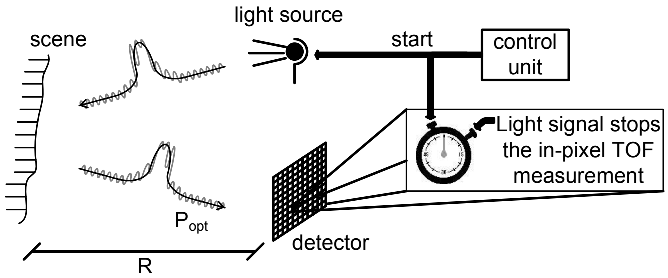
Interferometry: This kind of distance measurement exploits the fact that light can be understood as an electromagnetic wave that propagates with the finite speed of light $c \approx 3 \cdot 10^8$m/s in vacuum and having a wavelength of λ. The delay between the optical signal that travels along two different optical paths corresponds to the targeted



(a) Active triangulation.



(b) Michelson interferometer.



(c) Time-of-Flight.

Figure 2.2: Principle set-ups for the contactless distance measurement.

distance. Figure 2.2 (b) shows the principle set-up of a Michelson interferometer using a coherent light source, a beam splitter and a mirror in the reference path. Other well-known interferometer types are for example the Mach-Zehnder interferometer or the Sagnac interferometer [86]. They mainly differ in the characteristics of the optical paths. The intensity Υ on the detector, defined as the absolute square of the complex amplitude of the electromagnetic wave, is the result of the constructive or destructive superposition of the two optical waves coming from the reference path and the object path. It can be described as [60, 86]

$$\Upsilon = \Upsilon_1 + \Upsilon_2 + 2\sqrt{\Upsilon_1 \Upsilon_2} \cos\left(\frac{2\pi(R-x)}{\lambda}\right). \tag{2.2}$$

The distance R can easily be extracted when the intensity relations between the optical paths are known. Using two-dimensional pixel arrays as a detector, real-time 3D-images of the surface can be acquired without the need for scanning components.

While high distance resolution in the range of a fraction of the wavelength λ can be achieved - generally resolutions of $\lambda/100$ or $\lambda/1000$ are possible, which is in the sub-nanometer range - classical interferometry with coherent light has two drawbacks:

1. the non-ambiguity range is limited to $\lambda/2$ and 2. no absolute measurements are possible due to the small non-ambiguity range. Also, interferometric techniques like electronic speckle pattern interferometry (ESPI) [93] deal with this problem.
Multiwavelength interferometry is one possible method to overcome the two problems of getting relative distance data with a small non-ambiguity distance [109, 125]. Two close optical wavelengths are used for the measurement of the distance, where the beat frequency determines the enhanced non-ambiguity range.
Another method to enhance the non-ambiguity measurement range is white-light interferometry, also called low-coherence interferometry [19]. This technique combined with smart CMOS-pixels can even be used for multi-reflection measurements leading to a tomographic measurement system [13,14,19]. These systems have proven real-time capability with high distance resolutions in the sub-micrometer range.

Basically, interferometric systems are more suitable for short distance measurements of high precision. However, the 3D-TOF system, subject of this work, is targeted for measuring short and long distances in the range of several centimeters to several meters, for which the time-of-flight principle is much more suited.

Time-of-Flight: Both time-of-flight and interferometry belong to the group of delay measurements. While interferometric techniques use the shift of the carrier wave for the extraction of the distance, time-of-flight techniques accomplish a modulation of the carrier wave and hence, the delay of the modulation signal is measured. Consequently, the limitations of interferometric systems with respect to non-ambiguity and resolution are given by the carrier wavelength; in the case of the time-of-flight measurement the limitations are set by the modulation signal applied.

Figure 2.2 (c) shows the basic set-up of a time-of-flight 3D-imaging system. An intensity-modulated light signal is emitted by a light source. After being reflected by the object, the signal is detected by the sensor triggering the time-of-flight measurement. The distance is computed as

$$R = \frac{c \cdot \tau}{2}, \tag{2.3}$$

with the known speed of light c and the measured time-of-flight τ. In a real-time 3D-imaging system each pixel of the sensor has to have the capability of measuring the time-of-flight, as indicated in figure 2.2 (c).

With regard to the kind of modulation scheme used, one distinguishes betwen three methods of time-of-flight measurement [93]:

1. Pulse modulation: A short pulse of optical power is emitted and the time-of-

flight is directly measured. The advantage of this modulation scheme is that no distortions due to multi-path propagation occur. The influence of the background light can be kept low because high short-term optical signal-to-noise (and signal-to-background) ratio is attained while maintaining a low mean value of optical power [60]. This is an important issue with respect to eye safety. However, high bandwidths and high dynamic range of the electronic components are required. For example, a simple counter could be used for the time measurement. To obtain distance resolutions of one millimeter, a clock frequency of more than 150GHz would be necessary. Attenuations in the optical path lead to flattening of the pulse's slope, making the distance measurement more imprecise. Generally, lasers need to be used to generate optical pulses of high power and short rise and fall times. Pulse modulation is widely used in laser scanning systems. Example products can be found in [3, 5, 17, 30].

2. Continuous-wave (CW) modulation: A time-continuous signal is applied as intensity modulation, where all variations of signals are imaginable as used for the projection patterns in triangulation measurements [93]. The phase shift between emitted and detected signal is measured instead of the time-of-flight directly. We distinguish between heterodyne and homodyne measurements.
   Heterodyne measurements perform a mixing between two signals, whose frequencies are different. The beat frequency of the bandpass signal corresponds to the distance. Generally, heterodyne distance measurement allows for measuring long distances without ambiguity. The generation of the signals, however, is not straight-forward. An example of a modulation signal leading to heterodyne distance measurement is the chirp modulation technique, which is widely used in radar technology [101].
   Homodyne measurements use the correlation between two sinusoidal signals of the same frequency f_m. The phase shift φ between the emitted modulation signal and the reference signal corresponds proportionally to the time-of-flight according to
   $$\varphi = -2\pi f_m \tau. \tag{2.4}$$
   From the electronic point of view, the generation of a sinusoidally modulated light signal can conveniently be performed with laser or standard light emitting diodes (LEDs). Though high bandwidth is not required for the electronic components, nonetheless a lack of low-pass filtering would have no effect on the phase at all. Hence, systematic distortions of the distance measurements are expected to be small, only resulting from an imperfect generation of the sinusoidal modulation signal.

There are some optical distance measurement systems based on CW-modulation, commercially available. More in-depth information can be found in [6, 8, 17, 54].

3. Pseudo-noise (PN) sequences: These are sequences that show almost perfect auto-correlation/cross-correlation properties like white noise. The orthogonality between different signals can be used for suppressing signals of other cameras. The time shift between emitted and detected signal is generally measured by an evaluation of the significant Dirac-pulse-like auto-correlation function. The discrete nature of pseudo-noise sequences requires high bandwidths of the electronics. Furthermore, the signal generation needs more computational power for ensuring the randomness of the sequence.
   Distance measurement systems using PN modulated signals, can particularly be found in micro-wave radar systems [101]. An optical point-wise distance measurement system using PN-sequences, is presented in [56]. There, a PN-sequence is emitted and correlated with its own replica after having been reflected by the object. The distance corresponds to the chip rate of the reference sequence, which is controlled by a voltage controlled oscillator (VCO). The complexity of the system, however, makes it unsuitable for parallel distance measurements in a 3D-imaging system.

Combinations between pulse modulation and continuous-wave modulation could make use of the advantages of each technique with regard to the signal-to-noise ratio, signal-to-background ratio, eye safety and multi-path propagation.

In the following sections, time-of-flight measurement using sinusoidal optical intensity-modulation is discussed in more detail. It also forms the basis for the characterization of the demodulation pixels presented in chapter 5. Pseudo-noise modulation for real-time 3D-imaging is investigated in chapter 3 treating the multi-camera capability.

2.2 TOF Measurement Based on Sinusoidal Modulation

The ideal emitted and received optical signals of a time-of-flight measurement system based on sinusoidal optical modulation, are shown in figure 2.3. The emitted signal $P_e(t)$ is a sine wave of frequency f_m and its amplitude and offset are equal, according to

$$P_e(t) = P_A \cdot [1 + \cos(2\pi f_m t)]. \tag{2.5}$$

The detected or received signal is attenuated by a factor k due to any optical losses in the measurement path, phase-shifted by φ due to the travel time of the light signal and it has experienced a superposition of an additional offset component $P_B(t)$ due to background light in the environment. It is described by

$$P_r(t, \tau) = P_B(t) + k \cdot P_A \cdot [1 + \cos(2\pi f_m(t - \tau))]. \tag{2.6}$$

This consideration is idealized in such a way that the amplitude is assumed not to show any fluctuations and the modulation contrast of the light source is 1 because amplitude and offset of the emitted signal are the same. Thus, the light source is assumed to have ideal properties.
Using equations 2.3 and 2.4, the distance is computed as

$$R = \frac{c}{4\pi f_m} \cdot \varphi. \tag{2.7}$$

The phase of the detected sinusoidal modulation signal can only be non-ambiguously extracted as long as $0 \leq \varphi < 2\pi$. Consequently, the non-ambiguity range of the distance measurement is defined as

$$R_0 := \frac{c}{2f_m}. \tag{2.8}$$

Two methods for the extraction of the phase shift of the received optical signal are explained in the following subsections. The first method performs a sampling of the optical signal. Based on the samples, the amplitude and phase spectra can be calculated. The second method consists of a correlation process. The time-shift of the correlation function corresponds to the desired phase shift φ.
However, a brief explanation of the general pixel model is given beforehand in order to simplify the understanding of why both approaches can be used with the same demodulation pixel.

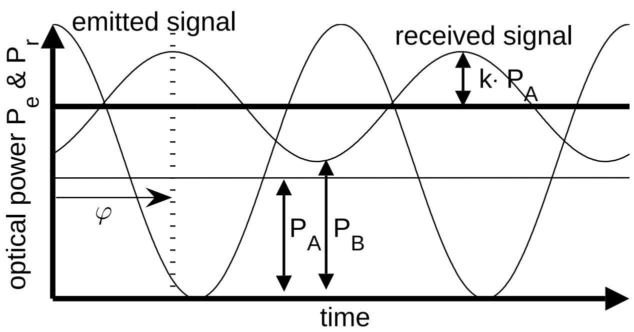


Figure 2.3: Optical signals in a TOF system based on sinusoidal modulation.

2.2.1 Simplified Pixel Model

The demodulation pixel converts the incident optical signal of power $P_r(t)$ into electrons which are directed to one of the N_t integration nodes. This is illustrated in figure 2.4. The integration nodes are also called taps, why t is used as subscript. The relation between the electron flux I and the optical power is described by a proportionality factor R_{ph} which is called the responsivity of the pixel [86]. By dividing the photo-current by the elementary charge q, the corresponding number of electrons is obtained. The N_t charge packets at the pixel's outputs are referred to as $A_0 ... A_{N-1}$. Mathematically they are described as

$$A_g = \int_{t_0}^{t_0+T} S_g(t) \cdot \frac{R_{ph} P_r(t)}{q} dt, \tag{2.9}$$

where S_g is the control function of the switch over time in position g. The switch has in its ideal implementation two states: 1$\widehat{=}$on, 0$\widehat{=}$off. The integration is started at an arbitrary time t_0 and it takes a total duration of T.

2.2.2 Demodulation by Sampling

The sampling of the photo-generated sinusoidally modulated signal by N_s equally spaced samples per period is shown in figure 2.5. Since each sample is noisy in practice, the signal-to-noise ratio is increased by accumulating each sample over several periods. One sample is described as

$$A_g = \sum_{l=1}^{m} a_{g,l}, \tag{2.10}$$

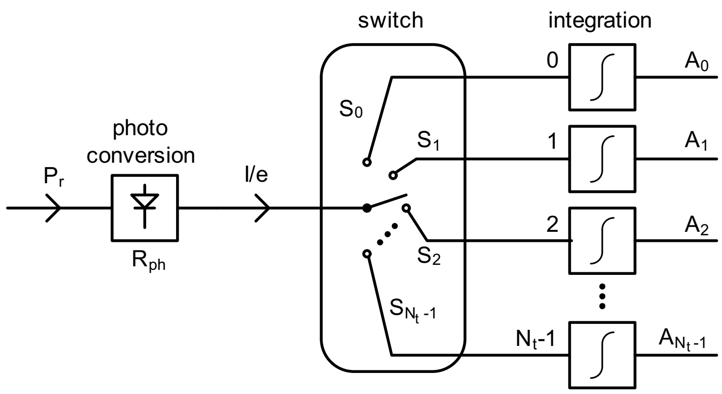


Figure 2.4: Simplified pixel model.

where $a_{g,l}$ is the g-th sample of the l-th period and $m = T/T_m$ is the total number of modulation periods. Following the theory of discrete Fourier transform (DFT), the sampling of one period of a continuous and periodic signal by an appropriate minimum number of samples according to the Nyquist theorem allows for unambiguous reconstruction of the signal. The discrete Fourier transform is defined as [79]

$$\hat{F}_n = \frac{1}{N_s} \sum_{g=0}^{N_s-1} A_g \cdot e^{-j2\pi ng/N_s}. \tag{2.11}$$

Since the modulation is sinusoidal, the first harmonic at $n = 1$ is of special interest. Its amplitude A and phase φ are computed as [60, 103]

$$A = \frac{2}{N_s} \sqrt{\left[\sum_{g=0}^{N_s-1} A_g \cos\left(2\pi \frac{g}{N_s}\right)\right]^2 + \left[\sum_{g=0}^{N_s-1} A_g \sin\left(2\pi \frac{k}{N_s}\right)\right]^2} \tag{2.12}$$

$$\varphi = -\arctan\left[\frac{\sum_{g=0}^{N_s-1} A_g \sin\left(2\pi \frac{g}{N_s}\right)}{\sum_{g=0}^{N_s-1} A_g \cos\left(2\pi \frac{g}{N_s}\right)}\right]. \tag{2.13}$$

The constant offset is expressed as

$$B = \frac{1}{N_s} \sum_{g=0}^{N_s-1} A_g. \tag{2.14}$$

With regard to the general pixel model in figure 2.4, the control function of the switch S describes a rectangular function in each position, so that each sample can be rewritten as a convolution

$$A_g = m \frac{R_{ph}}{q} \left[\mathrm{rect}\left(\frac{t - g\frac{T_m}{N_s}}{\Delta t}\right) * P_r(t)\right], \tag{2.15}$$

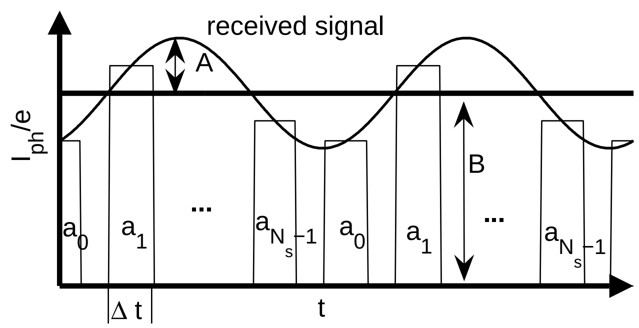


Figure 2.5: Sampling of the detected optical signal being converted into a current of electrons.

where the ratio between the modulation period T_m and the number of samples N_s is the equidistant space between the samples. The sampling time is denoted by Δt. The convolution with the rect function in the time domain means a multiplication with the sinc function in the frequency domain. While the amplitude is expected to be equal to the offset component when no background light is present, the low-pass effect due to the integration process reduces the amplitude by the so-called demodulation contrast c_d. In practice additional non-idealities of the demodulation process, for example inclined optical incidence, reduce the demodulation contrast even more, for which the signal theory provides an upper bound:

$$c_{d,max} = \mathrm{sinc}(\pi f_m \Delta t)) := \frac{\sin(\pi f_m \Delta t)}{(\pi f_m \Delta t)}. \tag{2.16}$$

For a sampling time of $\Delta t = T_m/4$ we expect a maximal demodulation contrast of 0.9, while $\Delta t = T_m/2$ results in a maximal contrast of 0.64. In the following, the parameter c_d denotes the pixel's demodulation contrast, taking into account the upper bound $c_{d,max}$, which is determined by the signal theory, and any non-idealities of the pixel's demodulation process c_p, so that

$$c_d = c_p \cdot c_{d,max}. \tag{2.17}$$

Demodulation with 4 Samples: Following the Nyquist theorem, three discrete samples of a sine wave are already sufficient for the unambiguous reconstruction of the sinusoidal signal. The practical implementation of the sampling stages within the pixel is more convenient to realise with an even number of samples. Hence, the equations for the extraction of the signal parameters based on four samples are summarised here. The amplitude, offset and phase are calculated as

$$A = \frac{\sqrt{(A_3 - A_1)^2 + (A_0 - A_2)^2}}{2} \tag{2.18}$$

$$B = \frac{A_0 + A_1 + A_2 + A_3}{4} \tag{2.19}$$

$$\varphi = \arctan\left[\frac{A_3 - A_1}{A_0 - A_2}\right]. \tag{2.20}$$

The sampling of the optical signal by four successive samples can be performed with any pixel structure that can be modeled according to figure 2.4 having either 1, 2 or 4 integration nodes. If just 1 or 2 integration nodes are incorporated, the optical signal will have to be sampled subsequently twice or three time, respectively. A pixel structure with 4 integration nodes is called a 4-tap structure and it enables the acquisition of all four samples within one integration cycle.

2.2.3 Demodulation by Correlation

In the following a pixel structure according to the model of figure 2.4 is considered under the assumption that it has incorporated just two integration nodes, for example A_0 and A_2. Assuming an integration time that is a multiple of the modulation period T_m, the difference Φ of these two values can be expressed with the use of equation 2.9 as

$$\begin{aligned} \Phi(\tau) &= A_0 - A_2 \\ &= \tfrac{R_{ph}}{q} \int\limits_{t_0}^{t_0+T} (S_0(t) - S_2(t)) \cdot P_r(t,\tau) dt \ , \\ &= G \cdot (r \otimes P_r)(\tau) \end{aligned} \tag{2.21}$$

where $\otimes$ represents the correlation of two power signals [79] according to

$$(P \otimes r)(\tau) := \lim_{\Delta\to\infty} \frac{1}{\Delta} \int\limits_{-\Delta/2}^{+\Delta/2} P(t+\tau) \cdot r(t) dt$$

and G corresponds to the generation efficiency defined as

$$G := \frac{R_{ph} T}{q}, \tag{2.22}$$

describing the number of electrons generated by an incident optical power of 1W. The difference of the switch control functions is abbreviated by the reference function $r(t)$. The equivalent signal model of this mathematical description is depicted in figure 2.6.

The correlation property of the difference of two samples is exploited for the phase measurement. If the optical signal is sinusoidally modulated, an appropriate reference function is described as

$$r(t) = \mathrm{sign}(\cos(2\pi f_m t)), \tag{2.23}$$

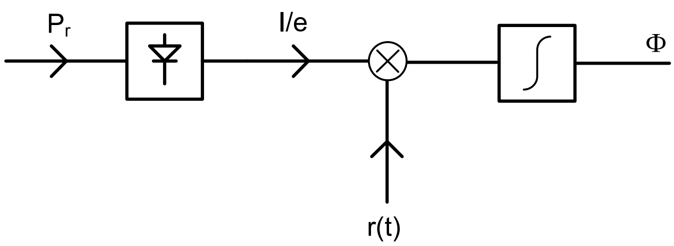


Figure 2.6: Signal flow of the correlation process.

where

$$\operatorname{sign}(x) = \begin{cases} 1 & \text{if } x > 0 \\ 0 & \text{if } x = 0 \\ -1 & \text{if } x < 0 \end{cases}.$$

This reference function is a periodic rectangular function synchronized with the non-delayed version of the modulation signal as shown in figure 2.7. It is obvious that the two switches in the pixel are being alternately activated and the second sample is weighted by -1. The outcome of the difference is the correlation value between the reference signal $r(t)$ and the optical signal $P_r(t,\tau)$, which is

$$\Phi(\tau) = \frac{2}{\pi} G k P_A \cdot \cos(2\pi f_m \tau). \tag{2.24}$$

By taking two samples Φ_0 and Φ_1 of the correlation function, the phase can be calculated. The first sample is the correlation value between $r(t)$ and $P_r(t,\tau)$, and the second one is the correlation value between $P_r(t,\tau)$ and a delayed version of the reference function $r(t - T_m/4)$, as illustrated in figure 2.7. Φ_0 and Φ_1 are expressed as

$$\begin{aligned} \Phi_0 &= \tfrac{2}{\pi} G k P_A \cdot \cos(2\pi f_m \tau) \\ \Phi_1 &= \tfrac{2}{\pi} G k P_A \cdot \sin(2\pi f_m \tau). \end{aligned} \tag{2.25}$$

While $\Phi_0 = A_0 - A_2$, the second correlation value Φ_1 is given by $A_1 - A_3$ because a second correlation process has to be performed. With $\varphi = -2\pi f_m \tau$ the phase is computed by

$$\varphi = \arctan\left[\frac{A_3 - A_1}{A_0 - A_2}\right], \tag{2.26}$$

which is the same result as obtained for the sampling approach in equation 2.20.

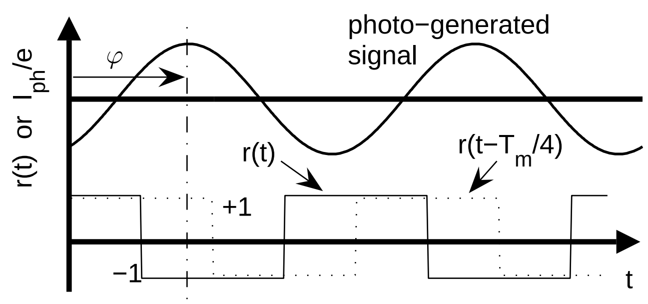


Figure 2.7: Depiction of the photo-generated modulation signal and the reference signals used for the correlation processing.

2.2.4 Distance Resolution

Since the demodulation process is affected by any noise source within the signal chain, the achievable resolution of the measurement is limited. Based on equation 2.7 the distance resolution is estimated in general terms as

$$\sigma_R = \frac{c}{4\pi f_m} \cdot \sigma_\varphi, \tag{2.27}$$

where σ_φ is the standard deviation of the phase measurement. The phase is a function of the samples $A_0 \cdots A_{N_s-1}$ according to equation 2.13, so that the expected error is given by the rules of Gaussian error propagation [21, 60] as

$$\sigma_\varphi = \sqrt{\sum_{g=0}^{N_s-1} \left(\frac{\partial \varphi}{\partial A_g}\right)^2 \mathrm{Var}(\hat{A}_g)}, \tag{2.28}$$

where $\hat{A}_g$ are the noisy sample values.

Contributing noise components: Figure 2.8 depicts a simplified noise model of the demodulation pixels [25]. It distinguishes between the noise sources n_i that are integrated on the capacitances and the noise sources n_a added subsequently.
Integrated noise sources are photon shot noise and dark current shot noise. Added noise sources are independent of the integration time. The most important ones are kTC-noise, injection noise, thermal and flicker noise of the pixel-inherent amplification stages, transfer noise and trapping noise, electro-magnetic-interference and quantization noise.

The detected numbers of electrons $\hat{A}_g$ show a distinct dependency on the incident light intensity as depicted for example in figure 2.9 for generic CCD sensors [38, 106]. The only noise component that depends on the light intensity is photon shot noise [106]. All other noise components act as a constant noise floor for a given integration time. Depending on the incident light intensity, the noise floor or the shot noise dominates the measurement.

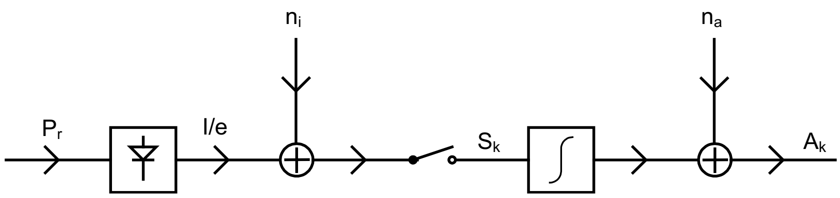


Figure 2.8: Simplified noise model.

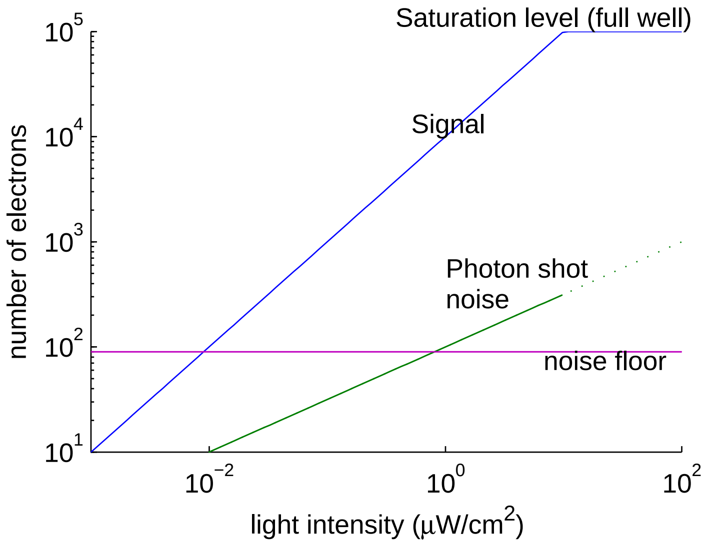


Figure 2.9: Illustration of the different noise sources of a CCD imager, in terms of number of electrons and as a function of the incident light intensity [38].

Shot noise makes up the ultimate physical limitation of the devices because it is the only noise source whose effect on the distance measurement cannot be mitigated at all. Reset noise comprising kTC-noise and injection noise by the reset switch, can be eliminated by correlated double sampling [106]. Thermal noise, flicker noises and dark current shot noise can be reduced by optimizing the pixel's design and operating the devices at low temperatures. The same applies to transfer and trapping noise.
To simplify the calculation of equation 2.28, first the distance measurement is assumed to be limited only by photon shot noise before other noise components are added subsequently as an additional offset. There photon-shot-noise limited measurement means that the numbers of integrated electrons are Poisson distributed so that [21]

$$\mathrm{mean}(\hat{A}_k) = \mathrm{Var}(\hat{A}_k). \tag{2.29}$$

Distance resolution with 4 samples: The statistical deviation of the distance measurement with four samples based on equation 2.20 or 2.26 is derived as [60]

$$\sigma_R = \frac{R_0}{2\sqrt{2}\pi} \cdot \frac{\sqrt{B}}{c_m c_d A_{sig}}, \tag{2.30}$$

with $B = A_{sig} + BG + BD + N_{ps}$, $A = c_d A_{sig}$ and $c_d = c_p c_{d,max}$. There, c_m denotes the modulation contrast of the light source, which is not necessarily 1 in practice, meaning that the amplitude is not equal to the offset component. While the integrated noise sources are separately accounted for, all added noise sources are modelled by one general parameter N_{ps} of so-called pseudo-noise electrons. The integrated noise sources consist of the mean numbers of generated electrons by the signal $A_{sig} = A/(c_m c_d)$, the background light BG and the dark current BD.

The achievable distance resolution can be improved by increasing the demodulation contrast, the optical power or the modulation frequency. The optimization of the demodulation contrast is directly related to the particular pixel architecture, and it is limited by the upper bound $c_{d,max}$. It should be noted here that the optical power cannot arbitrarily be increased because eye-safety has to be ensured in most applications. The most important parameter for improving the accuracy is the modulation frequency. For that reason new pixel structures allowing high modulation frequencies are presented in chapter 5.

2.2.5 Systematic Phase Errors

An analysis of the systematic measurement errors is provided in [60]. The investigation of aliasing effects and non-linearities of the system's output chain is based on the four sample-approach of the signal demodulation. The results are briefly summarized below.

Aliasing

If four samples per period are acquired, the relationship between sampling and modulation frequency is $f_s = 4f_m$. In an ideal system the modulation frequency is the highest frequency component, so that no errors due to aliasing are expected. As soon as additional frequency components f_n greater than the Nyquist frequency of $2f_m$ are included in the modulation signal, the phase measurement will be incorrect.

The received optical power, consisting of an arbitrary number of sinusoids with amplitude P_g and frequency that is an integer multiple of the modulation frequency gf_m, can be expressed as

$$P_r(t,\tau) = P_B + kP_A \cdot [1 + \cos(2\pi f_m(t-\tau))] + \sum_{g=2}^{\infty} P_g \cos(2\pi g f_m(t-\tau_g)). \qquad (2.31)$$

Its spectrum consists of discrete frequency components at $f = 0$, $f = f_m$ and integer multiples of f_m. Natural sampling means a convolution of the power signal with a rect function. The ideal sampling process is described by a multiplication with an infinite

Dirac-pulse chain. Thus, the sampled optical signal is expressed as

$$\mathfrak{c}(t) = \left\{ P_r(t, \tau) * \mathrm{rect}\left(\frac{t}{\Delta t}\right)\right\} \cdot \sum_{l=-\infty}^{\infty} \delta\left(t - l\frac{T_m}{4}\right). \tag{2.32}$$

The Fourier transform of the sampled signal has a spectral distribution

$$\mathfrak{C}(f) = \{\Delta t P_r(f) \cdot \mathrm{sinc}(\pi f \Delta t)\} * \frac{4}{T} \sum_{l=-\infty}^{\infty} \delta(f - l \cdot 4f_m). \tag{2.33}$$

The term in the parenthesis describes a sampling of the sinc function because $P_r(f)$ consists only of discrete frequencies according to the superposed harmonics. The second term is an infinite sum of Dirac pulses that leads to a shift of the frequency distribution of the first term by $4f_m$.

As a result, the odd harmonics will superpose with the modulation frequency at $f = f_m$, producing a phase error of the measurement. In contrast, even harmonics will not superpose at $f = f_m$ and hence, they do not affect the phase measurement at all. Furthermore, it is obvious that Δt has a direct impact on the attenuation of the harmonics. Generally, longer durations Δt will sharpen the peak of the sinc function and hence, the harmonics are attenuated more effectively. If $\Delta t = T_m/g$, the g-th harmonic is completely suppressed. Table 2.2 shows simulated phase measurement errors for different modulation signals [60].

Non-linearities

In practice, the analogous transfer characteristics of the pixel's output stage, subsequent amplification stages, analogue to digital conversion etc. are not ideally linear. In a first step, a quadratic behaviour of the output transfer characteristics has been studied in [60], meaning that rather $a + bA_\mu + cA_\mu^2$ is measured instead of A_μ. The measured phase is then

$$\varphi_{meas} = \arctan\left[\frac{a - a + bA_3 - bA_1 + cA_3^2 - cA_1^2}{a - a + bA_0 - bA_2 + cA_0^2 - cA_2^2}\right]. \tag{2.34}$$

With the use of equation 2.19, the equation above can be simplified. The constants and the quadratic parts are removed so that

$$\varphi_{meas} = \arctan\left[\frac{A_3 - A_1}{A_0 - A_2}\right] = \varphi, \tag{2.35}$$

which means that there is no impact of the non-linearity transfer characteristics on the phase measurement.

Table 2.2: Expected phase errors for a 4-tap DFT measurement depending on the harmonics of the modulation signal and the integration time Δt of one sample [60].

Modulation signal	Harmonics	Error in degrees
$\cos(x)$	-	0
$\cos^2(x)$	-	0
	even harmonics	
$\cos(x) + \cos(2x)$	2	0
$\cos(x) + \cos(4x)$	4	0
$\cos(x) + \cos(6x)$	6	0
	odd harmonics	
$\cos(x) + \cos(3x)$, $\Delta t = T_m/2$	3	± 19.5
$\cos(x) + \cos(5x)$, $\Delta t = T_m/2$	5	± 11.5
$\cos(x) + \cos(7x)$, $\Delta t = T_m/2$	7	± 8.2
$\cos(x) + \cos(3x), \Delta t = T_m/3$	3	0
$\cos(x) + \cos(5x), \Delta t = T_m/3$	5	± 11.5
$\cos(x) + \cos(5x), \Delta t = T_m/5$	5	0
square wave with $\Delta t = T_m/2$: $\cos(x) + \frac{\cos(3x)}{3} + \frac{\cos(5x)}{5} + ...$	$3, 5, 7, 9, ...$	± 4.2
square wave with $\Delta t = T_m/3$: $\cos(x) + \frac{\cos(3x)}{3} + \frac{\cos(5x)}{5} + ...$	$3, 5, 7, 9, ...$	± 3.4
triangle with $\Delta t = T_m/2$: $\cos(x) - \frac{\cos(3x)}{9} + \frac{\cos(5x)}{25} - \frac{\cos(7x)}{49} + - ...$	$3, 5, 7, 9, ...$	± 2.6

However, this conclusion is only valid if all four samples use the same physical output channels, which restricts the design of pixels to structures that have only one common output channel. Non-linearity effects become significant in pixel architectures having several output channels for fast parallel read-out. In this case mismatch effects between the output channels can only be reduced by a symmetric design and high process quality so that the characteristics of the different physical channels become similar.

2.3 Power Budget

The time-of-flight distance measurement system is an active measurement system. The emitted optical power determines the amount of photo-generated electrons in the pixel. As a consequence, the number of electrons specifies the achievable distance accuracy as shown in section 2.2.4. The necessary emitted optical power for achieving a specific distance resolution is described by the power budget, as calculated in [60]. The relationship is essential for the dimensioning of the optical time-of-flight measurement system.

Figure 2.10 shows the power budget of the time-of-flight measurement system. A specific number of photo-generated electrons N_{el} requires the detection of N_{ph} photons. The number of electrons and photons are related by the quantum efficiency QE which depends on the optical wavelength λ. We can deduce the energy per sensitive pixel area E_{sense} with the energy of one photon, which is hf/λ. h denotes Planck's constant. The fill factor FF, which describes the ratio between the sensitive pixel area A_{sense} and the total pixel area A_{pix}, and the integration time T allow the calculation of the optical power P_{pix} on the pixel. The optical power on the sensor P_{image} is computed with the ratio between the pixel area A_{pix} and the image area in the sensor plane A_{image}. With the definition of k_{lens} being all losses of the lens and filters, the optical power in front of the lens is calculated. The optical power P_{obj} reflecting from the object surface is related to P_{lens} via the distance R and the lens diameter D. It is assumed that the reflection characteristics at the object's surface corresponds to Lambertian emission, meaning that the same optical intensities are radiated into all directions. With the use of the reflection coefficient ρ, we can calculate the required optical power of the light source P_{ls}.

Two basic relations are derived from the power budget. First, following from the assumption that the object's surface acts as a Lambertian reflector, the relation between the optical power density on the pixel denoted by P'_{pix} and the optical power density

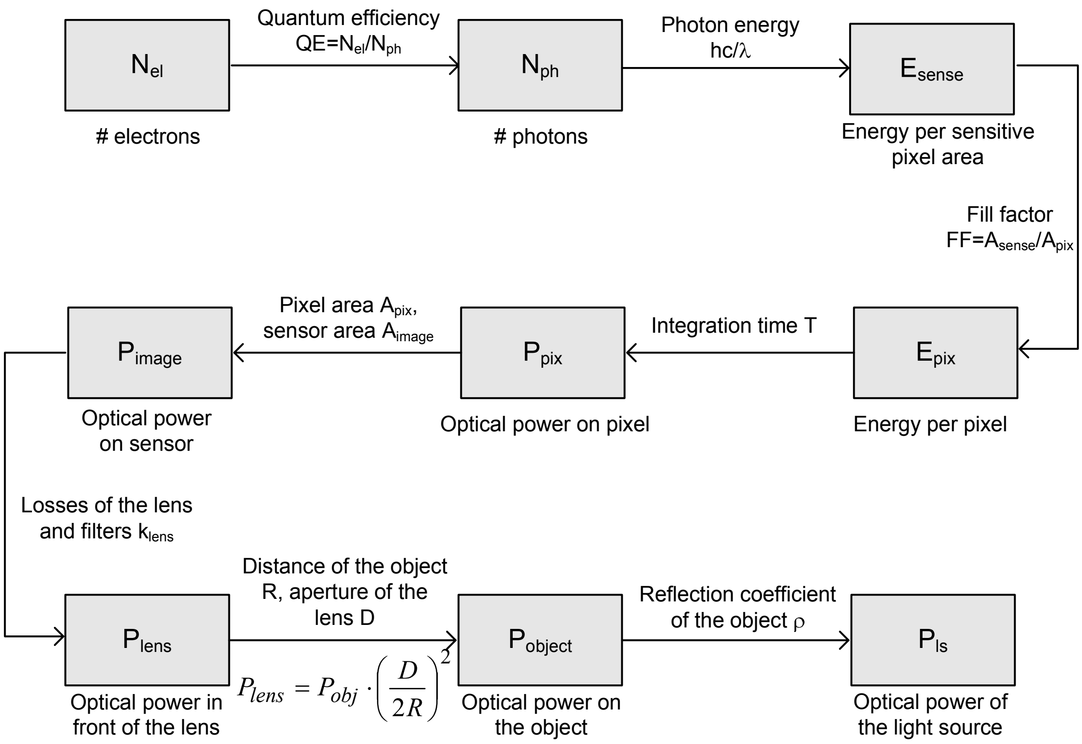


Figure 2.10: Optical power budget.

on the object P'_{obj} is derived as [60]

$$P'_{obj} = \frac{(2F/\#)^2}{k_{lens}} \cdot P'_{pix}, \tag{2.36}$$

with $F/\#$ being the F-number or speed of the optics. Secondly, the number of electrons is related to the optical power of the light source according to [60]

$$\begin{aligned} P_{ls} &= \frac{P'_{obj} \cdot A_{image}}{\rho\left(\frac{l_f}{R}\right)^2} \\ &= \frac{N_{el} hc (2F/\#)^2 A_{image}}{\rho k_{lens} \left(\frac{l_f}{R}\right)^2 QE(\lambda) \cdot \lambda \cdot T \cdot FF \cdot A_{pix}}, \end{aligned} \tag{2.37}$$

or the conversion to the number of electrons gives

$$N_{el} = \frac{P_{ls} \rho k_{lens} \left(\frac{l_f}{R}\right)^2 QE(\lambda) \cdot \lambda \cdot T \cdot FF \cdot A_{pix}}{hc(2F/\#)^2 A_{image}}. \tag{2.38}$$

Equation 2.37 in combination with equation 2.30 allows the estimation of the necessary optical power for a desired distance resolution when a four-tap sampling algorithm is used and a Lambertian reflection at the object's surface is assumed. The reflection coefficient ρ depends on the target's material and surface and it is generally between 0 and 1, when the object is a Lambertian reflector. For directed reflection, for example by retro-reflectors, the reflection coefficient is greater than 1 [60].

2.4 Dynamic Range

Luminance image sensors are characterized by their dynamic range. This parameter is defined as the ratio between the full well level and the noise floor (see also figure 2.9) [106]. The full well level describes the maximum light intensity that can be detected, while the noise floor determines the minimum level of detectable light intensity. Typical values for standard pixels in CMOS or CCD technology are about 60 to 80dB.

However, the intra-frame dynamic range of 3D-TOF sensors is differently determined. Following the extreme environmental conditions in a time-of-flight measurement system as depicted in figure 2.11, the 3D-TOF sensors have to deal with high optical power due to the reflection from close and highly-reflective targets and with low optical power reflected from far and lowly-reflective targets. The specification of a minimum distance

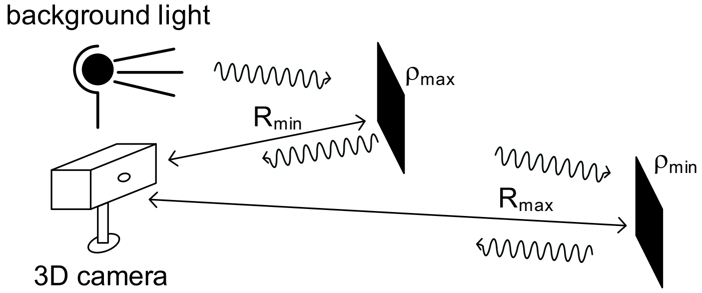


Figure 2.11: Illustration of the extreme optical conditions in the time-of-flight 3D-imaging system.

resolution defines the integration time and optical power of the light source to be used in order to achieve this resolution for the far, lowly reflective objects. Since the integration time is generally a globally set parameter and applies to all pixels, highly reflecting, close objects within the scene still have to be detected without driving the pixels into saturation. However, the full-well level FW of the pixel's integration nodes limit the range of the object's distance detection. Thus, the dynamic range of a 3D-TOF sensor for a particular distance resolution σ_R is determined as

$$DR_{se}[\text{dB}] = 20 \cdot \log_{10}\left[\frac{FW}{N_{el}(\sigma_R)}\right], \tag{2.39}$$

where N_{el} is the minimum amount of electrons that needs to be accumulated in order to reach the specified distance resolution σ_R at maximum distance R_{max} and low reflectivity ρ_{min} of the object. The link between the number of electrons that have to be stored and the mean number of photo-generated electrons per tap is given as

$$N_{el} = (1 + c_d) \cdot A_{sig}, \tag{2.40}$$

where A_{sig} depends on the specified distance resolution and can be calculated with equation 2.30.

The dynamic range of the scene can be highly demanding to the sensor technology depending on the optical conditions with regard to both the signal component and the background light. This is discussed in the next two subsections.

2.4.1 Signal Range

The scene that has to be perceived in its three dimensions sets the demands on the 3D-TOF imager with regard to the dynamic range. The dynamic range of the scene is determined by the ratio between the maximum and minimum optical power received.

Referring to figure 2.11, the maximum optical power reflected by a target is expected for minimal distances R_{min} and highest reflectivities ρ_{max}. On the other hand the minimum power is detected for objects at the maximum distance R_{max} with minimum reflectivity ρ_{min}. Following the equation 2.38 in section 2.3 regarding the power budget, the optical power detected by the sensor and depending on the particular object's distance and reflectivity can be expressed in terms of number of electrons using the same integration time. Thus, the dynamic range of the scene is defined as [23]

$$\begin{aligned} DR_{sc}[\mathrm{dB}] &= 20 \cdot \log_{10}\left[\frac{N_{el,1}(P_{ls},T,R_{min},\rho_{max})}{N_{el,2}(P_{ls},T,R_{max},\rho_{min})}\right] \\ &= 20 \cdot \log_{10}\left[\left(\frac{R_{max}}{R_{min}}\right)^2 \cdot \frac{\rho_{max}}{\rho_{min}}\right], \end{aligned} \tag{2.41}$$

where $N_{el,1}$ and $N_{el,2}$ correspond to the number of electrons generated by the close, highly reflective target and the far, lowly reflective target, respectively.

The dynamic range that can be handled by the 3D-TOF imager is directly related to the desired measurement accuracy (see equation 2.39), which is given by a minimum number of photo-generated electrons. The required full well of a 3D-TOF sensor to capture a scene of a specific dynamic range is calculated by the product of the scene's dynamic range and the minimum number of electrons required for the given distance resolution. The following example demonstrates the high demand on 3D-TOF sensor regarding the dynamic range.

Example: If the imaged scene consists of distances from $20\mathrm{cm}$ up to $5\mathrm{m}$ and the objects show a reflection coefficient from 5% up to 100%, then the demodulation pixels have to deal with a dynamic range of the scene of $DR_{sc} = 82\mathrm{dB}$. This is a factor of about 12500 between the minimal and maximal number of electrons that have to be stored in each sample node [23].
With the assumptions that a modulation frequency of $20\mathrm{MHz}$ is used, the demodulation contrast is 0.4 and no background light is present, each integration node has to be able to store about 620 electrons when a distance resolution of $10\mathrm{cm}$ is required. The dynamic range of $82\mathrm{dB}$, however, means that pixels detecting light that is reflected from close and highly-reflective targets have to deal with a maximum number of 7.75 million electrons. This is remarkably high compared to recently realised full-wells of 1 million electrons covering a dynamic range of just $DR_{se} = 64\mathrm{dB}$ at a distance resolution of $10\mathrm{cm}$.

It is not possible to design the integration stages arbitrarily large because the noise performance would be affected adversely. Alternative methods for the handling of high dynamic 3D-images have to be explored. This is the subject of chapter 5.

2.4.2 Background Light

The second challenging aspect with regard to the dynamic range of the sensor is background light. Though any offset components due to background light do not inherently affect the extraction of the phase at all in a sinusoidally modulated TOF-system due to the fact that the differences of two samples are taken, the photons of the background light are detected anyway and have to be handled by the pixel. Thus, the number of electrons that has to be stored in the pixel is even higher when background light is present.

The received power of background light mainly depends on the background light source, described by the spectral distribution of the power density Υ. As an example figure 2.12 shows the solar irradiance on earth within the wavelength range from $370\mathrm{nm}$ up to $1100\mathrm{nm}$, where silicon imagers are photo-sensitive. The solar irradiance is particularly significant when the 3D-camera system is used in outdoor applications.

The number of electrons additionally generated by background light is computed as

$$N_{el} = \frac{T \cdot FF \cdot k_{lens} \rho A_{pix}}{hc \cdot (2F/\#)^2} \cdot \int_{\lambda_0}^{\lambda_0 + BP} QE(\lambda) \cdot \lambda \cdot \Upsilon(\lambda) d\lambda, \tag{2.42}$$

where BP denotes the optical band-pass filter of the system and λ_0 is the lower boundary wavelength of the bandpass. The dynamic range of the scene is redefined as the ratio of the sum of electrons generated by the signal and the background light, and the minimum number of electrons generated exclusively by the signal components at a

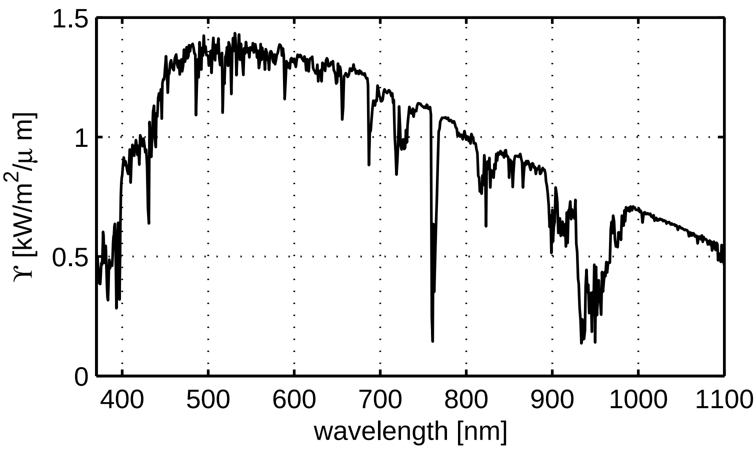


Figure 2.12: Terrestrial spectrum of sunlight. Data set taken from the American Society for Testing and Materials [1,4].

large distance and low reflectivity. The dynamic range is expressed as

$$DR_{sc}[\mathrm{dB}] =$$
$$20 \cdot \log_{10} \left[\frac{\rho_{max}}{\rho_{min}} \cdot \left(\frac{R_{max}}{R_{min}} \right)^2 + \frac{\rho_{max}}{\rho_{min}} \cdot \frac{R_{max}^2 \cdot A_{image}}{P_{ls} \cdot QE(\lambda_{sig}) \cdot \lambda_{sig} \cdot l_f^2} \cdot \int_{\lambda_0}^{\lambda_0 + BP} QE(\lambda)\lambda\Upsilon(\lambda)d\lambda \right], \tag{2.43}$$

where it is assumed that the optical modulation signal uses a single optical wavelength denoted by λ_{sig}.

At this point it is noted that background light not only affects the pixel's saturation level but also decreases the stochastic accuracy according to equation 2.30. Therefore, it is highly desirable to suppress the background light already in the optical domain, which means that fewer photons are detected and the stochastic performance is less influenced. Optical suppression methods include:

- Using narrow band-pass light source and optical filter. *Remark:* The bandwidth defines the spectral width demanded from the light source for the modulation. Its spectral distribution should be within the bandwidth of the band-pass filter in order to exploit most of the optical energy of the light source.
- Reducing the field of view. *Remark:* This implies an increase of the optical signal power density on the object, so that the ratio between the power densities of the signal and the background light increases accordingly. The integration time can be lowered and, thus, fewer electrons will be generated by the background light.
- Operating the system at wavelengths where the background light source shows a reduced power density. *Remark:* In the case of solar irradiance, several local minima are found due to atmospheric absorption. However, the minimum power densities are located either in the region where silicon shows low photosensitivity ($\lambda > 900\mathrm{nm}$) or the light is visible (for example at $\lambda = 760\mathrm{nm}$).
- Burst modulation. *Remark:* Short pulses of optical sinusoidally modulated signals with high power are emitted. While the mean optical power is not increased, the signal-to-background light becomes significantly larger and the integration time can be lowered. This technique is very effective, but the modulation light source needs to support high peak power.

In addition to the optical suppression methods, background light can be electronically suppressed in the pixel as well. A new concept is presented in chapter 5. However, it is obvious that electronic suppression methods only avoid the pixel saturation due to background light but the degrading influence of the background light on the stochastic measurement performance is not avoided.

2.5 Conclusion

Based on a simplified model of the demodulation pixel consisting of an element for the photo-conversion, N_t switches and storages nodes, the time-of-flight principle has been described. Two different theoretical approaches have been pursued: the demodulation process can be regarded either as a sampling process or as a correlation process. Both theories lead to the same calculation of the phase shift between the emitted and received optical signals.

Since sinusoidally-modulated signals can conveniently be generated, this kind of modulation serves as the basic modulation technique for the performance evaluation of the demodulation pixels presented here. The investigation of several error sources, directly linked to the modulation scheme, provide a fundamental understanding of the system's behaviour. The effects of the superposition of harmonic signals, non-linearities in the output transfer characteristics and photon shot noise on the distance measurement have been discussed.

The photon shot noise is of special interest because it shows the achievable physical limitation of a time-of-flight range imaging system in terms of distance accuracy. Only by increasing the modulation/demodulation frequency significantly can even more accurate distance measurements be achieved. This in turn necessitates the development of new pixel structures that are capable of demodulating high-frequency optical signals. Recent pixel structures, forming the core component of the 3D-imaging system are mainly limited by their demodulation speed of some tens of Megahertz.

Further analysis of the optical circumstances under which a time-of-flight range imaging system operates clearly shows the demanding dynamic range requirements of the pixels. Since the system is actively driven, a large number of electrons photo-generated by the modulation signal and any background light has to be stored. Moreover, differences in the scene concerning the object's distance and reflectivity increase the dynamic range remarkably.
The highly demanding operating conditions have been shown by a simple example. It makes clear that particularly designated pixel structures need to be developed in order to operate under strongly varying light conditions. Pixel structures for both the

improvement of the dynamic range of the signal component as well as the active suppression of background light are subject of this work.

Another aspect of the system's restriction when using sinusoidally-modulated signals is the impracticality of concurrent camera operation. So-called multi-camera operation, however, is highly desired for enabling new applications of the time-of-flight range imaging where several cameras are operating in an uncontrolled environment without any synchronisation or spatial separation.

In order to first complete the signal theory portion of this work, the multi-camera operation using pseudo-noise modulation is discussed in the next chapter and a detailed analysis of the achievable performance is provided. The pixel-related challenges of the handling of the high-dynamic signal range, background light suppression and high-speed demodulation and the corresponding physical fundamentals are treated in the chapters 4 and 5.

Chapter 3

Multi-Camera Operation

Since the time-of-flight 3D-camera is an active system, the simultaneous operation of several cameras can lead to interference between the cameras due to the superposition of the optical signals. Figure 3.1 shows a general multi-camera environment where several 3D-cameras are located at different positions and acquire the scenes in their three-dimensional geometry. The cameras are numbered from 1 to N_c. The fields of view of the cameras are partially overlapping. An accurate distance extraction is only possible when each camera unambiguously identifies its own signal with which the phase measurement has to be performed.

In the first section, it is shown that sinusoidal modulation does not allow for measuring the distance when several camera signals are superposed, all using either the same or odd integer multiples of the frequency. Therefore, fundamental concepts for avoiding

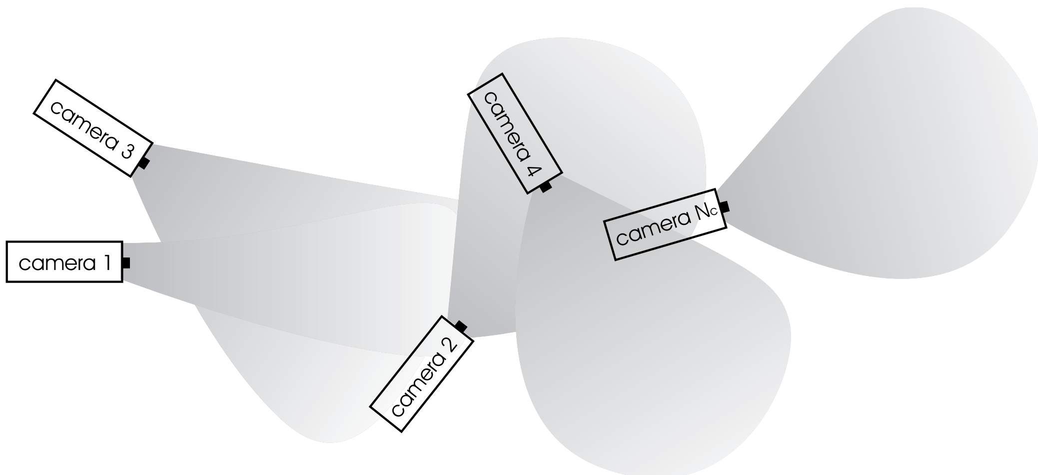


Figure 3.1: Optical interferences in a general multi-camera environment.

measurement errors due to the superposition of several modulation signals are briefly discussed with regard to the multi 3D-camera operation in the second section. These concepts are well-known from mobile communication and optical data transmission.
Since most concepts are directly related to very special environmental conditions or only suitable for particular applications with, for example, a limited number of cameras, in the subsequent chapters the focus lies on an optical spread-spectrum modulation scheme based on pseudo-noise sequences, that provides the same performance for each camera, allows for a random multi-camera set-up and can be adapted to any theoretically arbitrary number of operational cameras. We provide a detailed analysis of the pseudo-noise optical modulation scheme for 3D-imaging with a comparison to the sinusoidal modulation. The analysis covers the effects of several error sources in the distance computation, including the interference by concurrently running cameras, the error due to photon shot noise, the distortion due to non-linear output characteristics, the system's response to background illumination and the effects of low-pass modulation signals and multi-path propagation. We verify the basic theoretical relationships concerning the multi-camera operation and the photon shot noise limitation in a single-camera environment, by performing measurements with an state-of-the-art 3D-camera. The last section discusses the advantages and deficiencies of pseudo-noise optical modulation in a 3D-imaging system. It provides an inspiration for considering other modulation schemes for multi-camera operation.

3.1 Sinusoidal Modulation in a Multi-Camera Environment

The inapplicability of sinusoidally modulated signals of the same frequency and optical wavelength to a multi-camera environment can conveniently be illustrated by the exploitation of complex phasor diagrams. As depicted in figure 3.1, a camera operating in a multi-camera environment detects its own signal and the signals of other cameras, attenuated according to the distance and reflectivity of the observed objects. If the camera's modulation is sinusoidal and each camera uses the same modulation frequency, the received signal of camera 1 is mathematically expressed as

$$\begin{aligned} P_r(t) &= P_{r,1}(t) + \sum_{g=2}^{N_c} k_g P_{A,g} \cdot \cos(2\pi f_m - \varphi_g) \\ &= P_B(t) + P_{A,\Sigma} \cdot \cos(2\pi f_m t - \varphi_\Sigma), \end{aligned} \tag{3.1}$$

where k_g denotes the attenuation of the amplitude $P_{A,g}$ of the g-th camera. The resulting power signal consists of the background light $P_B(t)$ and a sinusoidally modulated signal of amplitude $P_{A,\Sigma}$ and phase φ_Σ. The sinusoid corresponds to a summation of the particular sinusoids of the contributing cameras in a phasor diagram. Figure 3.2 is a general depiction of the phasor addition in the complex plane. While φ_1 is the searched phase corresponding to the correct distance value, the summation of several phasors will produce a phasor within the depicted circle, which has a radius $r_\Sigma = \sum_{g=2}^{N_c} P_{A,g}$ and its origin is at the end of phasor 1. Depending on the number of contributing signals and their amplitudes, an arbitrary phase φ_Σ is generated, resulting in an erroneous distance calculation.

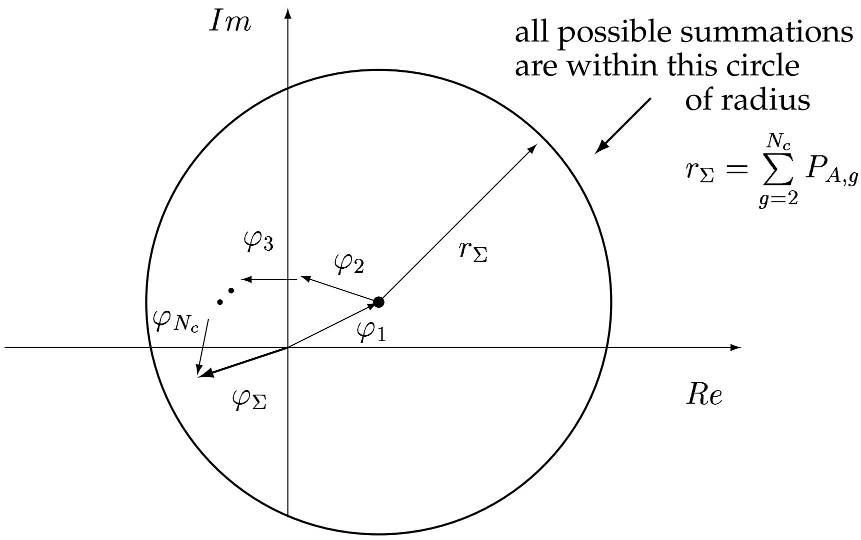


Figure 3.2: Summation of several cosine signals in the complex plane.

3.2 Concepts of Multi-Camera Operation

The superposition of data signals resulting in an incorrect measurement is a well-known problem in mobile communication systems and optical data transmission. Fundamental concepts exist that allow several users to access the same communication channel [32, 86, 88]: space division multiple access (SDMA), time division multiple access (TDMA), frequency division multiple access (FDMA), code division multiple access (CDMA) and wavelength division multiple access(WDMA). These concepts can be adopted for the 3D-camera technology and are briefly discussed in the following paragraphs.

It is obvious that all concepts of multi-camera operation can be combined, similar to the accomplishment in modern communication systems [32]. Such combinations enhance even more the possible number of cameras operating in a multi-camera environment.

3.2.1 Space Division Multiple Access

The overlapping of the illumination lobes of the different cameras is avoided by clearly separating the cameras' operation areas in space. The basic idea of SDMA is illustrated in figure 3.3. All cameras can use the same modulation frequency and optical wavelength. No particular controll of the cameras in time is required. SDMA is useful in all applications where no overlapping of the imaged scenes is expected, for example in security applications where each camera observes one pre-defined area. Applications where the cameras are moving and unexpectedly interfere each other are not properly addressed by the SDMA solution.

3.2.2 Time Division Multiple Access

This technique assigns a time slot to each camera when it may acquire a 3D-image of the scene. Figure 3.4 depicts the basic idea of time division multiple access, where a static assignment of the time slots to the cameras has been performed. Also a dynamic distribution of the time slots could be realised.
The system needs either a master device that accurately controls the acquisition times of the cameras, or a communication between the cameras in order to avoid the simultaneous acquisition of the scene by two or more cameras. TDMA enhances the multi-camera operation to applications where several cameras take images of the same scene so that

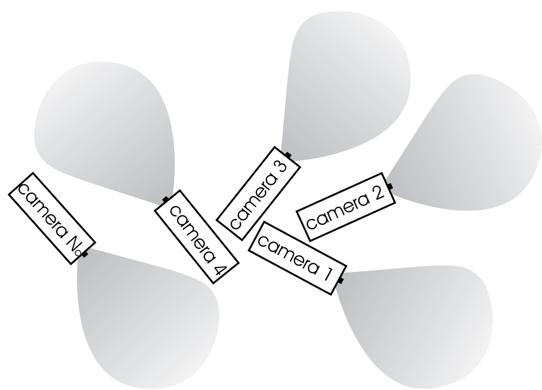


Figure 3.3: Space division multiple access.

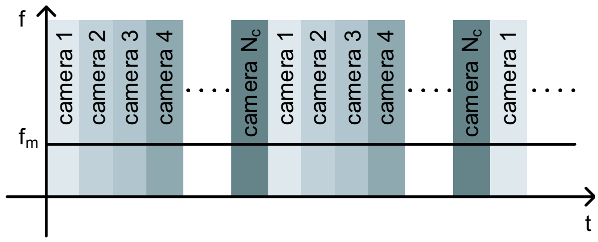


Figure 3.4: Time division multiple access.

the illumination lobes would overlap.
Two restrictions of the TDMA system are noticed: 1. An unexpected approach of a camera that is not synchronised with the TDMA camera system leads to interferences and thus, distance errors are produced. We deduce that the pure TDMA system is exclusively suitable for closed multi-camera environments that use a pre-defined camera set-up and acquisition sequence. 2. The more cameras are operating, the lower the expected frame rate. For a large number of N cameras, the real-time functionality is increasingly lost.

3.2.3 Wavelength Division Multiple Access

By assigning a unique optical wavelength λ to each camera, a robust multi-camera environment can be set up without the need of any synchronisation between the cameras. Figure 3.5 shows the principle of wavelength division multiple access. Each camera obtains 3D-images all the time. The signal of interest is passed by an optical filter and all other signals are blocked. The quality of the optical band-pass filter and of the modulating light sources determines the necessary spacing between the optical wavelengths used.

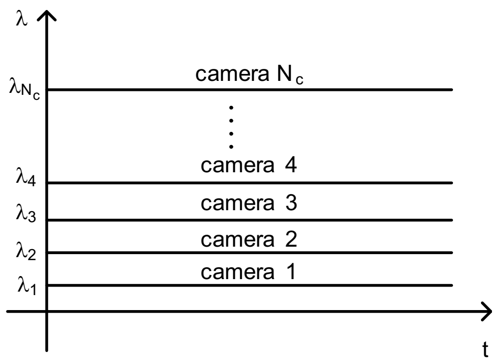


Figure 3.5: Wavelength division multiple access.

The main drawback of this technique is that the cameras will not achieve the same performance for different wavelengths. The light sources behave differently dependent on the wavelength; the sensor's sensitivity is strongly depending on the wavelength as well. Additionally, operating the cameras in the non-visible optical range, which is an important assumption for many applications, limits the number of simultaneously working cameras significantly.

3.2.4 Frequency Division Multiple Access

The available range of modulation frequencies is segmented into a corresponding number of discrete frequencies. Each camera has one modulation frequency f_m assigned, as shown in figure 3.6. This technique allows the operation of several cameras in an arbitrary multi-camera environment without the need of synchronisation between them. However, the modulation frequency directly affects the achievable distance resolution according to equation 2.27, so that a lower bound is given by specification of the particular system's accuracy. A second aspect should be noted: the necessary distance between the modulation frequencies is determined by the shortest integration time. This effect has already been investigated by Lange [60] and is briefly discussed in the following. The present work even extends the investigation by an analysis of the expected phase error in the case of spurious sinusoidal signals.

Frequency Distance Between Two Signals

The longer the integration time, the better is the frequency selectivity of the demodulation process: Instead of demodulating a perfect infinite sinusoid, a time-limited sinusoid is processed. The time limitation means a multiplication of the detected optical

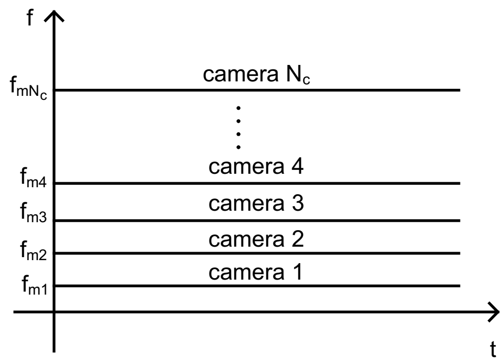


Figure 3.6: Frequency division multiple access.

signal with a rectangular function of width equal to the integration time T. If two cameras are considered, then the detected signal consists of two neighbouring modulation frequencies f_{m1} and f_{m2}. Assuming they have the same amplitude, the detected signal $s(t)$ is expressed in a simplified form as

$$\mathfrak{s}(t) = [\cos(2\pi f_{m1} t) + \cos(2\pi f_{m2} t)] \cdot \mathrm{rect}\left(\frac{t}{T}\right) \tag{3.2}$$

just taking into account the signal components, without any phase shifts and normalized to their maxima. In the frequency domain, this is expressed by a convolution as

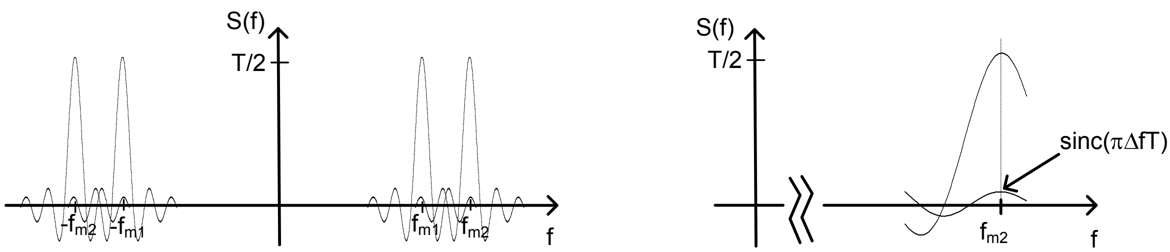
$$\mathfrak{S}(f) = \frac{T}{2} \cdot [\delta(f + f_{m1}) + \delta(f - f_{m1}) + \delta(f + f_{m2}) + \delta(f - f_{m2})] * \mathrm{sinc}(\pi f T). \tag{3.3}$$

This mathematical relation is depicted in figure 3.7. The impact of a second neighbouring modulation frequency is given by $\mathrm{sinc}(\pi \Delta f T)$, where $\Delta f = |f_{m1} - f_{m2}|$.

If we postulate an attenuation of the second frequency by 40dB (factor 100) [60], the necessary minimum integration time T_{min} depends on the difference of the modulation frequencies Δf according to

$$T_{min} = \frac{100}{\Delta f}. \tag{3.4}$$

The basic relationship between the minimum integration time and the distance between two modulation frequencies for the above-mentioned assumption of 40dB suppression of the spurious signals is shown in figure 3.8.



(a) Impact of spurious signals on the signal of interest, depicted in the frequency domain.

(b) Zoom into the frequency spectrum around the second modulation frequency f_{m2}.

Figure 3.7: Frequency selectivity in the presence of spurious signals.

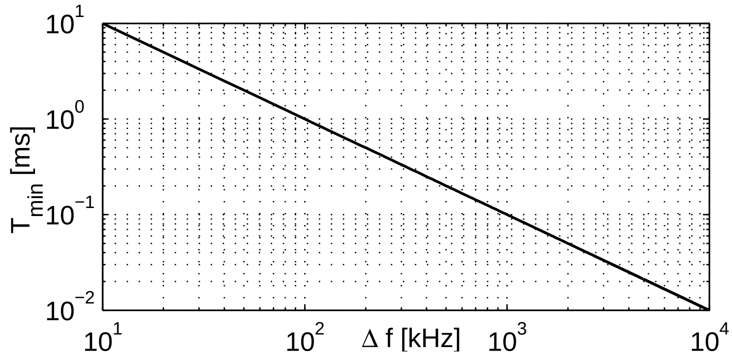
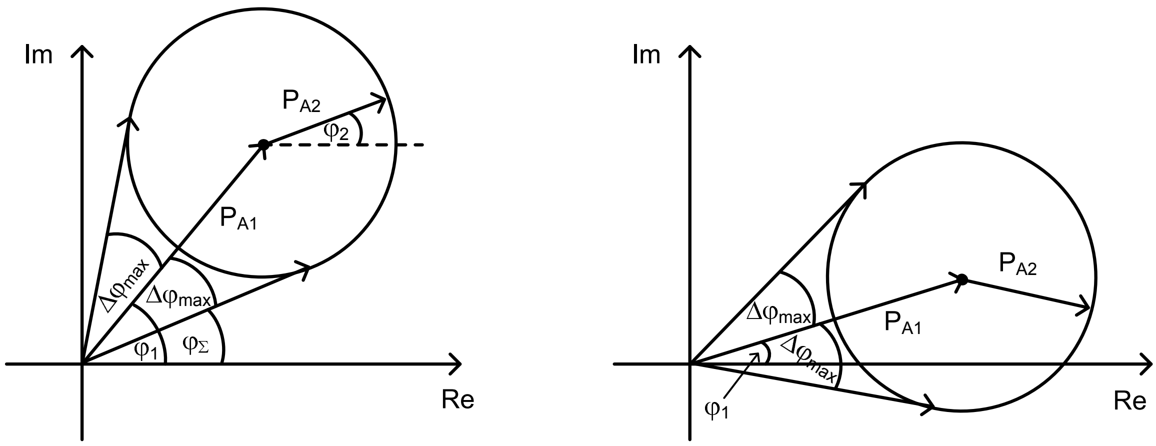


Figure 3.8: Relationship between the necessary minimum integration time and the distance between two modulation frequencies for achieving a suppression of the spurious signal by at least a factor of 100.

Phase Error

The postulation of a 40dB suppression of spurious signals determines the upper bound of the expected maximum phase error. Similar to the considerations in section 3.1, the addition of two sinusoids of the same frequency is illustrated in the complex phasor diagram. Figure 3.9 (a) shows the phasor diagram for two sinusoids with the amplitudes P_{A1} and P_{A2} and the phases φ_1 and φ_2, respectively. Since P_{A2} is attenuated, the phasor addition leads to phasors whose end points lie within the depicted circle. Based



(a) General summation of two sinusoids.

(b) Summation of two sinusoids for small angles φ_1.

Figure 3.9: Depiction of the erroneous phase measurement in the complex phasor diagram when two sinusoids of the same frequency are added together.

on trigonometric relations, the expected maximum phase error is calculated as (see appendix A.1)

$$\Delta\varphi_{max} = \arctan\left[\frac{P_{A2}/P_{A1}}{\sqrt{1-(P_{A2}/P_{A1})^2}}\right]. \quad (3.5)$$

For weak spurious signals ($P_{A2}/P_{A1} \to 0$), equation 3.5 is approximated by

$$\Delta\varphi_{max} \approx \frac{P_{A2}}{P_{A1}}. \quad (3.6)$$

With regard to our postulation of suppressing the spurious signal by a factor of at least 100, the maximum error is expected to be about 0.01 rad. Using a $20\mathrm{MHz}$ modulation signal, this error corresponds to maximum distance error of about $1.2\mathrm{cm}$. The measured erroneous phase φ_Σ is between $\varphi_1 - \Delta\varphi_{max}$ and $\varphi_1 + \Delta\varphi_{max}$.
Another aspect needs to be mentioned. If phase φ_1 is small, the circle's area could cover both a part of the positive and of the negative imaginary region of the complex plane. In this case, the spurious signal potentially even produces phase hops. This situation is illustrated in figure 3.9 (b). The range of the measured phase error covers the positive and negative cases: $0 \leq \varphi_\Sigma < \varphi_1 + \Delta\varphi_{max}$ or $2\pi + \varphi_1 - \Delta\varphi_{max} < \varphi_\Sigma < 2\pi$. The distinction is important due to the modulo-2π problem in the computation of the angles.

It is concluded that a good suppression of spurious signals is required in order to keep the phase errors as small as possible. Long integration times and large distances between the neighbouring modulation frequencies ease this suppression.

3.2.5 Code Division Multiple Access

Another possibility of extracting the camera's own signal of interest is to encode the optical signal by an appropriate coding sequence. By correlating the detected optical signal with the reference signal, only the signal of interest will be extracted, assuming that the modulation signals are orthogonal to each other. Such a correlation process can be performed with the dedicated demodulation pixels by evaluating the difference of two integration nodes, as shown in section 2.2.3. Based on the pixel model, the principle of the CDMA technique is depicted in figure 3.10. The light source of the g-th camera is intensity-modulated by the modulation signal $M_g(t)$ and emitted to the object. The delayed version of the optical signal is detected by the sensor, where it is mixed with the reference signal and integrated over a multiple of the modulation period. The reference signal $r(t)$ is equal to the modulation signal $M_g(t)$.

In the presence of N_c cameras that illuminate the same scene, the g-th camera detects the sum of all modulated light signals. Accordingly to equation 2.21, the correlation

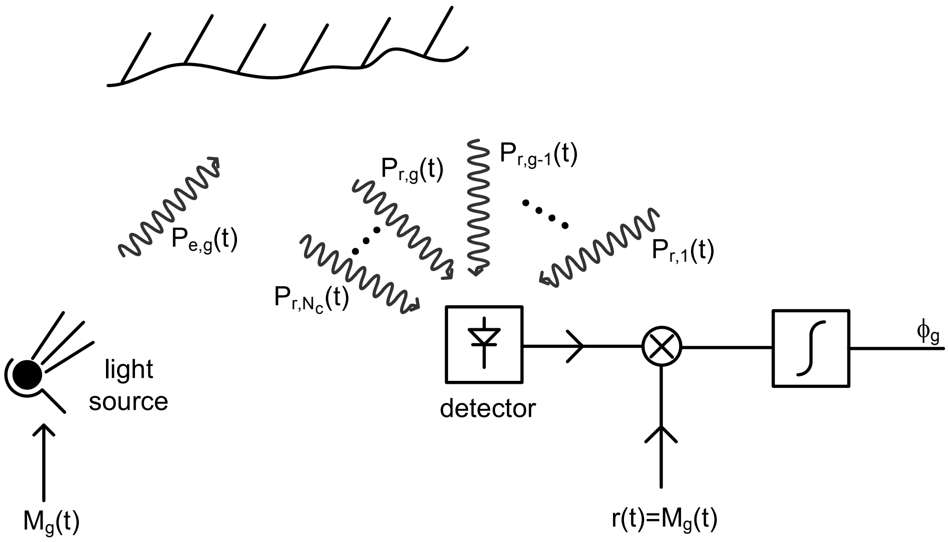


Figure 3.10: Code division multiple access.

value Φ_g is expressed as

$$\begin{aligned} \Phi_g &= G \cdot (r \otimes P_r)(\tau) && (3.7) \\ &= G \cdot (M_g \otimes \sum_{l=1}^{N_c} P_{r,l})(\tau). && (3.8) \end{aligned}$$

With the assumption of orthogonal modulation signals

$$M_g(\tau) \otimes P_{r,l}(\tau) = 0 \qquad (3.9)$$

for $g \in [1; N_c]$ and $g \neq l$, the correlation value Φ_g is only affected by its own signal components M_g and $P_{r,g}$, and it can be expressed as

$$\Phi_g = G \cdot (M_g \otimes P_{r,g})(\tau). \qquad (3.10)$$

Spread Spectrum Property

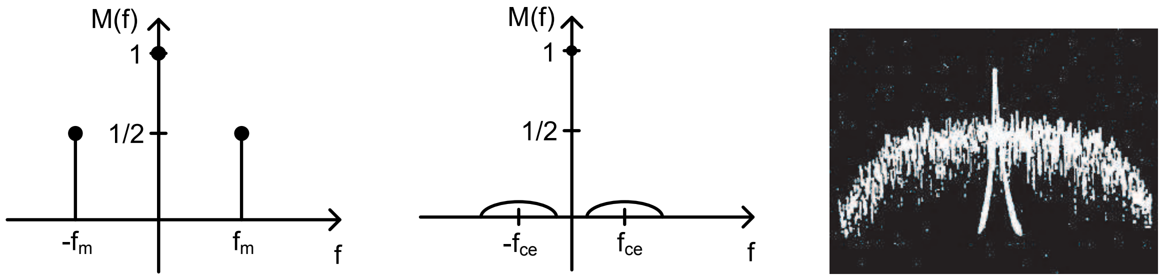
These encoding techniques have the common property of spreading the emitted optical power over a rather wide modulation frequency range instead of concentrating the power in just one modulation frequency. That is the reason why these techniques are called spread-spectrum techniques. An introduction to spread spectrum techniques is provided in [33]. The general behaviour of spreading the optical signal spectrum is depicted in figure 3.11 [32]. Figure (a) shows the spectrum of the modulation signal when continuous wave modulation is applied, so that $M(t) = 1 + \cos(2\pi f_m t)$; subfigure (b) schematises the spectral distribution when a spread spectrum technique is applied, so that $M(t) = 1 + ss(t)$ with $ss(t)$ being any spread spectrum function. For the most common forms of spread spectrum signals, the spectral distribution corresponds to a $\mathrm{sinc}^2(x)$ characteristic [33] around the center frequency f_{ce}. Figure 3.11 (c) shows a direct comparison of measured spectral distributions of continuous-wave and spread spectrum modulations in a communication system with a code rate of $1\mathrm{Mbit/s}$ [33].

The spreading of the spectrum of the optical modulation signal has one important drawback: the electronic components (driver electronics, light source etc.) need to support a larger bandwidth than in the pure continuous wave mode where the electronic components can be optimized to the particular modulation frequency. Errors of the distance measurement are directly generated by any filtering processes of the spread spectrum, which is inevitable to a certain degree.

Distance Evaluation

Since the spectral distribution of a signal is directly linked to its correlation function by the Wiener-Khintchine theorem [79], the autocorrelation functions of a spread spectrum modulation signal have a significant peak characteristic, as depicted for example in figure 3.12. The sampling of the autocorrelation's peak allows the computation of the time-of-flight by applying the inverse function Φ^{-1} to the sampled correlation value $\Phi(\tau)$ where Φ is unambiguously invertible. Thus, the distance is calculated as

$$R = \frac{-c \cdot \Phi^{-1}(\Phi(\tau))}{2}. \tag{3.11}$$



(a) Amplitude spectrum of the continuous wave modulation signal.

(b) Spectrum of a modulation signal that spreads the power over a wide frequency range around the center frequency f_{ce}.

(c) Main lobe of the direct sequence signal compared to equal power continuous wave signal in frequency space [33].

Figure 3.11: Spread spectrum technique.

Figure 3.12: Typical behaviour of the periodic autocorrelation function (ACF) of a spread spectrum signal, normalized to its maximum.

Since the total number of generated electrons, which corresponds to the correlation energy, is generally unknown, the unambiguous determination of the time-of-flight necessitates the acquisition of at least two correlation values within the peak. The more values of the autocorrelation function are sampled, the less the reconstruction of the function is affected by any noise sources. However, the requirement of real-time capability limits the system to taking as few samples as possible for one measurement.

CDMA Methods

Four types of spread spectrum techniques are distinguished [33]:

- Direct sequence (DS-CDMA): A code sequence, also called spreading sequence, [32] is used for the intensity modulation. Orthogonal codes allow for multi-camera operation.
- Frequency hopping (FH-CDMA) [32]: The optical intensity is continuously modulated. However, the frequency is permanently changed following a strict hopping rule. If orthogonal spreading sequences are used as hopping sequence, multi-camera operation becomes possible.
- Time hopping (TH-CDMA): Short optical bursts are emitted, while the emission duration and the pauses between are controlled via a hopping function similar to that in FH-CDMA.
- Chirp: This technique is the only CDMA technique that does not use spreading code sequences. By continuously sweeping the modulation frequency, so-called chirp pulses are generated. Special sweep functions have been created in the field

of communication theory, allowing for multi-user access of the same communication channel. Such sweep functions are discussed for example in [35, 47, 53] and could be adopted to the multi-camera time-of-flight measurement.

The implementation of chirps needs additional analogue components for the generation of monotonous frequency sweeps such as voltage-controlled oscillators (VCO) or phase-locked loops (PLL), while spreading sequences can easily be realized with digital electronics such as linear shift registers. A very convenient way of implementation and distance extraction is accomplished by the DS-CDMA technique. The autocorrelation functions have well-defined characteristics that allow for a straightforward extraction of the time-of-flight. In the next section, we focus on the DS-CDMA technique based on pseudo-noise spreading sequences.

3.3 Binary Pseudo Noise Optical Modulation for Multi-Camera Operation

In this section binary pseudo-noise optical modulation is presented for use in 3D-imaging with the DS-CDMA technique in a multi-camera environment. A detailed performance evaluation of this modulation technique is provided in the next section regarding different error sources and direct comparison with the sinusoidal modulation scheme. The performance is measured with existing demodulation pixels that provide two integration values, corresponding to one correlation value per acquisition. The outline of this section is the following: First a summary of the fundamentals of pseudo-noise, meaning the criteria for randomness, the terminology conventions and the classification of PN-sequences, is given. Then the properties of m-sequences are treated, where we pay special attention to the elaboration of the auto- and cross-correlation properties. We introduce a general system-model that acts as the basic tool for the mathematical description of the distance extraction and the relationship between the amount of the totally photo-generated electrons and the integration values.

Why *binary* modulation sequences? With respect to the pixel model in section 2.2.1, the controlling of the in-pixel switches corresponds to the multiplication of the incoming photo-generated current by a unipolar binary sequence. Further, the subtraction of two in-pixel integration nodes enables the determination of the correlation value between the photo-current signal and a bipolar binary reference signal. Thus, binary sequences are well suited for the demodulation pixels.

A general definition of a pseudo random sequence is found in [124]:

A deterministic signal shall be referred to as a pseudo random signal if the degree of similarity of its properties in terms of particular characteristics come sufficiently close to the characteristics of a random signal.

This definition implies an orthogonality of the signals, which in turn allows for applying simulative deterministic random sequences as a modulation signal in 3D-TOF-imaging.

3.3.1 Pseudo Noise Sequences: Randomness, Terminology and Classification

This subsection gives an overview of the criteria of pseudo randomness, the terminology and the classification of pseudo random sequences.

Criteria for Pseudo Randomness

Deterministic binary sequences are considered as pseudo-random when they fulfill at least one of the following three criteria [124]:

Balance criterion: The difference between the numbers of ones and zeros within one period of the pseudo random sequence is less than or equal to 1. If $H(x)$ is the frequency of occurrence of x, the balance criterion is described for unipolar sequences as

$$|H(0) - H(+1)| \leq 1. \tag{3.12}$$

Run criterion: A run refers to a sequence of only ones or zeros. The occurrence frequency of runs with the same length is the same for both zeros and ones

$$H(0) = H(1), \quad H(00) = H(11), \quad H(000) = H(111), ..., \tag{3.13}$$

where $H(x)$ denotes the occurrence frequency of x. One-half the runs are of length 1, one-fourth are of length 2, one-eighth are of length 3, and so on.

Correlation criterion: The periodic autocorrelation function shows properties similar to those found in true white noise signals, so that

$$\Phi(t) = \begin{cases} 1 & \text{if } t = 0 \\ a & \text{if } t \neq 0 \end{cases}, \tag{3.14}$$

where a is required to be a small constant.

Terminology

Two names are found in the literature in association with sequences that have noise-like properties: pseudo-random and pseudo-noise sequences. In the past, a strict distinction was made between both of them. Pseudo-random sequences defined all sequences which fulfilled at least one of the above-mentioned criteria, and pseudo-noise sequences had to fulfill all three criteria. However, today it is well accepted in the scientific world to use both names in a "generic sense covering the many classes of sequences with pseudo randomness properties" [124]. Therefore, in the remainder of this work pseudo random sequence and pseudo noise sequence are used in the same synonymous generic sense.

Another convention is the use of the word "chip" as a designator for one bit of the binary PN-sequence. This is a common terminology in communication systems and, thus, it is used in this work as well.

Classification of Pseudo Noise Sequences

Figure 3.13 gives an overview of the diversity of known binary random sequences and their combinations. All of them have their advantages and disadvantages and, hence, they are particularly suited to some special applications. The optimization of the sequences is performed in view of obtaining either good autocorrelation or good cross-correlation properties. It is not possible to realize sequences having both good auto- and cross-correlation properties [124].
Maximal-length sequences have, for example, good autocorrelation properties but the

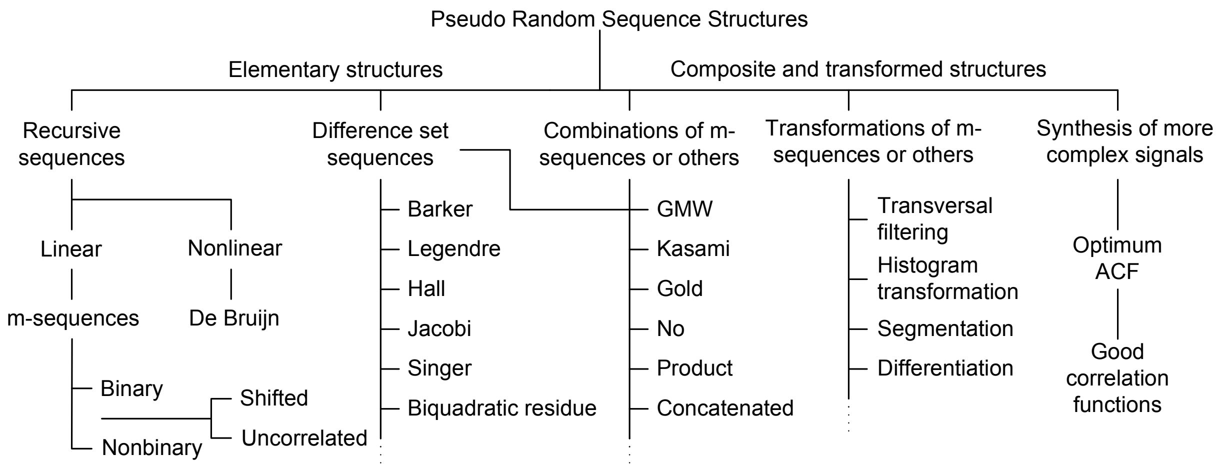


Figure 3.13: Classification of selected pseudo random sequences. [124]

cross-correlation values are not negligible. Gold sequences, in turn, have optimised cross-correlations, but the autocorrelation function is unsatisfactory for the 3D-imaging application.

No further description of all PN-sequences itemized in figure 3.13 is given here. The interested reader is referred to [124]. We selected for a further in-depth investigation PN sequences of maximal length (m-sequences) due to their excellent autocorrelation property, which is a necessary requirement for the convenient distance extraction in the TOF system, and due to the easy and straightforward hardware implementation.

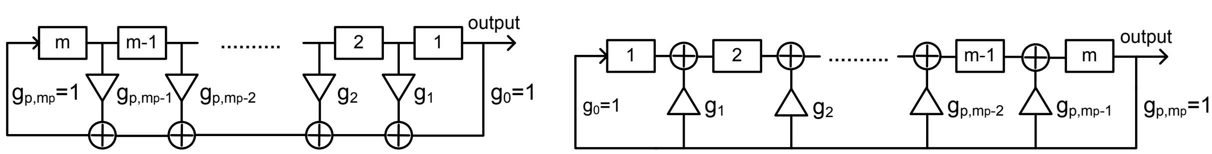
3.3.2 Properties of Maximal-Length PN-Sequences

m-sequences are generated with linear feedback shift registers (LFSR), as depicted in figure 3.14. Two implementations are known: Fibonacci and Galois implementation. [50] Both of them produce the same binary sequences when their feedback taps are of reversed nature. The Fibonacci approach has been used for the tests in the multi-camera environment due to the possible exploitation of standard blocks for the implementation in the programmable logic device.
If appropriate feedback combinations are used, PN sequences of maximal length

$$n_p = 2^{m_p} - 1 \tag{3.15}$$

are generated. Basically, the initial bit assignment of the registers can be arbitrary with the exception of the zero bit assignment.
Each possible m_p-bit combination, except for the m_p-zero combination, occurs within the m-sequence when a window of width m_p is slid over the sequence. This is illustrated in figure 3.15 for a PN-sequence of length 7.

The PN sequence produced fulfills all randomness criteria presented in section 3.3.1. Since the m_p-zero combination does not occur in the sequence, the number of ones is larger by one bit than the number of zeros. This unbalance property is an important



(a) Fibonacci implementation. **(b)** Galois implementation.

Figure 3.14: Linear feedback shift register (LFSR). [50]

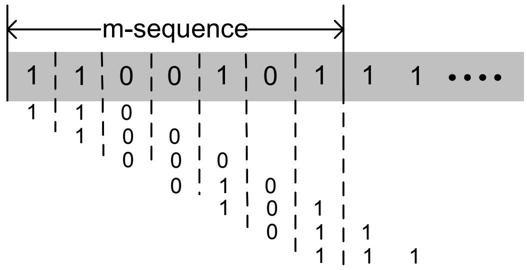


Figure 3.15: Example of a PN sequence that has been generated with a 3-stage LFSR. All possible 3-bit combinations occur as demonstrated.

issue for the performance evaluation of background light and the stochastic accuracy, as shown in section 3.4 later on .

Galois Field Mathematics

The fundamental mathematics for the description of pseudo-noise sequences is based on the theory of Galois fields [85, 124]. More precisely, if binary sequences are treated, Galois fields with basis 2 are exploited.

The feedback of the LFSR is described by the so-called generator polynomial of degree m_p. The general form of a generator polynomial is expressed as

$$GP(x) = g_{p,m_p}X^{m_p} + g_{p,m_p-1}X^{m_p-1} + g_{p,m_p-2}X^{m_p-2} + ... + g_{p,2}X^2 + g_{p,1}X^1 + g_{p,0}, \quad (3.16)$$

where $g_{p,0}...g_{p,m_p}$ are the polynomial coefficients restricted to the range of the modulo calculation with the basis of the Galois field. The generation of maximal sequences requires primitive generator polynomials. Such polynomials are irreducible polynomials of degree m_p whose order is at the same time $n_p = 2^{m_p} - 1$. The order describes the smallest degree of a polynomial of which the generator polynomial is a divider. The irreducible polynomials are mathematically described as the roots of $X^{n_p} + 1$. There are often roots of degree less than m_p. Recapitulating, the generator polynomial is of degree m_p and is a divider of $X^{(2^{m_p}-1)} + 1$ [124]. Hence, only a small fraction of all irreducible polynomials are of the order m_p and, hence, are primitive polynomials.
By calculating an arbitrary power of the generator polynomial, it is possible to generate all elements of the polynomial Galois field extension. This processing is performed by linear feedback shift registers and that is the reason why maximal length sequences are generated when the feedback is determined by a primitive polynomial.
The generation of binary m-sequences is described by polynomials of the Galois field GF(2). It means that all computations are limited to a class ring determined by modulo 2 calculations.

Example: A LFSR with $m_p = 3$ stages is used for the generation of m-sequences of length 7. The factorization of $X^7 + 1$ leads to three irreducible polynomials, according to

$$X^7 + 1 = (X+1) \cdot (X^3 + X + 1) \cdot (X^3 + X^2 + 1). \tag{3.17}$$

However, only the second and third polynomial factors are primitive because they are of the degree $m_p = 3$.

Tables of the feedback taps determining the primitive polynomials can be found for example in [50, 124, 127, 128]. The total number of primitive polynomials M for a given register length m_p is calculated as [124]

$$M_p(m_p) = \frac{\Phi_{Euler}(2^{m_p} - 1)}{m_p}, \tag{3.18}$$

where Φ_{Euler} is the Euler's totient. A convenient method for calculating the Euler's totient is

$$\Phi_{Euler}(2^{m_p} - 1) = (2^{m_p} - 1) \cdot \left(1 - \frac{1}{p_1}\right) \cdot \left(1 - \frac{1}{p_2}\right) \cdot \dots \cdot \left(1 - \frac{1}{p_k}\right), \tag{3.19}$$

where $p_1, p_2, ..., p_k$ are the prime factors of $2^{m_p} - 1$.

Example: A LFSR with 15 register stages can be used to generate 1800 different m-sequences. *Proof:* $2^{15} - 1 = 32767$ is expressed by its prime factors as $7 \cdot 31 \cdot 151$. The Euler's totient gives

$$\Phi_{Euler}(32767) = 7 \cdot 31 \cdot 151 \cdot \left(1 - \frac{1}{7}\right) \cdot \left(1 - \frac{1}{31}\right) \cdot \left(1 - \frac{1}{151}\right) = 27000,$$

and the number of available m-sequences is

$$M_p(15) = \frac{27000}{15} = 1800.$$

Autocorrelation Function

The correlation function between the emitted optical modulation signal and the reference signal of the pixel-inherent switches is interesting for the distance extraction. The autocorrelation function is related to the correlation between the unipolar and bipolar versions of the modulation signal $M_u(t)$ and $M_b(t)$, respectively. Figure 3.16 illustrates the characteristics of the pure modulation signals and the result of their periodic correlation. The significant peak of the ACF is mathematically described as

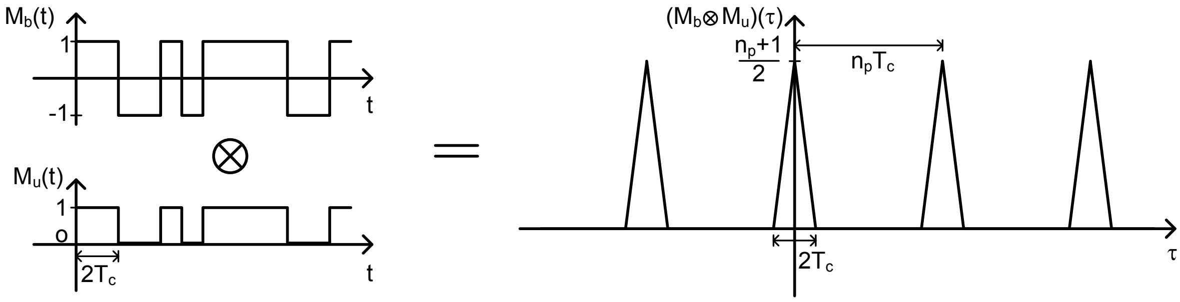


Figure 3.16: Periodic autocorrelation function of m-sequences, when the correlation corresponds to the processing of the unipolar and bipolar versions of the m-sequence.

$$(M_u \otimes M_b)(\tau) = \begin{cases} \frac{n_p+1}{2}\left(1 - \frac{|\tau - n_p g T_c|}{T_c}\right) & \text{for } \epsilon_- < \tau < \epsilon_+ \\ \\ 0 & \text{otherwise,} \end{cases} \tag{3.20}$$

with the bounding parameters $\epsilon_+ = T_c(n_p g + 1)$, $\epsilon_- = T_c(n_p g - 1)$ and $g \in \mathbf{Z}$. T_c denotes the duration of one chip. The width of the peak is $2T_c$ and the distance between two peaks is $n_p T_c$.

Since the autocorrelation function is independent of the particular PN sequence, a sampling of this peak allows for unambiguously extracting the distance with a generic formula, which is described below.

Cross-correlation

The cross-correlation properties of the PN sequences used determine the capability of suppressing the optical signals of concurrently running cameras. Unfortunately, the cross-correlation properties of m-sequences are not as good as the autocorrelation properties. Other pseudo noise sequences such as Gold sequences or orthogonal Hadamard-Walsh codes [32] are optimized with respect to the cross-correlation values. However, with increasing length, the cross-correlation properties of m-sequences and, hence, their applicability in multi-camera environments improve.

Table 3.1 shows the dependencies of m-sequences between LFSR stage number m_p, the corresponding PN sequence length $n_p = 2^{m_p} - 1$, the cross-correlation properties and the available number of m-sequences M_p for the given LFSR length.

The increasing length of the PN-sequences for larger LFSR has two drawbacks. The first one is that the minimum integration time becomes longer because it is determined by one period of the sequence. Also, the step size for longer integration times is de-

Table 3.1: Properties of m-sequences.

m_p LFSR stages	PN length n_p	Θ_{max}	Θ_{prob}	Θ_{avg}	nr. of m-sequ. M_p
3	7	0.5	$7.14 \cdot 10^{-2}$	$1.4 \cdot 10^{-1}$	2
4	15	0.5	$3.33 \cdot 10^{-2}$	$6.7 \cdot 10^{-2}$	2
5	31	0.375	$5.38 \cdot 10^{-3}$	$3.2 \cdot 10^{-2}$	6
6	63	0.375	$5.29 \cdot 10^{-3}$	$1.6 \cdot 10^{-2}$	6
7	127	0.313	$2.62 \cdot 10^{-3}$	$7.9 \cdot 10^{-3}$	18
8	255	0.375	$2.45 \cdot 10^{-4}$	$3.9 \cdot 10^{-3}$	16
9	511	0.219	$2.45 \cdot 10^{-4}$	$2.0 \cdot 10^{-3}$	48
10	1023	0.375	$1.63 \cdot 10^{-5}$	$9.8 \cdot 10^{-4}$	60
11	2047	0.141	$3.89 \cdot 10^{-5}$	$4.9 \cdot 10^{-4}$	176
12	4095	0.344	$3.39 \cdot 10^{-6}$	$2.4 \cdot 10^{-4}$	144
13	8191	0.086	$2.71 \cdot 10^{-6}$	$1.2 \cdot 10^{-4}$	630
14	16383	0.344	$8.07 \cdot 10^{-8}$	$6.0 \cdot 10^{-5}$	756
15	32767	0.150	$3.39 \cdot 10^{-8}$	$3.0 \cdot 10^{-5}$	1800

fined by the duration of one period in order to fulfill the mathematical definition of the correlation process to integrate over an integer number of modulation periods. The second drawback means a higher influence of the variation of the background light because longer runs of zeros or ones occur during longer PN sequences. The aspect of the integration time can generally be said to be negligible for sequence lengths up to 32767 chips. For example, a PN sequence of 32767 chips with a chip duration of $T_c = 50\text{ns}$ would have the smallest step size of the integration time of $\backsim 1.7\text{ms}$, which is acceptable in most applications.
The advantages of using long m-sequences are shown by the columns 3 to 6 in table 3.1. All three parameters quantify the cross-correlation performance of the m-sequences in a certain way. The first parameter Θ_{max} defines the ratio between the maximum of all possible cross-correlation values $\Phi_c(\tau)$ of all possible sequence combinations and the maximum peak value of the autocorrelation function $\Phi_a(0)$, so that

$$\Theta_{max} := \max\left\{\frac{\Phi_c(\tau)}{\Phi_a(0)}\right\}. \tag{3.21}$$

Although there is a lack of strict mathematical theory describing the behaviour of Θ_{avg}, it seems to tend to zero for long m-sequences. Basically, the use of an odd number of shift register stages is preferred because the maximum values of these m-sequences are lower.
In addition, the probability of occurrence of the maximum cross-correlation value strongly decreases for longer m-sequences. This is described by the parameter Θ_{prob}, which is the ratio between the number of occurrences of the maximum cross-correlation value $H(max\{\Phi_c(\tau)\})$ among m-sequences of the same set and the total number of possible discrete cross-correlation values

$$\Theta_{prob} := \frac{H(max\{\Phi_c(\tau)\})}{n_p \cdot M_p}. \tag{3.22}$$

The third parameter Θ_{avg} is often used in the performance evaluation of communication systems [124] and may be the most meaningful parameter. It is the expectation value of the ratio between the cross-correlation values and the peak value of the autocorrelation, which is, in mathematical notation

$$\Theta_{avg} := \text{mean}\left\{\frac{\Phi_c(\tau)}{\Phi_a(0)}\right\}. \tag{3.23}$$

The simulation of the expectation value agrees well with theory, which says that it behaves inversely proportional to the sequence length (see appendix A.2 and [87])

$$\Theta_{avg} = \frac{1}{n_p}. \tag{3.24}$$

From this it is concluded that the expectation of signal distortion due to additional cameras is low. However, a more system-specific simulation and analysis is given in the following subsections.

The size of the available set of m-sequences (6-th column) for a given LFSR is calculated by equation 3.18 or is found in literature [70, 87]. It shows that for long shift registers a large set of sequences is available that all perform in a similar way. Although it is unlikely that several hundred cameras are operated simultaneously, such a demodulation method allows for distinguishing an arbitrarily high number of cameras. The probability that two cameras use the same PN sequence is very low.

3.3.3 General System-Model Based on PN-Modulation

In this subsection we introduce a general system-model that describes in a mathematical sense the relationships between the mean number of photo-generated electrons per tap, the integration values of each tap and the correlation value, which is the difference of two taps.

Time-of-flight distance measurement based on PN optical intensity-modulation uses the correlation approach as described in section 2.2.3. The pixel model of figure 2.4 is reduced to two integration nodes, which are denoted in this section by the generic expressions A_μ and A_ν. As shown in the same section, the pixel-inherent switch can be replaced by a multiplication stage whose inputs are the photo-generated current and the reference function. This reference function $r(t)$ is the periodic bipolar m-sequence

$$r(t) = M_b(t). \tag{3.25}$$

The optical signal's relationships are shown in figure 3.17 and can mathematically be expressed in a similar way to that performed for sinusoidally modulated signals. While the emitted signal is described as

$$P_e(t) = P_{PN} \cdot M_u(t), \tag{3.26}$$

with the unipolar version of the periodic PN sequence $M_u(t)$ and the peak-to-peak power P_{PN}, the received signal is expressed as

$$P_r(t) = P_B + kP_{PN} \cdot M_u(t - \tau), \tag{3.27}$$

where P_B denotes the background light, k describes all attenuations in the optical path and τ is the time-of-flight.

The read-out charge packets A_μ and A_ν of the two integration nodes depend on the delay τ which is sought by the distance measurement. Their general time-characteristic

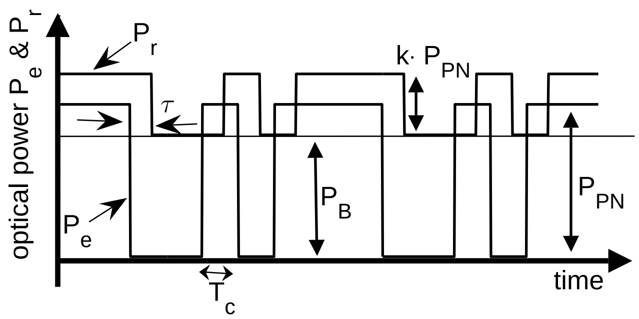


Figure 3.17: Delay measurement with pseudo-noise intensity-modulated optical signals.

is shown in figure 3.18 and can be described by the following two equations

$$A_\mu(t) = \begin{cases} 2A_{sig}\cdot\left(1-\frac{|t|}{2T_c}\right) - A_{sig}\cdot\left(1-\frac{|t|}{T_c}\right)\cdot(1-c_d) \\ \qquad\qquad +BG\cdot\frac{n_p+c_d}{n_p} & \text{for } -T_c < t < T_c \\ A_{sig} + BG\cdot\frac{n+c_d}{n} & \text{else,} \end{cases}$$

$$A_\nu(t) = \begin{cases} A_{sig}\cdot\frac{|t|}{T_c} + A_{sig}\cdot(1-c_d)\left(1-\frac{|t|}{T_c}\right) + BG\frac{n_p-c_d}{n_p} & \text{for } -T_c < t < T_c \\ A_{sig} + BG\frac{n_p-c_d}{n_p} & \text{else,} \end{cases} \tag{3.28}$$

where the periodic repetitions have not been considered. In the following it is assumed that two correlation values are acquired for one distance measurement, A_{sig} is the mean number of photo-generated electrons per integration node, c_d denotes the demodulation contrast and BG is the mean number of electrons generated by the background light. The number of electrons photo-generated by the signal component can be expressed as

$$A_{sig} = \frac{n_p+1}{4n_p}\cdot GkP_{PN} = \frac{1}{2}GkP_A, \tag{3.29}$$

where the imbalance property of m-sequences has been taken into account. G is the generation efficiency as already defined by equation 2.22, and P_A denotes the mean emitted optical power of the light source similar to that for the sinusoidally modulated

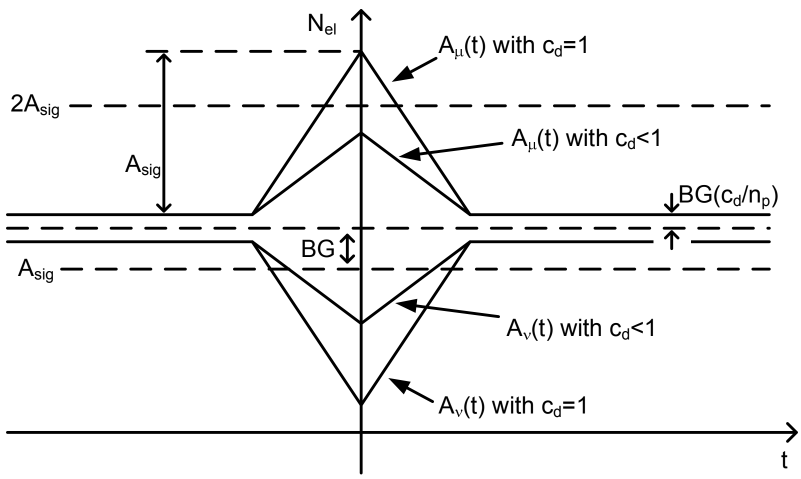


Figure 3.18: Time dependency of the output charge packets A_μ and A_ν of the pixel, when pseudo-noise modulation is used.

TOF system. The number of electrons generated by the background light is given as

$$BG = \frac{GP_B}{2}. \tag{3.30}$$

The division by 2 comes from the fact that A_{sig} and BG denote mean numbers of electrons per tap. During one acquisition over the integration time T, two taps are processed.

The correlation value corresponds to the difference of the two charge packets, which leads to

$$\Phi_a(t) = \begin{cases} 2c_d A_{sig}\left(1 - \frac{|t|}{T_c}\right) + 2c_d\frac{BG}{n_p} & \text{for } -T_c < t < T_c \\ \\ 2c_d\frac{BG}{n_p} & \text{else.} \end{cases} \tag{3.31}$$

This corresponds to the expected characteristics given by equation 3.20.

3.3.4 Distance Extraction with Maximal-Length PN-Sequences

In order to deduce a generic formula for the extraction of the distance we assume no background light $P_B = 0$. The effect of the background light is separately studied in the subsequent sections. Thus, the autocorrelation function (3.31) is written as

$$\Phi_a(t)|_{P_B=0} = \begin{cases} 2c_d A_{sig}\left(1 - \frac{|t|}{T_c}\right) & \text{for } -T_c < t < T_c \\ \\ 0 & \text{else.} \end{cases} \tag{3.32}$$

For the extraction of the time-of-flight τ, the significant triangular characteristics of the autocorrelation function are sampled twice at $t_0 = -\tau$ and $t_1 = T_c - \tau$, as depicted in figure 3.19. The first sample is obtained when no additional artificial time-shift is added to the reference function $r(t)$ and the optical signal has experienced a delay of τ. The second sample requires an artificial time-shift of the reference function by one chip duration T_c. Based on the two samples, the time-of-flight is computed as

$$\tau = \frac{\Phi_{a,s1}}{\Phi_{a,0} + \Phi_{a,s1}} \cdot T_c, \tag{3.33}$$

where $\Phi_{a,s0} = \Phi_a(-\tau)$ and $\Phi_{a,s1} = \Phi_a(T_c - \tau)$ are the two samples. The subscript s indicates that these samples of the autocorrelation function refer to the correlation

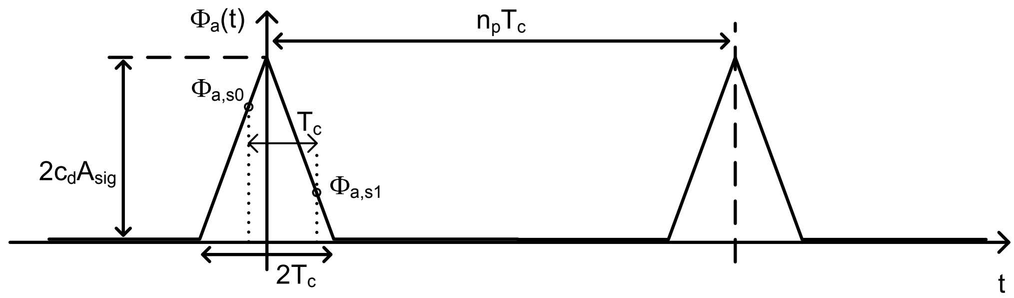


Figure 3.19: Sampling of the autocorrelation function with two samples $\Phi_{a,s0}$ and $\Phi_{a,s1}$ that are separated by one chip duration T_c. No background light is assumed in this depiction.

without taking into account any background light. In terms of number of electrons, the samples are expressed as

$$\begin{aligned} \Phi_{a,s0} &= 2c_d A_{sig} \left(1 - \tfrac{\tau}{T_c}\right), \\ \Phi_{a,s1} &= 2c_d A_{sig} \tfrac{\tau}{T_c}. \end{aligned} \tag{3.34}$$

The distance is computed with the use of equation 2.3 as

$$R = \frac{\Phi_{a,s1}}{\Phi_{a,s0} + \Phi_{a,s1}} \cdot R_{max}, \tag{3.35}$$

where $R_{max} = (cT_c)/2$ is the maximum detectable distance. Distances greater than R_{max} are not detected. Hence, there is no non-ambiguity problem similar to that in the case of sinusoidal modulation.

As long as only the pure signal component is taken into account, equation 3.35 provides physically reasonable results. However, if additional signals interfere, the correlation values can become negative. In order to avoid negative distances, the distance formula is modified to

$$R = \left(1 + \frac{\Phi_1 - \Phi_0}{|\Phi_1| + |\Phi_0|}\right) \cdot \frac{R_{max}}{2}. \tag{3.36}$$

The absolute values of the autocorrelation values in the denominator ensure the avoidance of negative distances, which would be have no physical meaning. In the subsequent error analysis, the last formula is particularly important for the analysis of the multi-camera operation. The other channel effects do not require the consideration of the absolute values, so that the raw formula of equation 3.35 leads to the same results.

3.4 Distance Error Analysis of PN Modulation

In this section we provide a detailed analysis of several error sources and their effects on the distance measurement when PN modulation is used. We begin with the performance analysis of pseudo noise sequences in a multi-camera environment, which was the basic motivation for the application of pseudo noise sequences as modulation signal. This investigation is followed by the performance analysis in a single camera environment and a comparison with sinusoidal modulation. The error aspects itemized below are studied in the following subsections:

- multi-camera interference (MCI),
- distance error due to non-linearities in the output transfer characteristics,
- the stochastic noise performance,
- error due to multi-path propagation (MPP) of the optical signal,
- error due to background light (BL),
- error due to low-pass (LP) filtered modulation or reference signals.

A general signal model with respect to the several sources of errors is depicted in figure 3.20. The optical signal of interest is denoted by $P_r(t)$ and the electronic reference signal is $r(t)$.

The analysis of the distance errors is based on the definition of the relative error

$$\delta = \frac{R_{meas} - R_t}{R_{max}}, \tag{3.37}$$

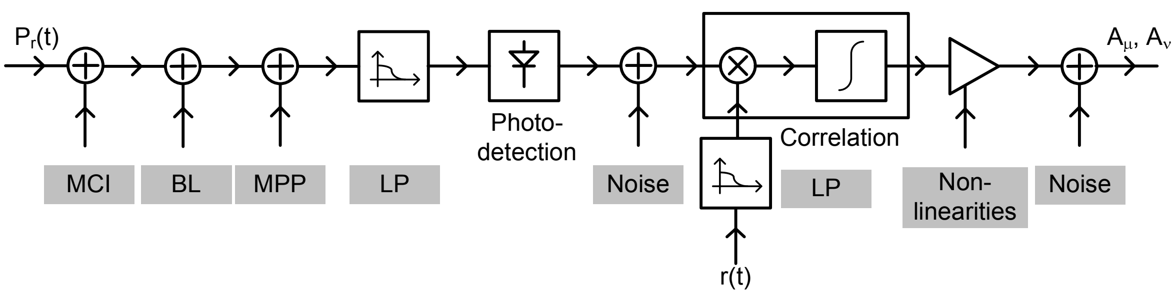


Figure 3.20: Signal model depicting the different error sources.

where R_{meas} is the measured and potentially erroneous distance value and R_t denotes the 'true' distance value, which is

$$\begin{aligned} R_t &= \tfrac{\Phi_{a,s1}}{\Phi_{a,s0}+\Phi_{a,s1}} \cdot R_{max} \\ &= \tfrac{\Phi_{a,s1}}{2c_d A_{sig}} \cdot R_{max}. \end{aligned} \tag{3.38}$$

It is recalled that the subscript s indicates that only the signal component may be considered for obtaining the true distance value.

Figures of merit: The performance of the modulation scheme is characterized in general by the distance error δ, as defined in equation 3.37. However, a more precise characterization of the different error sources necessitates the elaboration of the following two parameters:

$\bar{\delta} := \mathrm{mean}(\delta)$ is the expectation value of the distance error. This kind of parameter describes a systematic offset in the distance measurement depending on the particular error source considered.

$\sigma_\delta := \mathrm{std}(\delta)$ is the standard deviation of the distance error. This figure of merit describes the expectation of the distance fluctuations due to the particular error source.

The evaluation of the different measurement errors both in a multi-camera as well as in a single-camera environment is performed in the following subsections. It turns out that the sinusoidal modulation generally performs better in a single-camera environment, while pseudo noise sequences have their advantages in a multi-camera environment.

3.4.1 Multi-Camera Operation

The fundamental motivation for the use of pseudo noise sequences as a modulation signal is the enabling of the operation of several cameras at the same time without the need for camera synchronisation or separation in space. Since no strict mathematical descriptions of the cross-correlation functions of m-sequences exist, an exhaustive simulation has been set up for analysing the expected measurement behaviour [36] in a real-world multi-camera environment.

The distance measured by the g-th camera is influenced by the cross-correlation values between the unipolar modulation signal of interest $M_{u,g}(t)$ and the sum of the optical signals of concurrently running cameras. In a general form, the measured distance is given as

$$R_{meas} = \left(1 + \frac{\Phi_{a,s1} + \Phi_{c,\Sigma 1} - \Phi_{a,s0} - \Phi_{c,\Sigma 0}}{|\Phi_{a,s0} + \Phi_{c,\Sigma 0}| + |\Phi_{a,s1} + \Phi_{c,\Sigma 1}|}\right) \cdot \frac{R_{max}}{2}, \tag{3.39}$$

where $\Phi_{c,\Sigma 0}$ and $\Phi_{c,\Sigma 1}$ are the sum of the cross-correlation values at a particular delay τ_l:

$$\begin{aligned} \Phi_{c,\Sigma 0} &= G \cdot \sum_{l=0, l\neq g}^{N_c-1} (M_{b,g} \otimes P_{r,l})(-\tau_l) \\ \Phi_{c,\Sigma 1} &= G \cdot \sum_{l=0, l\neq g}^{N_c-1} (M_{b,g} \otimes P_{r,l})(T_c - \tau_l). \end{aligned} \tag{3.40}$$

As already shown in subsection 3.3.2, these cross-correlation values decrease for longer m-sequences, so that the error of the distance measurement δ is expected to have a similar dependency on the sequence length. Since no synchronisation between the cameras exists, the error can fluctuate due to the change of the cross-correlation values. Therefore, both the expected systematic distance error $\bar{\delta}$ as well as its fluctuations described by σ_δ are of interest.

Simulation Boundary-Conditions

The simulation has been set up according to figure 3.21. An arbitrary number of N_c cameras illuminates a target located at distance R. All cameras have the same distance to the target, emit the same optical power and use m-sequences of the same lengths but which are generated with different generator polynomials. The cameras are not synchronised with each other, so that the time delays between the m-sequences are randomly changed.

The simulation takes the before-mentioned conditions into account. For short m-sequences, the performance evaluation is based on the computations of all combinations between the m-sequences as well as the time delays. For long m-sequences, a restricted number of random combinations between m-sequences and time delays has been evaluated in order to reduce the computation time reasonably. Although this restriction makes the simulation of long m-sequences less accurate, the basic behaviour is still retained. The output of the simulation consists of the expectation value of the relative distance measurement error and the standard deviation of the error.

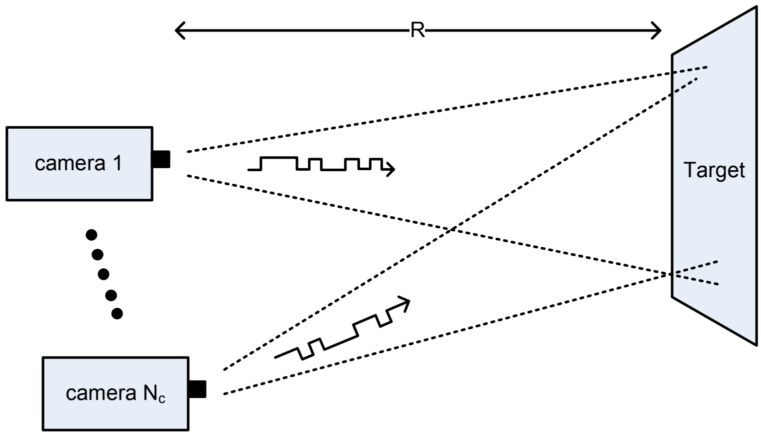
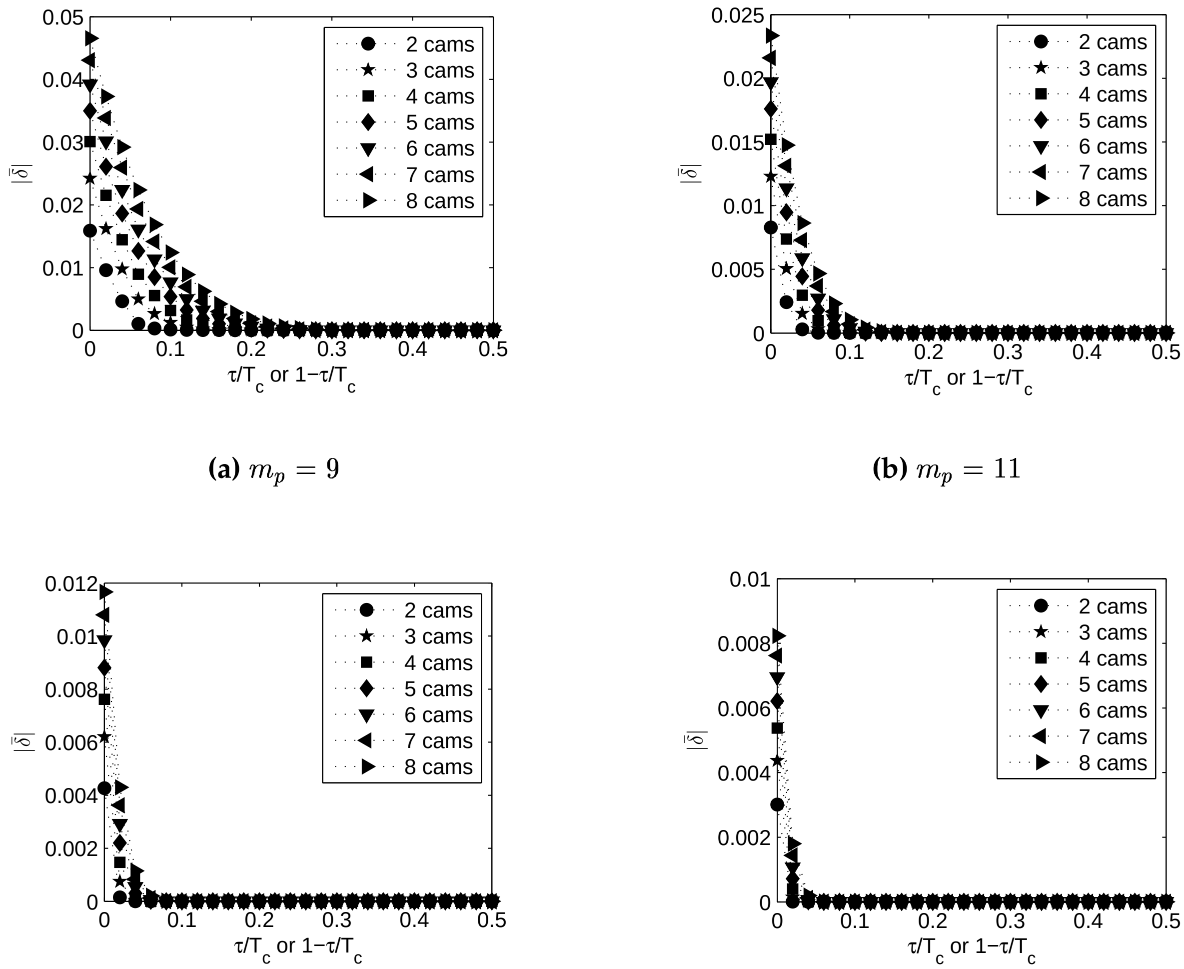


Figure 3.21: Multi-camera environment.

Simulation Results

The simulation has been performed for shift registers with an odd number of stages because it is known in advance that the maximum cross-correlation values are smaller than for even stage numbers, as shown in table 3.1 of the subsection 3.3.2.
Figures 3.22 (a)-(d) show the absolute of the expectation values of the distance error δ depending on the number of operating cameras and on the length of the LFSR. The depiction of the absolute values makes the graphs clearer. However, it should be mentioned that the mean error behaves in a complementary manner for delays of $\tau \in [0; 0.5]$ and $\tau \in [0.5; T_c]$. In the first interval of $\tau \in [0; 0.5]$ the error is positive so that its behaviour corresponds to that depicted: $\bar{\delta} = |\bar{\delta}|$. If $\tau \in [0.5; T_c]$, then the error curves would have a negative sign: $\bar{\delta} = -|\bar{\delta}|$. At $\tau = 0.5T_c$ the expectation value is perfectly 0 for any number of cameras or PN sequence length. Around this "perfect measurement point" the error tends to zero for a specific range of the time delay that depends directly on the sequence length. The worst case with the highest systematic measurement error is obtained at $\tau = 0$ or $\tau = T_c$.
Intuitive explanation of the error's behaviour: At the edges of the delay's range, any additions of correlation values to the autocorrelation function produce errors of the same sign. However, the more the delay tends to $0.5T_c$, the more there are combinations of cross-correlation values that completely compensate each other, and the distance error produced can be of positive or negative sign with the same probability. Hence, the main distance error tends to 0 for delays for which the cross-correlation values no longer produce negative summations.
The standard deviation of the measurement error, which is a measure of the expected fluctuations of the distance, is shown in the figures 3.23 (a)-(d). The longer the se-



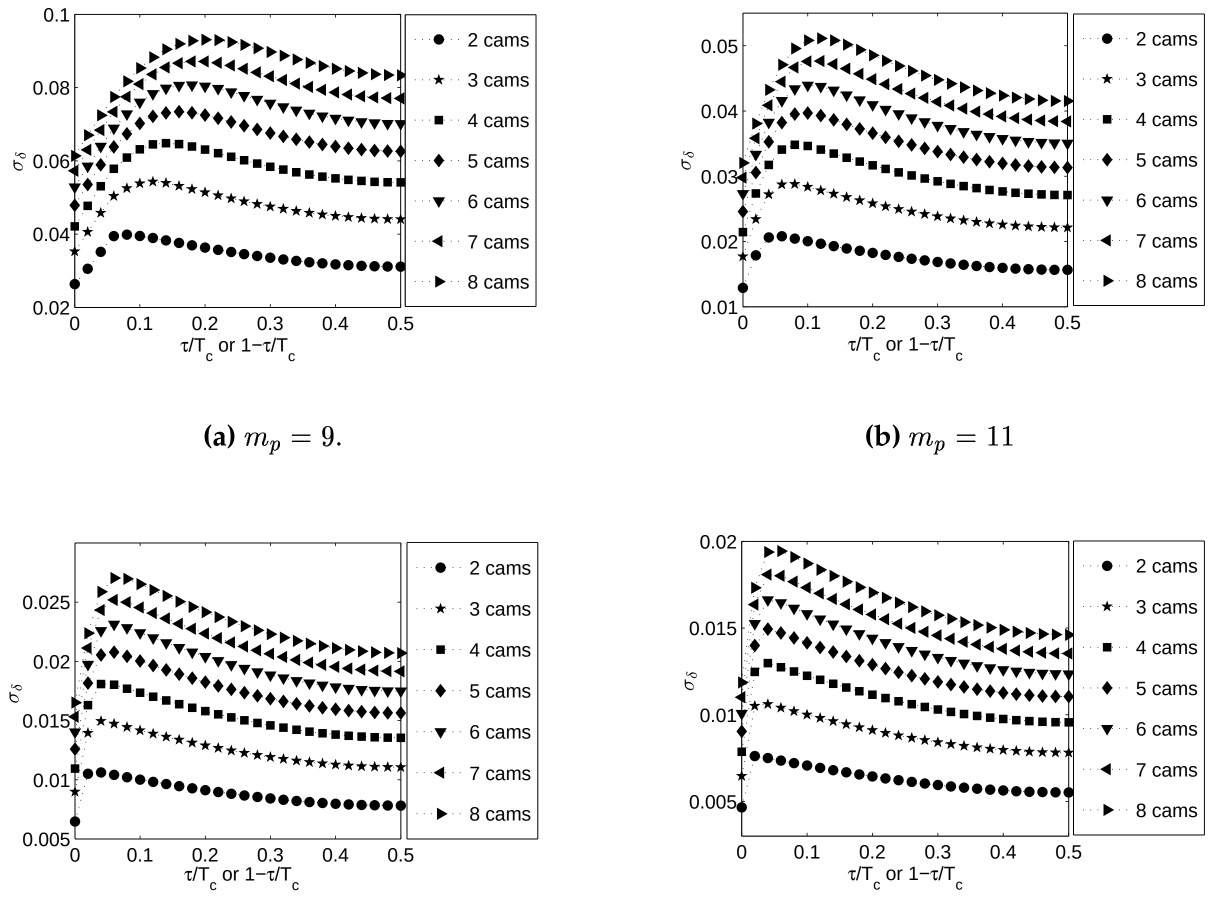
(a) $m_p = 9$ **(b)** $m_p = 11$

(c) $m_p = 13$ **(d)** $m_p = 15$

Figure 3.22: Expectation of the distance measurement error $|\bar{\delta}|$ for different numbers of cameras and lengths of the LFSR.

quence used, the better is the multi-camera performance with respect to this figure of merit. Best-case is found at a delay of $\tau = 0$ or $\tau = T_c$.

A summarising evaluation of the simulation is shown in figure 3.24. The left figure depicts the worst-case expectation-value of the distance measurement, which occurs at $\tau = 0$ (positive maximum error) or $\tau = T_c$ (negative maximum error). Figure 3.24 (b) shows the mean expected fluctuations, with the corresponding ranges from the minimum to maximum standard deviations.



(a) $m_p = 9$.

(b) $m_p = 11$

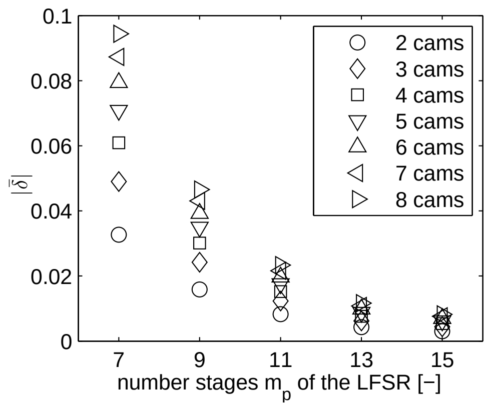
(c) $m_p = 13$

(d) $m_p = 15$

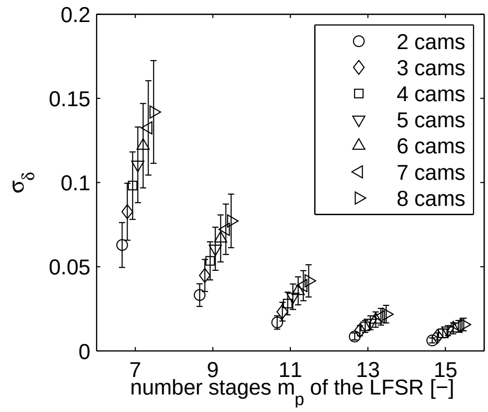
Figure 3.23: Depiction of the standard deviation of the distance measurement error σ_δ for different numbers of cameras and lengths of the LFSR.

Comparison with Sinusoidal Modulation

As shown in subsection 3.1 the superposition of several sinusoids of the same frequency can be treated in the complex plane. The resulting phasor is of arbitrary phase and lies within the circular area whose radius is the sum of the signal power of all other cameras. If only two cameras are operating with the same optical power, the circle limits the maximum relative error to ± 0.25, the expectation value is about 0 for time delays between $\tau = T_m/4$ and $\tau = 3T_m/4$, where the modulo-2π problem does not occur, and the standard deviation of the relative error is calculated as $1/\sqrt{48} \approx 0.144$ (see appendix A.3). With regard to the simulation conditions, giving each camera the same optical power, the case of two operating cameras is special, because the resulting phase



(a) Worst-case expectation values of the distance measurement $|\bar{\delta}|$ due to multi-camera interference. Worst-case occurs at $\tau = 0$ (positive maximum error) and $\tau = T_c$ (negative maximum error).



(b) Standard deviation of the relative distance errors σ_δ due to multi-camera interference, based on numerical simulations. Since σ_δ depends on the time delay of the signal of interest, the minimum and maximum boundaries are included by the bars as well.

Figure 3.24: Simulation results of the multi-camera interference for PN modulation.

is limited to $\varphi_\Sigma \in [\varphi_1 - \pi/2; \varphi_1 + \pi/2]$. When 3 cameras or more are operating at the same time, the resulting circle will describe all phasors of completely arbitrary phase between $]0; 2\pi]$, so that for such a case the standard deviation of the relative error tends rapidly to the maximum of $1/\sqrt{12} \approx 0.289$ (see appendix A.3), though the expectation value of the error is 0 because all phases occur with the same probability. Such significant distance error fluctuations are not tolerable at all. PN sequences perform much better in a multi-camera environment according to the simulation results.

3.4.2 Distance Resolution

The distance resolution of the time-of-flight measurement is computed as

$$\sigma_R = \sqrt{\sum_{k=0}^{3} \left(\frac{\partial R}{\partial A_k} \right)^2 \mathrm{Var}(\hat{A}_k)}, \tag{3.41}$$

where the distance is assumed to be expressed by the charge packets being read out. By performing the substitutions

$$\begin{aligned} A_0 &= A_\mu(-\tau) \\ A_1 &= A_\nu(-\tau) \\ A_2 &= A_\mu(T_c-\tau) \\ A_3 &= A_\nu(T_c-\tau) \end{aligned} \tag{3.42}$$

a corresponding expression is found for the distance as

$$\begin{aligned} R &= \tfrac{(A_\mu(T_c-\tau)-A_\nu(T_c-\tau))}{A_\mu(-\tau)-A_\nu(-\tau)+A_\mu(T_c-\tau)-A_\nu(T_c-\tau)} \cdot R_{max} \\ &= \tfrac{A_2-A_3}{A_0-A_1+A_2-A_3} \cdot R_{max}. \end{aligned} \tag{3.43}$$

Stochastic processes could produce negative distances, which are avoided in practice by taking the absolute values of the differences in the denominator (see equation 3.36). However, the calculation is performed without absolute values for simplicity. It is obvious that the effect of negative distances due to the stochastic variations only becomes noticeable for a very low number of electrons. However, the consideration of just 10 or 100 electrons does not make sense because huge standard deviations of the distance measurement would occur so that the system would not be applicable at all.

The computation of the distance resolution based on the Gaussian error propagation results in the standard deviation (the detailed computation can be found in the appendix A.4)

$$\sigma_R = \frac{R_{max}}{\sqrt{2}} \frac{\sqrt{A_{sig}+BG+BD+N_{ps}}}{c_d(A_{sig}+\frac{2}{n_p}BG)^2} \cdot \sqrt{\left(A_{sig}(1-\frac{\tau}{T_c})+\frac{BG}{n_p}\right)^2+\left(A_{sig}\frac{\tau}{T_c}+\frac{BG}{n_p}\right)^2}. \tag{3.44}$$

In order to refer to the figure of merit predefined at the beginning of section 3.4, it is noted that the expectation value of the distance is zero, so that

$$\bar{\delta} = 0. \tag{3.45}$$

As a consequence, the standard deviation of the distance error can be expressed in terms of the distance resolution as

$$\sigma_\delta = \frac{\sigma_R}{R_{max}}. \tag{3.46}$$

Conclusions: Three main consequences of the mathematical model should be emphasized before comparing the stochastic performance with a TOF system that is based on sinusoidal modulation.

1. A better distance resolution is achieved when using higher signal power or shorter chip durations, which in turn implies a smaller maximum detectable range R_{max}.
2. The measurement accuracy depends on the delay of the optical signal, where a best case is achieved for $\tau = T_c/2$ and a worst case for $\tau = 0$ or $\tau = T_c$.
3. As a consequence of the imbalance property of m-sequences, the sequence length n_p has a direct impact on the measurement accuracy.

Stochastic Performance with Sinusoidal and PN Modulation

A comparison between the distance resolutions, using either PN modulation or sinusoidal modulation, shows that the sinusoidal modulation provides a better stochastic performance. If the maximum detectable range of the PN modulation is set equal to the maximum non-ambiguity range of sinusoidal modulation $R_{max} = R_0$ and no background light is present, the standard deviation of a sinusoidally modulated system will be better by a factor of $2\sqrt{2}$ (best-case at $\tau = T_c/2$) to 4 (worst-case at $\tau = 0$ or $\tau = T_c$).

Simplifications for Special Environmental Conditions

By defining the background-to-signal light ratio as $BS := BG/A_{sig}$ and assuming that the dark current and any additional noise sources are negligible ($BD = N_{ps} = 0$), then equation 3.44 can be rewritten as

$$\sigma_\delta = \frac{\sqrt{1+BS}}{\sqrt{2A_{sig}}c_d(1+\frac{2BS}{n_p})^2} \cdot \sqrt{\left((1-\frac{\tau}{T_c}) + \frac{BS}{n_p}\right)^2 + \left(\frac{\tau}{T_c} + \frac{BS}{n_p}\right)^2}, \qquad (3.47)$$

with the utilisation of equation 3.46. Two cases are highlighted in the following.

1. $n_p >> BS$: In general the sequence length is much bigger than the ratio between background light and signal component $n_p >> BS$, so that the standard deviation is further simplified to

$$\sigma_\delta \approx \frac{\sqrt{1+BS}}{\sqrt{2A_{sig}}c_d} \cdot \sqrt{1 - 2\frac{\tau}{T_c} + 2\left(\frac{\tau}{T_c}\right)^2}. \qquad (3.48)$$

Figures 3.25 (a) and (b) show the dependencies of $\sigma_\delta = \sigma_R/R_{max}$ on the mean number of electrons generated by the signal component (A_{sig}) and the background-to-signal light ratio for the best-case at a delay of $\tau/T_c = 0.5$ and the worst-case at $\tau/T_c = 0$. Furthermore, an ideal demodulation contrast of $c_d = 1$ has been assumed for the computations. It is noted that the figures depict the mathematical limits according to the approximation given by $n_p >> BS$.

In addition, figure 3.25 (c) shows the exact dependence of the standard deviation on the number of electrons generated by the signal and the relative time delay τ/T_c when the background light is also not present at all ($BS = 0$), so that

$$\sigma_\delta = \frac{1}{\sqrt{2A_{sig}c_d}} \cdot \sqrt{1 - 2\frac{\tau}{T_c} + 2\left(\frac{\tau}{T_c}\right)^2}. \tag{3.49}$$

2. $BS >> n_p$: The case that the ratio BS between background light and signal is much bigger than the sequence length n_p, is rather improbable but it can occur. The consideration leads to the simplified transformation of equation 3.48 as

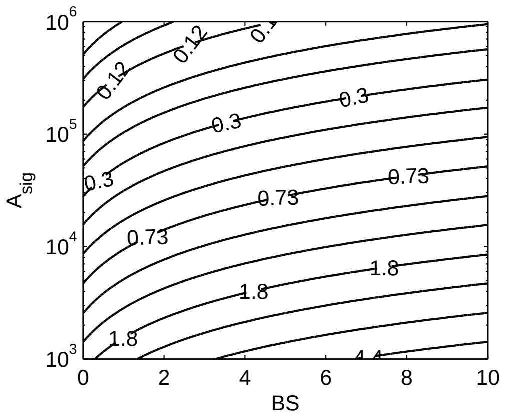
$$\sigma_\delta = \frac{n_p}{4c_d\sqrt{BG}} = \frac{n_p}{4c_d\sqrt{A_{sig}BS}}. \tag{3.50}$$

In this case, there is only a dependence of the distance resolution on the background light and the sequence length. This is rather an unconventional behaviour due to the imbalance property by 1 chip of the m-sequences. The more background light, the lower is the standard deviation because the background light acts like a signal component. As an example, figure 3.25 (d) depicts the relative errors for a sequence length of $n_p = 2^{15} - 1$ chips and a time delay $\tau = T_c/2$. The dependencies on the number of generated electrons and different background-to-signal light ratios have been incorporated.

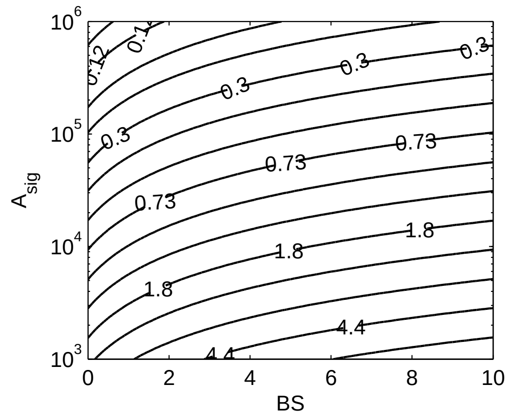
Influence of Multi-Camera Interferences on the Stochastic Performance

The detection of the light signals of additional cameras can be treated to a first approximation as additional background light that decreases the stochastic performance. The incorporation in the equations of the standard deviation results in

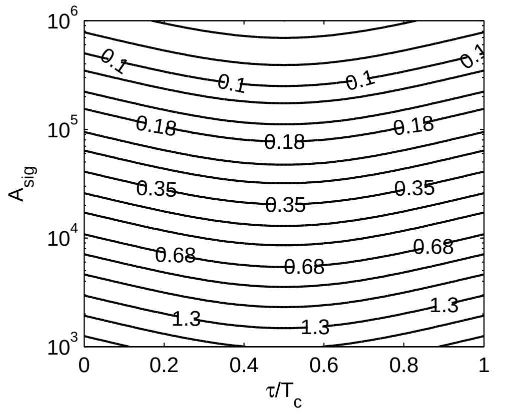
$$\sigma_R = \frac{R_{max}}{\sqrt{2}} \frac{\sqrt{A_{sig} + BG + BD + N_{ps} + BC}}{c_d(A_{sig} + \frac{2}{n}BG + \Phi_{c,\Sigma})^2} \cdot \sqrt{\left(A_{sig}(1 - \frac{\tau}{T_c}) + \frac{BG}{n} + \frac{\Phi_{c,\Sigma}}{2}\right)^2 + \left(A_{sig}\frac{\tau}{T_c} + \frac{BG}{n} + \frac{\Phi_{c,\Sigma}}{2}\right)^2}, \tag{3.51}$$



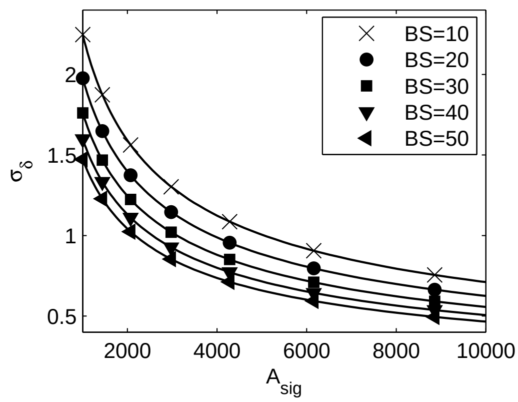
(a) Best-case for a delay of $\tau = 0.5T_c$. The graphs are approximations for $n_p >> BS$.



(b) Worst-case for a delay $\tau = 0$ or $\tau = T_c$. The values of σ_δ are greater by a factor of $\sqrt{2}$ than in the best-case at $\tau = 0.5T_c$. The graphs show the limits for the approximation of $n_p >> BS$.



(c) Relative error depending on the mean number of electrons A_{sig} generated by the signal component per integration node and the relative time delay τ/T_c. The graphs are based on the exact calculation for $BS = 0$.



(d) Exact calculation of the relative error for a PN sequence length $n_p = 2^{15}-1$ and a time delay $\tau/T_c = 0.5$. Sweeps over the mean number of electrons generated by the signal light A_{sig} and the background light to signal light ratio BS have been performed. The graphs show the behaviour for the extreme condition that $BS >> n_p$.

Figure 3.25: Standard deviation of the relative distance error $\sigma_\delta = \sigma_R/R_{max}$ as a consequence of stochastic noise contributions. The values have to be multiplied by $\cdot 10^{-2}$. The computations are based on an ideal demodulation contrast $c_d = 1$.

with $\Phi_{c,\Sigma} = (\Phi_{c,\Sigma 0} + \Phi_{c,\Sigma 1})/2$ and BC is the total number of electrons generated by the signals of concurrently running cameras. $\Phi_{c,\Sigma 0}$ and $\Phi_{c,\Sigma 1}$ are the sums of the cross-correlation values resulting from the interference signals, according to the definitions in equation 3.40.

3.4.3 Non-Linearity of the Output Transfer Characteristics

The effect of non-linearities of the output transfer characteristics on the distance measurement has been summarised in section 2.2.5 with regard to the sinusoidal modulation of the optical signal. It has been shown that there is no systematic error expected by first-order distortions of the output characteristics when the same output characteristics are assumed to be used for the read out process of each charge packet. The same effect can be examined in a similar way for PN modulated signals by defining the measured values to be rather $a + bA_g + cA_g^2$ instead of A_g. With the use of equation 3.43, the measured distance is then

$$R_{meas} = \frac{a + bA_2 + cA_2^2 - a - bA_3 - cA_3^2}{a + bA_0 + cA_0^2 - a - bA_1 - cA_1^2 + a + bA_2 + cA_2^2 - a - bA_3 - cA_3^2} \cdot R_{max}, \quad (3.52)$$

where the constant offsets a are automatically eliminated by the difference computations. Based on the fact that $A_0 + A_1 = A_2 + A_3 = 2(A_{sig} + BG)$, the following transformations are applied:

$$\begin{aligned} R_{meas} &= \frac{a(A_2-A_3)+2b(A_{sig}+BG)(A_2-A_3)}{a[(A_0-A_1)+(A_2-A_3)]+2b(A_{sig}+BG)[(A_0-A_1)+(A_2-A_3)]} \\ &= \frac{A_2-A_3}{(A_0-A_1)+(A_2-A_3)} \\ &= R_t. \end{aligned} \quad (3.53)$$

The measured distance corresponds to the true distance value R_t, or the relative error is equal to zero

$$\delta = 0. \quad (3.54)$$

It is concluded that quadratic distortions of the output transfer characteristics do not influence the distance measurement, either with sinusoidal or with PN optical modulation, when the same output transfer characteristics can be assumed for the readout of each charge packet.

3.4.4 Background Light

While the model in section 3.3.3 already describes the behaviour of the pixel's correlation output with additional background light, the generic distance formula, however, is based exclusively on the signal component when no background light is present. Thus, background light leads to an erroneous distance measurement.

The background light implies a positive offset of the autocorrelation function by $(2c_d BG)/n_p$, due to the (im)balance property of m-sequences. Taking the background light into account, the measured distance is

$$R_{meas} = \frac{\Phi_{a,s1} + \frac{2c_d BG}{n_p}}{2c_d A_{sig} + \frac{4c_d BG}{n_p}} \cdot R_{max}. \tag{3.55}$$

As a consequence, the relative error according to its definition in equation 3.37, is computed as

$$\delta = \frac{1 - 2\frac{\tau}{T_c}}{2 + \frac{n_p}{BS}}, \tag{3.56}$$

where $BS = BG/A_{sig}$ defines the ratio between background-light power and the signal-light power. The expected error, depending on the background-light to signal-light ratio, the sequence length and the particular delay τ of the measurement, is shown in figure 3.26. For long PN sequences or a low power of background light the error tends to zero. No error is obtained at a delay of $\tau = T_c/2$.

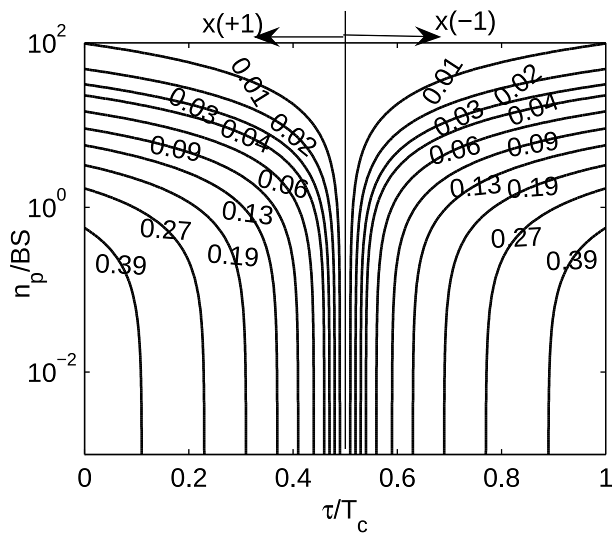


Figure 3.26: Relative distance error δ due to incident background light, normalized to the maximal detectable distance R_{max}.

3.4.5 Effect of Low-Pass Filtered Signals

Both signals, the optical modulation signal as well as the reference signal, experience a low-pass filtering due to capacitive loads of the electronic components. A low-pass generates a systematic distortion of the distance linearity. By modelling the low-pass system with an impulse response function $h(t)$, the correlation function is distorted as

$$\begin{aligned}\Phi_{a,lp}(t) &= G \cdot [(P \otimes S)(t)] * h(t) \\ &= \Phi_a(t) * h(t).\end{aligned} \tag{3.57}$$

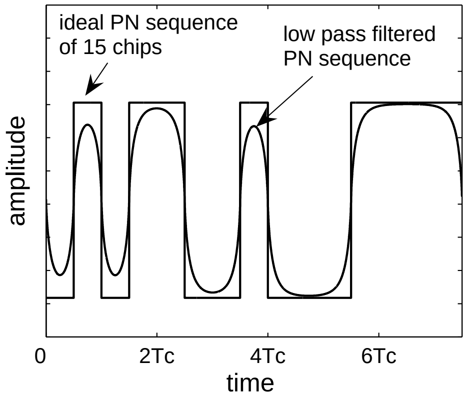
As a consequence of any low-pass effects in the signal chain of the modulation signals, the correlation function is directly affected by the same low-pass characteristics, so that a smoothing effect of the ideal peak is expected. For the assumption of a low-pass filtering by an RC element, the deformations of the ideal PN signal and of the autocorrelation function are depicted in figures 3.27 (a) and b, for the example of a 3dB cut-off frequency of $f_{cut} = 0.5f_c$, where f_c is the chip frequency. Filter-related time shifts have been compensated in this figure.

The relative error depends on the particular distance or time delay measured and on the low-pass characteristics of the system. Figure 3.27 (c) shows the theoretically expected relative distance error when an RC low-pass is assumed for the estimation. The error is positive if $\tau/T_c \leq 0.5$ and it is negative if $\tau/T_c > 0.5$.

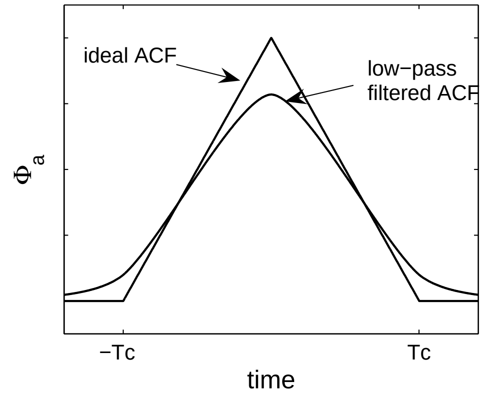
3.4.6 Multi-Path Propagation

Another source of error is multi-path propagation of the optical signal. A pulsed time-of-flight system is not influenced at all by multi-path propagations because the shortest transition path, which triggers the distance measurement, corresponds to the distance sought. However, a continuously modulated system such as the PN-modulated TOF system, is much more sensitive to superpositions of the same signal with different delays. The problem of multi-path propagation is depicted in figure 3.28, where two kinds of multi-path propagation are distinguished. The first one consists of multiple reflections in the scene as depicted in figure 3.28a. Secondly, multiple reflections occur between the sensor and the optics, which is generally known as the scattering-light effect. This kind of intra-pixel optical cross-talk is schematically illustrated in figure 3.28b.

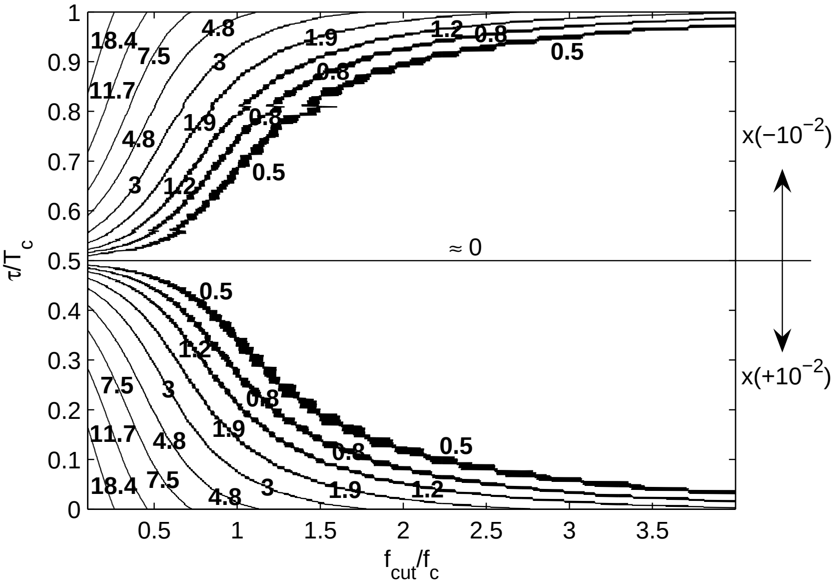
Both types of multiple reflections lead to an erroneous distance measurement. The scattering effect of the optical set-up can be minimized by



(a) PN sequence over time: ideal and low-pass filtered.

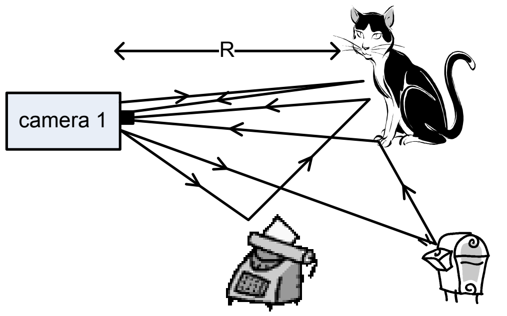


(b) Autocorrelation function of the ideal PN sequence and of the low-pass filtered sequence.

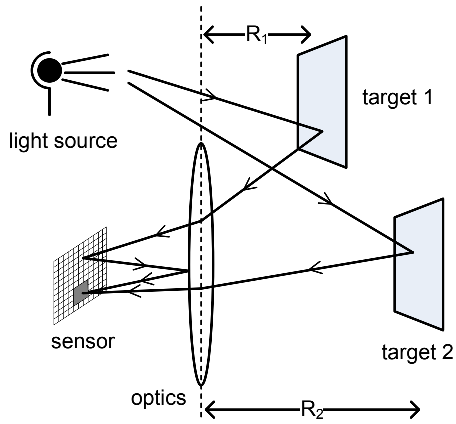


(c) Simulation of the relative distance error δ due to low-pass filtering depending on the ratio between cut-off frequency of an RC element f_{cut} and the chip frequency f_c and the relative time delay τ/T_c. This delay covers the full range from 0 to 1; the ratio between the cut-off frequency and the chip frequency is considered within the limited range from 0 to 4.

Figure 3.27: Effect of low-pass filtering with an RC element when $f_{cut} = 0.5 f_c$. The filter-related time delays have been compensated in these depictions.



(a) Multiple reflections in the scene.



(b) Multiple reflections between the sensor and the optics.

Figure 3.28: Illustration of the two types of multi-path propagation.

- appropriate optimization of the optics including
  - coating of the lens,
  - coating of the sensor,
  - improving the focus on the sensitive area of the pixels by the use of micro-lenses on the sensor,
- appropriate pixel/sensor design implying the avoidance of large areas of highly-reflective metal layers,
- by measuring the optical cross-talk and subsequently calculating the effect of the intensity distribution on the distance measurement so that a calibration can be effected.

The multiple reflections in the scene are difficult to minimize. The problem is well-known in mobile communication systems where multi-path propagation of the signal leads to significant bit error rates after reception. Depending on the operation environment, stochastic models such as Rayleigh and Rice distributions [32,85] are used for the description of the multi-path propagation. Similar models have to be incorporated into the 3D TOF system in order to reduce the distance error due to multi-path propagation. Though such models have not yet been applied, a general comparison of the distance error rate that is expected with sinusoidal or PN modulation is provided below.

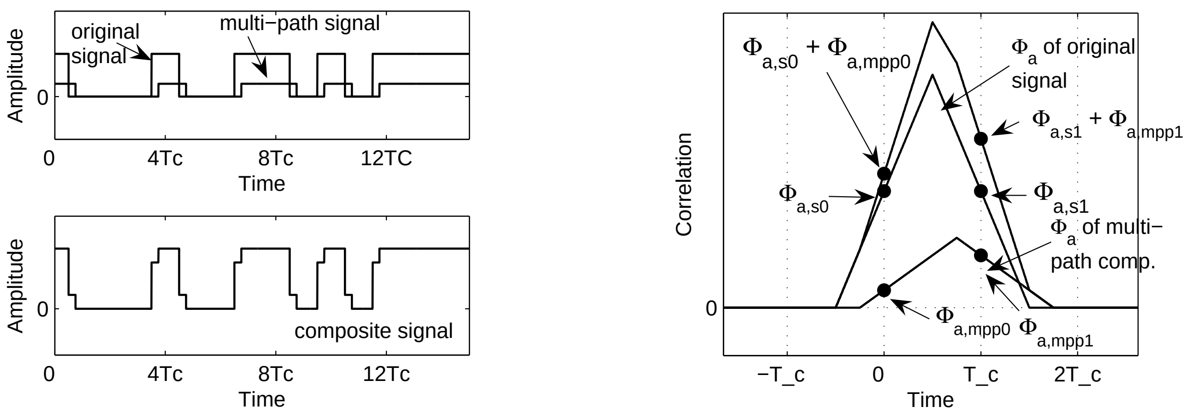
The superposition of the original signal and the multi-path propagated signal is shown in figure 3.29a assuming no background light illumination. The attenuation k_{mpp} denotes the ratio between the mean optical power of the original signal and the signal that is added due to the multi-path delay. In addition, the time-of-flight sought is given by τ, while the multi-path component has experienced a delay of τ_{mpp}. The effect on the autocorrelation function is shown in figure 3.29b for the case of PN modulation. Two autocorrelation values $\Phi_{a,mpp0}$ and $\Phi_{a,mpp1}$ are added to the unaffected values $\Phi_{a,s0}$ and $\Phi_{a,s1}$, respectively. $\Phi_{a,mpp0}$ and $\Phi_{a,mpp1}$ are the samples of the autocorrelation function that is multi-path delayed and not offset-shifted due to any background light. Consequently, the relative distance error is expressed in a general form as

$$\delta = \frac{\Phi_{a,s1} + \Phi_{a,mpp1}}{\Phi_{a,s0} + \Phi_{a,s1} + \Phi_{a,mpp0} + \Phi_{a,mpp1}} - \frac{\Phi_{a,s1}}{\Phi_{a,s0} + \Phi_{a,s1}}, \tag{3.58}$$

whith $\Phi_{a,mpp0} = k_{mpp}\Phi_a(-\tau_{mpp})$ and $\Phi_{a,mpp1} = k_{mpp}\Phi_a(T_c - \tau_{mpp})$.

Multiple Reflections in the Scene

By considering only multi-path components of a delay $\tau_{mpp} > \tau$, which is a valid assumption for the multi-path propagation in the scene, two dependencies of the error can be derived for two cases:



(a) Superposition of two pseudo-noise signals.

(b) Correlation functions of the original PN signal, the multi-path PN signal and the signal composed of both of them.

Figure 3.29: Multi-path propagation of pseudo-noise signals and the effect on the correlation function.

1. $(\tau_{mpp} - \tau) \in [0; T_c - \tau] \bigwedge \tau_{mpp} > \tau$:
   The measurement error results from the addition of both autocorrelation values $\Phi_{a,mpp0}$ and $\Phi_{a,mpp1}$ in
   $$\delta = \frac{k_{mpp}}{k_{mpp}+1} \cdot \frac{\tau_{mpp} - \tau}{T_c} \geq 0$$

2. $(\tau_{mpp} - \tau) \in [T_c - \tau; 2T_c - \tau] \bigwedge \tau_{mpp} > \tau$:
   The error is only affected by $\Phi_{a,mpp1}$ because $\Phi_{a,mpp0} = 0$. It is computed as
   $$\delta = \frac{k_{mpp}\left[2\left(T_c - \tau\right) - \tau_{mpp}\left(1 - \frac{\tau}{T_c}\right)\right]}{T_c\left(1 + k_{mpp}\left(2 - \frac{\tau_{mpp}}{T_c}\right)\right)} \geq 0$$

While the error is increasing in the first specified interval of τ_{mpp}, it is decreasing in the second interval but still greater than or equal to zero. The maximum error is found to be

$$\delta_{max} := \max|\delta| = \frac{k_{mpp}}{1+k_{mpp}} \cdot \left(1 - \frac{\tau}{T_c}\right), \tag{3.59}$$

for $0 < \tau/T_c < 1$ and $\tau_{mpp} = T_c$.
The attenuation factor k_{mpp} is limited to $[0; 1]$ for the multi-path in the scene.

Scattering Light

For scattering light the relation $\tau_{mpp} < \tau$ is also possible. The same absolute maximum error as before is obtained but two additional conditions have to be considered:

$$\delta_{max} := \max|\delta| = \begin{cases} \frac{k_{mpp}}{1+k_{mpp}} \cdot \left(1 - \frac{\tau}{T_c}\right), & \text{for } 0 < \tau/T_c \leq 0.5 \\ \frac{k_{mpp}}{1+k_{mpp}} \cdot \frac{\tau}{T_c}, & \text{for } 0.5 < \tau/T_c \leq 1 \end{cases}. \tag{3.60}$$

The multiple reflections of the optics can produce superpositions with an even stronger multi-path signal than the original, so that $k_{mpp} \in [0; +\infty]$.

Comparison

The differences between multi-path propagation in the scene and the optics are summarized in table 3.2 including the particular mathematical expressions. The expressions for the maximum error with sinusoidal modulation are derived from the corresponding considerations of the addition of two sinusoids in the complex plane, which has been shown in detail in appendix A.1. The maximum phase error is given as

$$\Delta\varphi_{max} = \arctan\left[\frac{k_{mpp}}{\sqrt{1-k_{mpp}^2}}\right] \tag{3.61}$$

In particular, the modulo-2π effect has been taken into account in the expressions of table 3.2. The modulation period of the sinusoidal modulation signal has been set equal to the chip duration $T_m = T_c$, meaning that the maximum detectable range in the PN case equals the unambiguity range in the sine case. Following this, all considerations are referred to the chip duration T_c.

The relationships between the maximum distance error, the time-of-flight and the attenuation coefficient are graphically shown in figures 3.30 and 3.31 for the propagation in the scene and for the scattering-light effect, respectively. The computations are based on the superposition of two signals.

It is obvious that the environmental conditions are decisive for finding the preference of one modulation scheme with respect to the MPP effect. The sinusoidal modulation performs better at close range if MPP in the scene is involved. Regarding the scattering-light effect, the sinusoidal modulation is better for relative distances from 0.2 up to 0.8 when the attenuation factor k_{mpp} can be assumed to be less than one. However, due to the modulo-2π problem, scattering light can also lead to significant distance errors when sinusoidal modulation is used.

Table 3.2: Multi-path propagation: in the scene and scattering light.

	MPP in the scene	**Scattering light**
Delays of the MPP signals	$\frac{\tau}{T_c} < \frac{\tau_{mpp}}{T_c} < 1$	$0 < \frac{\tau_{mpp}}{T_c} < 1$
Attenuation of the MPP signals	$0 < k_{mpp} < 1$	$0 < k_{mpp}$
Maximum error δ_{max} sinusoidal	$\frac{\Delta\varphi_{max}}{2\pi}$ for $0 < \frac{\tau}{T_c} < 1 - \frac{\Delta\varphi_{max}}{2\pi}$ $\frac{\tau}{T_c}$ for $1 - \frac{\Delta\varphi_{max}}{2\pi} < \frac{\tau}{T_c} < 1$	$\frac{\Delta\varphi_{max}}{2\pi}$ for $\frac{\Delta\varphi_{max}}{2\pi} < \frac{\tau}{T_c} < 1 - \frac{\Delta\varphi_{max}}{2\pi} \wedge k_{mpp} < 1$ $1 - \frac{\tau}{T_c}$ for $k_{mpp} > 1$ $1 - \frac{\tau}{T_c}$ for $0 < \frac{\tau}{T_c} < \frac{\Delta\varphi_{max}}{2\pi}(\wedge 1 > k_{mpp} > 0)$ $\frac{\tau}{T_c}$ for $1 - \frac{\Delta\varphi_{max}}{2\pi} < \frac{\tau}{T_c} < 1 \wedge 0 < k_{mpp} < 1$ no distance value for $k_{mpp} = 1$
Maximum error δ_{max} PN	$\frac{k_{mpp}}{1+k_{mpp}} \cdot \left(1 - \frac{\tau}{T_c}\right)$ for $0 < \frac{\tau}{T_c} < 1$	$\frac{k_{mpp}}{1+k_{mpp}} \cdot \left(1 - \frac{\tau}{T_c}\right)$ for $0 < \frac{\tau}{T_c} < 0.5$ $\frac{k_{mpp}}{1+k_{mpp}} \cdot \frac{\tau}{T_c}$ for $0.5 < \frac{\tau}{T_c} < 1$

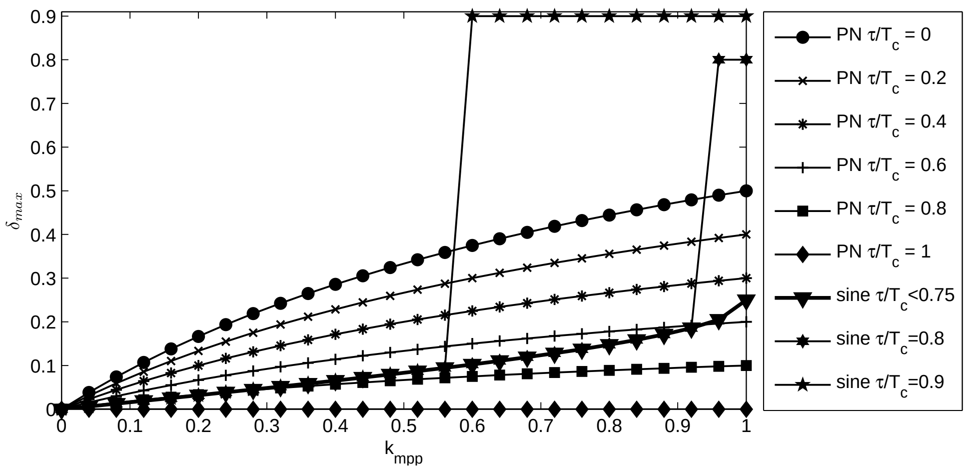
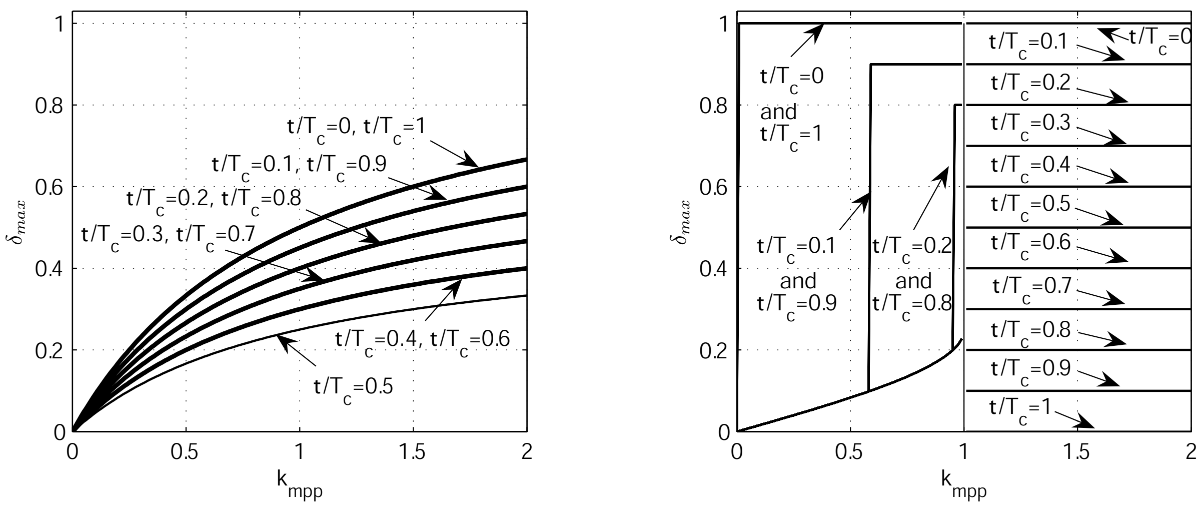


Figure 3.30: Maximum relative distance error. Two signals are added and one of them is additionally delayed and attenuated. This corresponds to the case of multi-path propagation in the scene.



(a) Pseudo-noise sequence modulation.

(b) Sinusoidal modulation.

Figure 3.31: Maximum relative distance error due to multiple reflections between the sensor and the optics. The computation corresponds to the superposition of two signals that can have arbitrary delays and amplitudes.

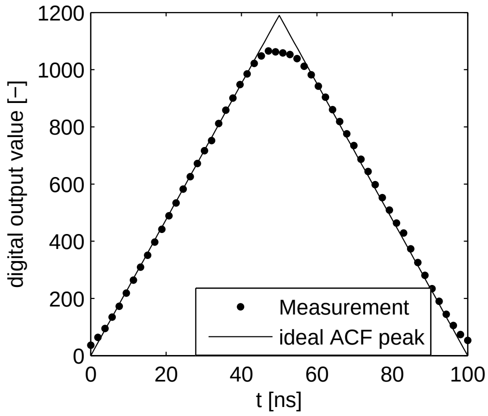
It is noted that the maximum error is only one figure of merit characterizing the performance of the modulation scheme with regard to the MPP effect. However, with knowledge of the environmental conditions concerning the optical reflections, estimations of the expectation values of the errors and their probabilities, based on stochastic models as used for example in mobile communications, would allow an even more thorough performance evaluation and comparison.

3.5 Experimental Results

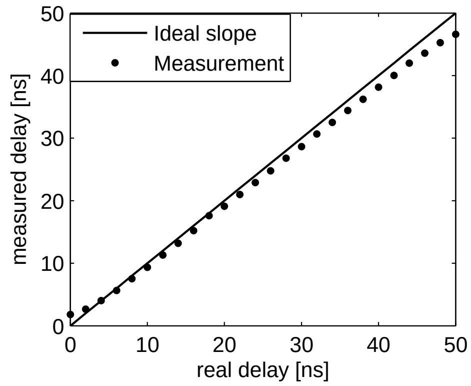
The measurements presented in this section have been performed with the SwissRanger 3D TOF camera [75], having a pixel resolution of 176×144 pixels. These measurements show that the system performs as predicted by the well-developed theory of pseudo noise modulated time-of-flight distance measurement. The different channel effects of stochastic distortion by shot noise sources and multi-camera interference have been measured and they show good agreement with the theory. The systematic errors could not be measured due to the dominating impact of the camera's non-linearities.

First, the principle of measuring distances with pseudo-noise modulated optical signals is verified by the measurement results in figure 3.32. This measurement has been performed without background light and no additional cameras. A maximal-length pseudo-noise sequence of 2047 chips has been used with a chip duration of $T_c = 50$ns, which allows for measuring distances up to $R_{max} = 7.5\mathrm{m}$. Figure 3.32 (a) shows the acquisition of the autocorrelation function when the delay is externally set by the operator and the actual distance to the target is kept constant. Figure 3.32 (b) shows the result of the delay extraction for delays ranging from 0 to $50\mathrm{ns}$. The non-linearities, which are particularly observed at long delays, are due to the unavoidable low-pass filtering of the modulation signal, as predicted in section 3.4.5. This effect is also recognised in the autocorrelation peak, which is slightly smoothed due to the low-pass filtering.

Based on the same measurements, the evaluation of the stochastic performance without background light is shown in figure 3.33 (a). The theoretical curve is based on equation 3.44 with $BG = N_{ps} = 0$ and $A_{sig} = \mathrm{const.}$. The demodulation contrast has been measured in advance to be about 62%.
Another verification of the noise model is given by figure 3.33 (b). The standard deviation of the distance has been measured for different background-light to signal-light ratios but at constant signal power. Both sinusoidal and PN modulation techniques have been investigated. The theoretical curves based on equation 2.30 for the sinusoidal modulation and equation 3.44 for the PN modulation are included in the figure. For both



(a) Measured autocorrelation function.

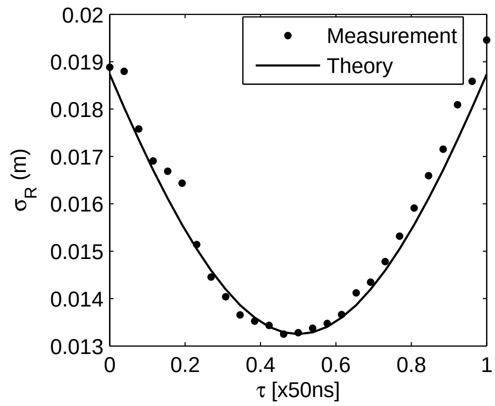


(b) Linearity measurement.

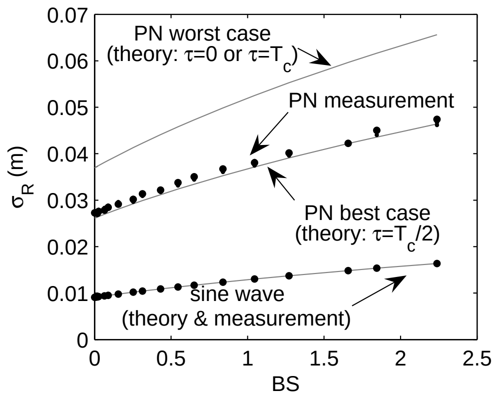
Figure 3.32: Proof of the measurement principle: TOF distance measurement with pseudo-noise intensity-modulated optical signals.

modulation schemes, good agreement between theory and practice is obtained. Since the PN measurement has been performed at a delay of $\tau = 0.5T_c$ the values are close to the best-case curve. The sequence lengths tested have 511, 2047, 8191 and 32767 chips. The difference in their performance is not recognisable in the figure. Furthermore, the sinusoidal modulation shows better performance for low background- to signal-light ratios when $T_c = T_m$, as theoretically predicted.

The last measurement shows the performance of the PN-modulation in a multi-camera environment, which was the initial motivation for the utilization of PN-sequences instead of sine waves for modulating the optical signal. Figure 3.34 shows the measured standard deviation of the distance error, which is a measure of the observable fluctuations of the distances. The measurement set-up corresponds to the simulation assumptions: the camera of interest receives the same amount of optical power as all other cameras together. The dependence of the distance error fluctuations on the number of cameras operating concurrently and on the length of the PN sequence is depicted. The evaluated camera is operating at a delay of approximately 25ns, which is half the chip duration T_c. The basic characteristics correspond to those simulated in figure 3.24 (b). Fluctuations of around 7cm have been measured when four additional cameras are running and a sequence length of 32767 chips has been used. The results are slightly better than those of the simulation because some worst-case conditions between m-sequences are avoided consciously in the measurements but not in the simulation.



(a) Dependence of the distance standard deviation due to photon shot noise. Measurement without background light and constant optical signal power.



(b) Dependence of the distance standard deviation on the background to signal ratio BS, with a comparison to theory and to the performance measurement of a sinusoidal modulated system. The modulation period is equal to the chip duration $T_c = T_m = 50\text{ns}$.

Figure 3.33: Stochastic performance of optical distance measurement using PN and sinusoidal modulation.

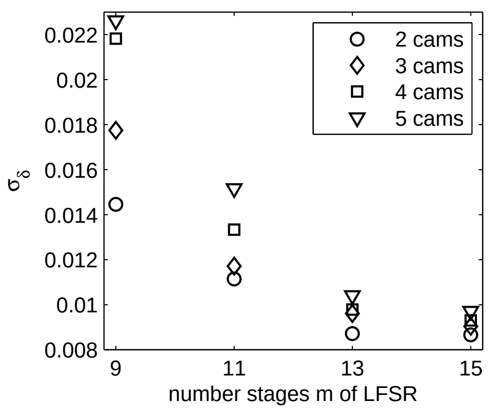


Figure 3.34: Measured standard deviation of the error signal due to multi-camera interference.

3.6 Discussion

A detailed investigation of the utilisation of pseudo-noise intensity-modulated optical signals for real-time 3D-imaging has been presented in the sections above, where the main focus was applicability to a multi-camera environment. Compared to the sinusoidal modulation, PN-sequences show their main advantage in the reduction of multi-camera interference and, thus, provide a suitable alternative modulation scheme for systems where several cameras can approach each other randomly and without the requirement of a time-synchronisation.

However, the performance of PN modulation in a single-camera environment is generally inferior to that of sinusoidal modulation signals. In particular, the stochastic performance with PN-modulation is worse when the maximum detectable range equals the maximum non-ambiguity range in the case of sinusoidal modulation. This observation is independent of the environmental conditions with regard to the delay or the background light. However, the nonexistent problem of non-ambiguity measurement for the PN modulation can be very advantageous for some applications, e.g. in the field of robotics, where distances greater than a particular distance are not interesting for the operator.

It is obvious that future work includes the search for modulation techniques that allow for measuring distances in both a single and a multi-camera environment without significant performance losses compared to the sinusoidal modulation in a single-camera environment.
On the one hand, research can continue to focus on pure encoding techniques like pseudo-noise modulation. On the other hand, alternative techniques such as the combination of sinusoidal modulation and pseudo-noise modulation could be investigated with the goal of combining the advantages of both. Some initial ideas are summarised below.

- *M-Sequences and background light susceptibility:* In section 3.4.4, the susceptibility of pure m-sequences to background light has been demonstrated, where a strong error-dependence on the particular delay or range has been observed. By padding the m-sequence with an additional zero at the end, the imbalance property of PN sequences could be avoided. This has been suggested in [27]. Though the autocorrelation properties do not change around the peak, there will be an impact on the multi-path propagation which has to be examined in more detail. Furthermore, the cross-correlation properties will change significantly, so that the multi-camera interferences have to be studied again.

- *M-Sequences and maximum range:* Extension of the maximum detectable range could be achieved by an evaluation of the discriminator function Φ_Δ, which is defined as the difference between the autocorrelation function and its replica that is time-shifted by $2T_c$. Figure 3.35 shows the construction of the discriminator function. Mathematically it is described as

$$\begin{aligned}\Phi_\Delta(t) &= \Phi_a(t) - \Phi_a(t - 2T_c) \\ &= G \cdot (P_r \otimes (M(t) - M(t - 2T_c)).\end{aligned}$$

  Since $M(t)$ is a binary signal having values of ± 1, the difference results in a reference signal that has five possible values 0, ± 1 and ± 2. Using pixels with five integration nodes, two subsequently acquired samples of the discriminator function are sufficient for its reconstruction. Each sample of the discriminator function is obtained by the sum of the integration values that are weighted each by one of the factors. Hence, the discriminator function can be exploited for the measurement of distances, where the maximum detectable range has been increased by a factor of 2, or the stochastic accuracy has become better when a shorter peak width has been used.

- *Signal combinations - PN and sinusoid:* Sinusoidal modulation shows the best performance in a single-camera environment. Hence, it is suggested that sinusoidal modulation be used as the basic scheme in combination with a technique that ensures low cross-correlation values with other signals, such as pseudo-noise. Basically, two parameters of a sinusoidal signal can be changed: the phase and the modulation frequency. If these parameters are randomly changed, both for the optical and the reference signal of one camera, the time delay as an informa-

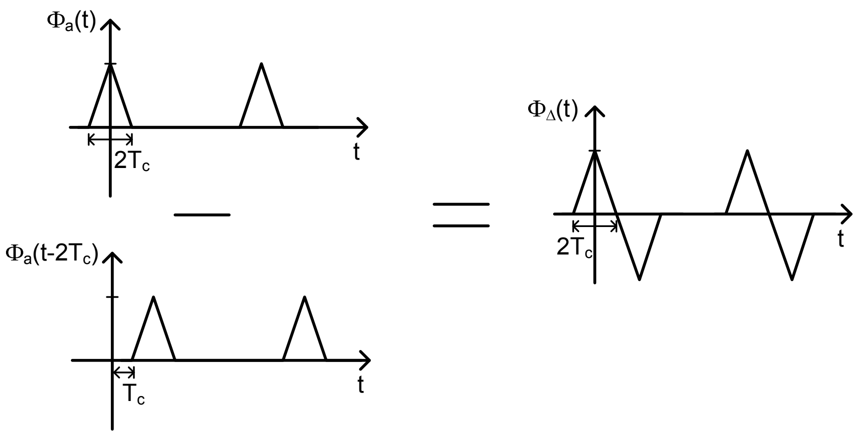


Figure 3.35: Construction of the discriminator function by the autocorrelation function and its time-shifted version.

tion for the distance to an object can still be extracted when the changes in phase between the two cameras are different.

- *Frequency chirping:* Chirp signals, which make use of sinusoidal waveforms with frequencies changing over time, combine the properties of sinusoidal signals with good distinguishability between other chirps. The rule of frequency sweep defines the multi-camera capability. Examples from mobile communication, where chirps are used in multi-user environments, are found in [35, 47]. Elaborations of either the heterodyne mixed signal or of the correlation function are possible methods for distance acquisition. However, a modification of the pixel itself and an adaptation of the distance extraction scheme probably become necessary.

Chapter 4

Fundamentals of Electronic Imaging

This chapter summarises those aspects of electronic imaging which are of particular importance for the TOF pixels: the electron-hole pair generation and separation, the fundamental image sensor technologies, the charge-coupled device concept for the charge storage and lateral transport, and the quantum efficiency of the MOS-gate structure. A general depiction of the photosensing chain is shown in figure 4.1 including the light source, the object, the imaging optics, an optical information processing unit, the detector array that converts the photons into an electric charge, a digital or analogue process sequence and the final units for displaying, interpreting or broadcasting the image. A digital conditioning of the signal comprises the analogue-to-digital (AD) conversion and a digital signal processing (DSP) unit, while an analogue path consists of analogue processing and storage units.

In the case of the 3D TOF system, the light source emits an intensity-modulated light signal, and the detector performs both the detection and the demodulation of the signal at the same time. While the subsequent chapter is dedicated to the pixel-inherent de-

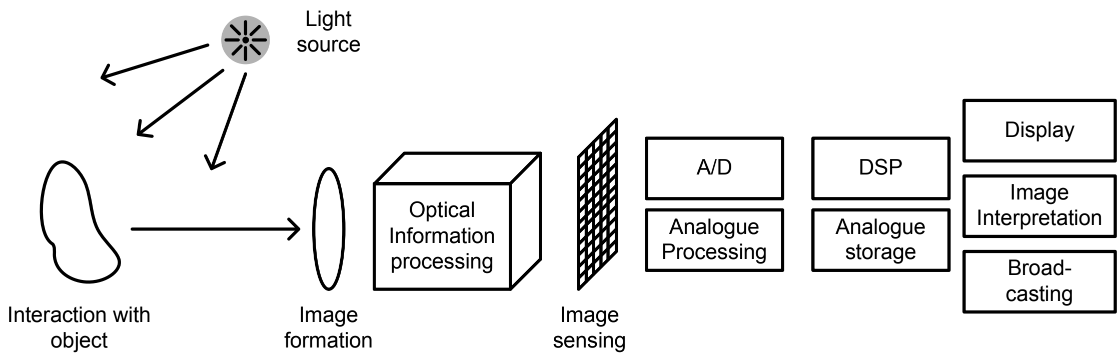


Figure 4.1: Electronic imaging chain. Originally extracted from [95] and subsequently modified.

modulation of the signal, the present chapter focuses on the general photon detection behaviour of solid-state sensors. This comprises the fundamental existing pixel and sensor architectures.

The first section presents the basic physical phenomena exploited in solid-state photosensing. Then the fundamental pixel concepts of recent sensor designs are explained. A brief discussion about the available standard imaging technologies is given. After that, the concept of charge-coupled devices is explained, covering the working principles of CCDs with regard to the storage and transport of photo-generated charges and the major types of transport channels. The subsequent section highlights the quantum efficiency of CCD gate structures, which is an important criteria for the process and design optimization of the TOF pixels in order to achieve low distance standard deviations (see chapter 2 and 3).

4.1 Photon Detection

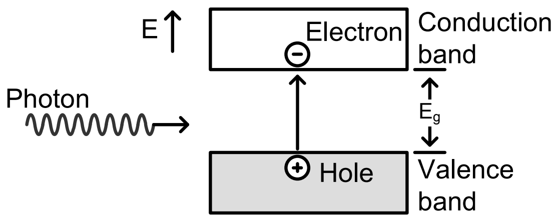
The detection of photons exploits the internal photo-electric effect of a semiconductor, which means the transmission of optical energy from impinging photons to the valence electrons of the material by exciting the electrons to a higher energy level within the conduction band. This situation of electron-hole pair generation in pure semiconductors is shown in figure 4.2 (a). The transition of an electron from the valence band to the conduction band is called an intrinsic transition [105]. Such an event as the consequence of the detection of a photon requires, from the energy point of view, the condition

$$E_{ph} = h\frac{c}{\lambda} > E_g, \tag{4.1}$$

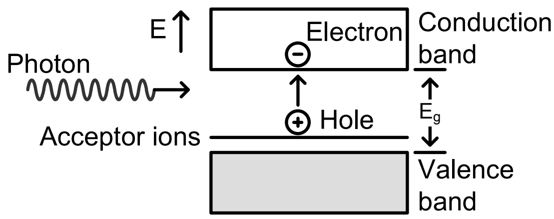
where E_{ph} is the energy of the photon, and the bandgap energy of the material is denoted by E_g, which is about 1.12eV for pure silicon material at room temperature. Following from this a maximum optical wavelength λ_c can be defined that determines the highest wavelength for which electron-hole pairs are still generated:

$$\lambda_c = \frac{hc}{E_g} \approx \frac{1.24\mu\mathrm{m}}{E_g[\mathrm{eV}]}, \tag{4.2}$$

where E_g is given in units of eV. The cut-off wavelength of silicon is about 1.1μm. Silicon is not photo-sensitive for larger wavelengths. However, this is only an approximation because silicon is an indirect semiconductor which means that the event of absorbing a photon requires both the transition of the bandgap as well as an adaptation of the electron's momentum [86]. Consequently, the required optical energy for the generation of an electron-hole pair is slightly increased and hence, the cut-off wavelength is



(a) Intrinsic transition.



(b) Extrinsic transition for the example of p-doped semiconductors. The acceptor ions provide an additional energy state close to the upper energy level of the valence band.

Figure 4.2: Internal photo-electric effect.

slightly smaller.
If the semiconductor is doped, the required energy for the generation of electron-hole pairs is reduced because additional energy states are present in the otherwise ”forbidden” energy zone between the valence and conduction bands. This transition is called an extrinsic transition and it is illustrated in figure 4.2 (b) [105].

Optical energies larger than the bandgap energy E_g generate electron-hole pairs, and, in addition to that, the excess energy $hc/\lambda - E_g$ is dissipated as heat. If energies are even larger (several eV), multiple electron-hole pairs are generated per incident high-energy photon, as exploited for example in X-ray detection. An upper bound of optical energy is mainly given by the absorbing nature of additional layers on top of the semiconductor device and surface effects such as very short carrier lifetimes [95]. This is discussed in more detail in section 4.4, treating of the quantum efficiency of the MOS-gate structure.

4.1.1 Recombination and Lifetime of Charge Carriers

After electron-hole pairs have been generated in the semiconductor by the incident photons, the charge pairs need to be separated in space in order to avoid recombination and the corresponding loss of information. The time that is used for separating electrons and holes is limited by the finite lifetime of the free charge carriers and depends on the dominant mechanism of recombination.

The effective and measurable lifetime is composed of the three recombination mechanisms: Shockley-Read-Hall (SRH) recombination [100], radiative recombination [105] and Auger recombination [89,105]. It is expressed as [89,90]

$$\tau_r = \frac{1}{\tau_{SRH}^{-1} + \tau_{rad}^{-1} + \tau_{Auger}^{-1}}, \tag{4.3}$$

where τ_{SRH}, τ_{rad} and τ_{Auger} denote the Shockley-Read-Hall lifetime, the radiative lifetime and the Auger lifetime, respectively. Figure 4.3, which has been extracted from [90], depicts measured recombination lifetimes versus the majority carrier density of p-type silicon. Two measurement data sequences have been incorporated and fitted with the mathematical expressions that can be found in literature, describing the particular recombination mechanism [89,90,100,105]. The SRH recombination depends on several material parameters, such as density of impurities or defects and the capture cross sections of electrons and holes. Depending on the specific piece of semiconductor, the SRH lifetime can vary very strongly. Figure 4.3 shows the characteristics of the lifetime for different orders of SRH lifetimes.
Obviously, the lifetime in silicon is controlled by Auger recombination at high carrier densities, while for low carrier densities the Shockley-Read-Hall mechanism dominates. Radiative recombination is relatively unimportant in silicon because direct band to band transition is unlikely for indirect-bandgap semiconductors. This is also the reason why silicon material is not exploited in devices for the emission of optical signals.

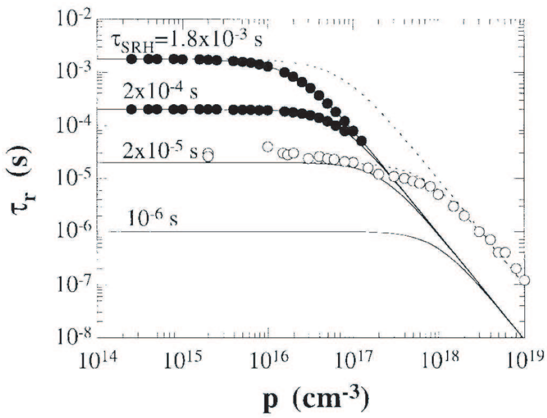


Figure 4.3: Recombination lifetime τ_r versus majority carrier density for p-type Si. The solid and dashed lines are based on theory using different proportionality constants. [90]

4.1.2 Thermal Diffusion

As long as no electric fields force the separation of photo-generated electrons and holes in space, these charge carriers will randomly move around, driven by their thermal energy. The average thermal velocity is deduced by equating the kinetic energy of the charge with the thermal energy, according to the equipartition theorem of the energy

$$\frac{1}{2}m_{eff}v_{th}^2 = \frac{3}{2}K_B\vartheta, \tag{4.4}$$

where K_B is Boltzmann's constant, ϑ is the absolute temperature and m_{eff} denotes the effective mass of the charge. The thermal velocity therefore is

$$v_{th} = \sqrt{\frac{3K_B\vartheta}{m_{eff}}}. \tag{4.5}$$

Statistically, the charge carriers recombine after a time corresponding to their lifetime τ_r. Until then they have travelled on average a certain distance called diffusion length L_{diff}, which is expressed as

$$L_{diff} = \sqrt{D \cdot \tau_r}, \tag{4.6}$$

with the diffusivity D. This parameter strongly depends on the temperature, doping concentration and whether an n-type or p-type semiconductor is used. This is obvious by considering the following expression of the diffusivity

$$D = \frac{1}{3}v_{th}^2\tau_c = \frac{K\vartheta}{m_{eff}} \cdot \tau_c. \tag{4.7}$$

Besides the temperature ϑ, the mean free time τ_c is included in the last equation, describing the mean time between two collisions of the charge carrier with the lattice or other carriers. The mean free time τ_c strongly depends on the temperature and doping concentration as well [105]. For p-doped silicon with low doping concentration less than $10^{16}\mathrm{cm}^{-3}$, the diffusivity is about $D \approx 37\mathrm{cm}^2/\mathrm{s}$ at a temperature $\vartheta = 300\mathrm{K}$.

The recombination lifetime respectively the diffusion length are important parameters of photo-sensitive devices. As shown in [103], the diffusion length has a direct impact on the modulation transfer function (MTF) [106, 117] of a sensor and the quantum efficiency (this will be discussed later in section 4.4). While the MTF suffers from long diffusion lengths, the quantum efficiency, in turn, improves. Since the feature size of opto-electronic processes determines the substrate doping concentration, the lifetime is also directly linked to the feature size. Processes of feature sizes less than $0.6\mu\mathrm{m}$ generally offer lifetimes of the order of microseconds [103, 119] or even up to three orders of magnitude lower, strongly depending on the defect density of the material [91]. The lifetime of one microsecond corresponds to a diffusion length of $61\mu\mathrm{m}$.

4.1.3 Charge Separation

In order to avoid the loss of information due to recombination of the photo-generated electron-hole pairs before they can be read out, special regions for the separation and collection have to be created in the semiconductor.

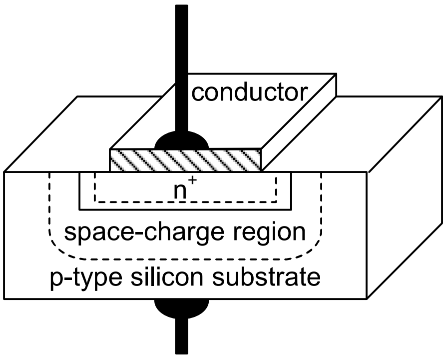
The electrons and holes are separated by the presence of electric fields. Two different design concepts for the external control of the electric fields are used in solid-state photo-sensor designs:

Reverse-biased photodiode: The basic set-up of a photodiode is shown in figure 4.4 (a), consisting of the pn-junction between the p-doped silicon substrate and a heavily n-doped implant at the top. When the diode has been reverse biased, a space charge region of width w (also called depletion region) is formed:

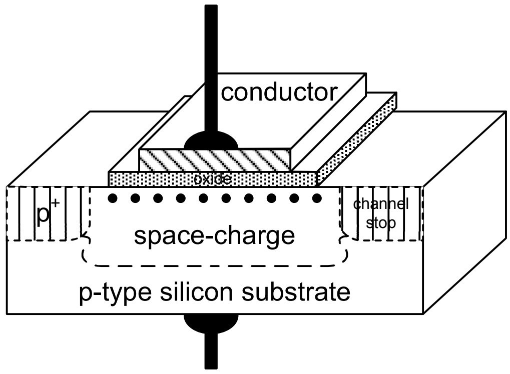
$$w = \sqrt{\frac{2\epsilon_0\epsilon_s}{qN_{low}} \cdot (V_{bi} + V_D)}, \tag{4.8}$$

with the permittivity ϵ_0, the dielectric constant ϵ_s of the semiconductor material, the built-in voltage V_{bi} of the junction, the reverse bias voltage V_D and the doping concentration N_{low} of the lower doped region. In this model a one-sided abrupt junction has been assumed. [105] Photo-generated electron-hole pairs in the space charge region are directly affected by the electric field, which is formed by the space charge itself. While the holes are moved to the substrate (in the case of a p-doped substrate), the electrons will drift to the n^+-implant. Depending on the particular pixel design, the electrons are stored in the photodiode's inherent depletion capacitance or in an external capacitance for read-out after a certain integration time; alternatively the photo-current is directly used for further processing.

Metal-oxide-semiconductor (MOS)-gate: Figure 4.4 (b) shows the second major type of electrical field generating devices, which is used for the separation of electron-hole pairs. It consists of a transparent gate, which is normally made of poly-silicon material, a dielectric insulator and the substrate below. By applying a voltage V_g to the gate (in the case of a p-doped substrate, the voltage has to be positive), a space charge region is created. The electric field in the space charge region is again exploited for the charge separation. The storage is performed by the same device. In the case of p-doped substrate, the photo-generated charge carriers that are accumulated at the surface of the semiconductor bulk are electrons instead of holes. Since this semiconductor device is the fundamental unit of charge-coupled devices (CCDs), more details about the exact position of the charge storage as well as the depletion width and transport channel location are provided in section 4.3.



(a) Photodiode, consisting of a reverse-biased pn-junction.



(b) MOS capacitance, consisting of a transparent electrode gate (generally poly-silicon material) on top of an dielectric insulator.

Figure 4.4: Two major types of semiconductor devices that are used for the separation and storage of the photo-generated charges. [95]

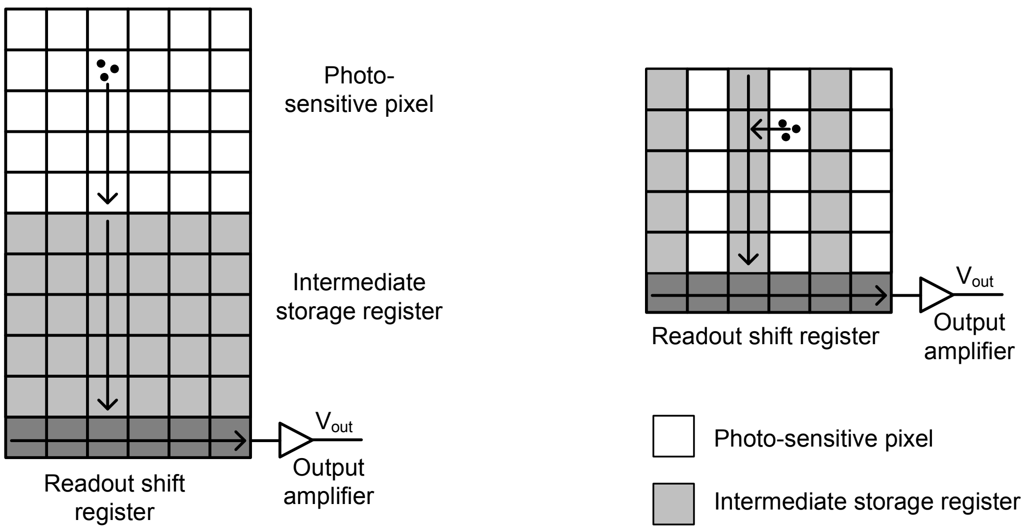
The detection of photo-generated charge carriers is not exclusively limited to the charges that are generated in the space charge region. Other carriers, which have been generated outside the space charge region, are not directly influenced by an electric field, but the thermal motion can let them diffuse into the space charge region, too, where they are affected by the electric field as well and hence, they can be collected and read-out.

4.2 Image Sensor Technology

Standard semiconductor process technologies experience a constant reduction of the feature sizes due to the great progression in lithography technology. This reduction enables circuitries of even higher density, which is exploited by image sensors. Two main technologies have to be distinguished for image sensors: historically, charge-coupled device (CCD) sensors formed the first electronic imaging chips; sensors in complementary-metal-oxide-semiconductor (CMOS) technology form the second major type of image sensors. A brief discussion of these two technologies follows.

4.2.1 CCD Image Sensors

The first major technology of solid-state sensors exploits charge-coupled devices (CCDs). The basic idea of these semiconductor devices is the storage and transport of charge packets using MOS gate structures. Several adjacent MOS-gates form a transport line for the charge packets, necessitating an appropriate synchronised control of the gate voltages. The charge transfer between the MOS gates benefits from overlapping poly-silicon layers, which are characteristic for recent CCD-processes. Also special channel implants improving the charge transport significantly are CCD-typical features. The physical description of how the charge is stored and transported through a CCD register is explained in the next section in more detail.
CCDs were invented by Boyle and Smith in 1970 [20]. They had already had the idea of using these devices in a wide field of applications such as shift registers, imaging devices, display devices and logic performance [20]. The practical proof of the CCD principle to store and transport single charge packets was accomplished in the same year [10]. In particular, information processing in the charge domain was expected to be very noise-efficient. In 1975, Mohsen et. al [69] published work about the investigation of CCD noise performance, where they could prove the operation of photon-detecting CCDs to be at the physical limitation given by photon shot noise.
Figure 4.5 shows two examples of the realisation of CCD sensors. Figure (a) shows the



(a) Frame transfer readout. **(b)** Interline transfer readout.

Figure 4.5: CCD sensor architectures.

arrangement of a frame transfer CCD sensor, where the pixel field is located at the top. After the integration process has finished, the charge is shifted into the non-sensitive intermediate storage register in order to avoid additional integration of photo-charges. The pixels are read out via the readout shift register line by line. Only one single stage for the final charge to voltage conversion is employed. In figure 4.5 (b) the non-sensitive intermediate storage registers are located beside each photo-sensitive pixel-column: this architecture is thus called interline-transfer CCD. The number of shifts for reading out the charge packets is reduced but the fill-factor of each pixel is less than 1. Another popular concept is the field-interline-transfer CCD, which is a combination of the two depicted sensor architectures. All CCD sensors usually use only one amplifier for the readout. More details about these architectures can be found for example in [95, 106].

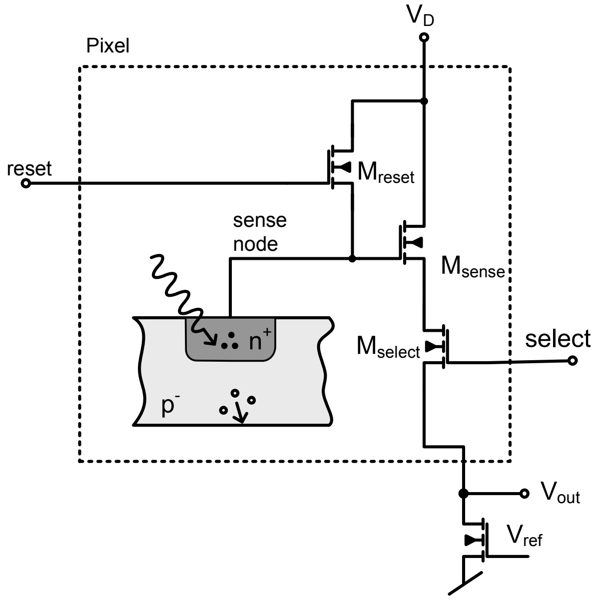
The important advantages of CCD sensors over CMOS sensors are the full operation in the low-noise charge domain during the optical detection and charge transport, as well as the low-noise contribution of only one single output amplifier, ensuring a low level fixed pattern noise due to process variations.

4.2.2 CMOS Image Sensors

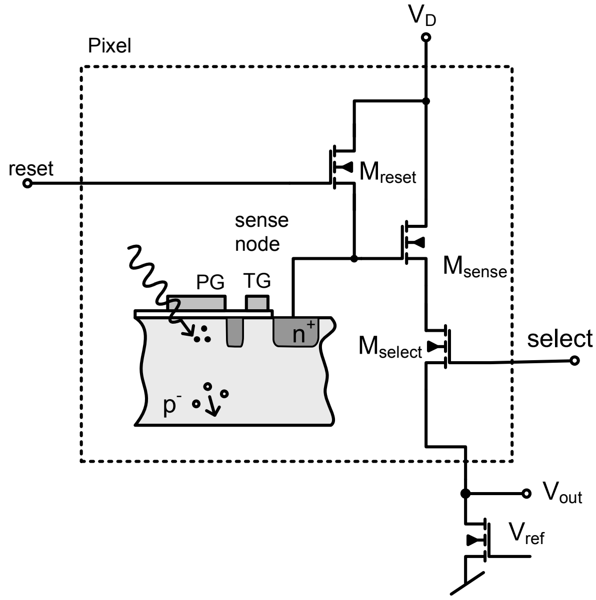
Passive CMOS pixel: The first realization of photo-sensors in standard CMOS technology consisted of an array of conventional photodiodes. These photodiodes were operated in reverse-bias mode. Under illumination, the photo-current was integrated on the photodiode's inherent capacitance. By activating a select transistor, the diode was directly connected to an amplifier common to one column or even to the whole sensor. The charge, which is a measure of the optical intensity, was read out by a charge amplifier. The basic problem of these first implementations was the high capacitance of the conduction lines between the pixel and the output amplifier, leading to very low voltage swings at the output. Hence, large noise contributions and low read-out frequencies were the consequences.

Fundamental active CMOS pixels: The improvement of CMOS-based pixels led to the invention of the CMOS active-pixel sensor (CMOS APS) that overcomes the problem of the large noise contribution due to the large capacitance of the readout line. Each pixel incorporates its own amplification stage. On the basis of some fundamental, today widely-used CMOS pixel structures, the concept of active-pixels is shown in figure 4.6. All pixels incorporate a select transistor $\mathrm{M}_{\mathrm{select}}$ and a sense transistor $\mathrm{M}_{\mathrm{sense}}$ as part of a standard source follower amplification stage. Those pixel structures, which are based on the integration of photo-generated charges and subsequent readout, need another transistor $\mathrm{M}_{\mathrm{reset}}$ that is used to reset the pixel. Other structures that measure the photocurrent directly do not need this reset transistor. The source follower is completed by an active load transistor, which does not need to be included in the pixel, moreover it can be shared by all pixels in a column.

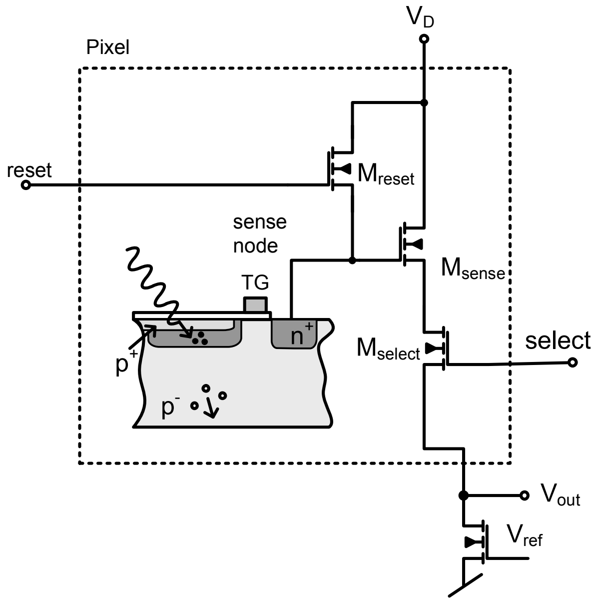
The basic photodetection process of the photodiode active pixel shown in figure 4.6 (a) is the same as in the case of the passive pixel architecture. After having reset the pn-junction to the reverse-bias voltage V_D by activating the reset transistor, the photocurrent discharges the photodiode's capacitance and the gate capacitance of the sense transistor. For readout purpose, the resulting voltage is intra-pixel buffered by the source follower circuit, which is capable of driving the large capacitance of the conduction line to the output amplifier (not shown in the figure 4.6).
Figure 4.6 (b) shows the principle of the photogate APS. There, the photo-generated charge is collected below the photogate PG. In order to read out the charge carriers, they are transferred via the transfer gate TG to the floating diffusion, where a conversion into a voltage and a subsequent amplification are performed. The transfer process is highly critical due to the implant region between the two gates PG and TG. This implant is the result of the source-drain diffusion step in standard CMOS processes. It leads to charge leakage, which implies image lag and excess noise [65, 121]. To minimize this negative impact on the image quality, special reset or timing sequences have been applied in the past [81, 121] or special structures with overlapping gates are introduced.
Another photodetection element is the pinned photodiode, which is shown in figure 4.6 (c). [41, 65] This kind of photodiodes consists of a p^+np^- structure where both p layers are at substrate potential. If no charge carriers have been accumulated, the device has been designed in such a way that the depletion layers of the two pn-junctions meet each other and create a potential in the n-layer that is pinned to a fixed value called pinned voltage. [65] During the exposure of the device charge carriers are stored in the depletion region, decreasing the potential of the photodiode below the pinned voltage. By activating the transfer gate TG, the accumulated charges are transferred to the sensing diffusion for subsequent read-out and the photodiode adopts automatically again its pinned voltage. While the full well of this device is reduced compared to the photodiode and photogate pixels described before, the dark current is advantageously reduced and the charge transfer to the read out node is improved, both due to the fact that the charge storage and transport are performed deeply in the substrate. Today's smallest pixels are typically based on pinned photodiodes.
The pixel structure in figure 4.6 (d) again makes use of a simple pn-junction-based diode as photodetection element. However, instead of implementing a reset transistor, which expects an externally fed reset signal, the transistor $\mathrm{M}_{\mathrm{log}}$ is operated in diode



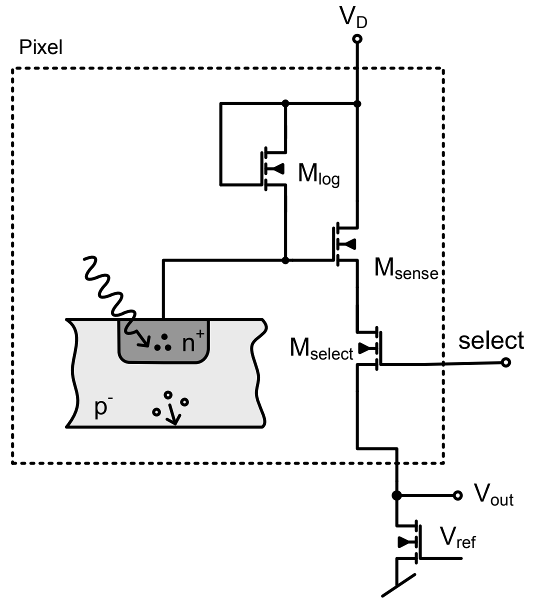
(a) Photodiode CMOS APS with 3 transistors inside the pixel.



(b) Photogate CMOS APS with 4 transistor elements inside the pixel.



(c) Pinned photodiode CMOS APS using 4 transistors inside the pixel.



(d) Logarithmic CMOS pixel providing almost instantaneous readout capability.

Figure 4.6: CMOS active pixel structures using fundamental photodetection elements.

configuration. Thus, the photo-generated current will be converted into a voltage signal following a logarithmic characteristics. This logarithmic compression of the photocurrent's value enables the acquisition of scenes where the dynamic range, which is the ratio between the darkest and brightest points in the scene, exceeds high values. Sensors covering dynamic ranges of 120dB and more have been reported [95]. Since no charge integration is performed by this pixel circuitry but rather the voltage at the drain of the logarithmic transistor element is measured as a result of the photocurrent, the pixel reacts quickly to changes in the light conditions. However, also this pixel structure shows a characteristic time-behavior, which is determined by the particular resistivity and capacitances of the transistor-diode circuitry. Under low light conditions it will take longer until the transistor reaches its equilibrium state than under bright illumination conditions. [95]

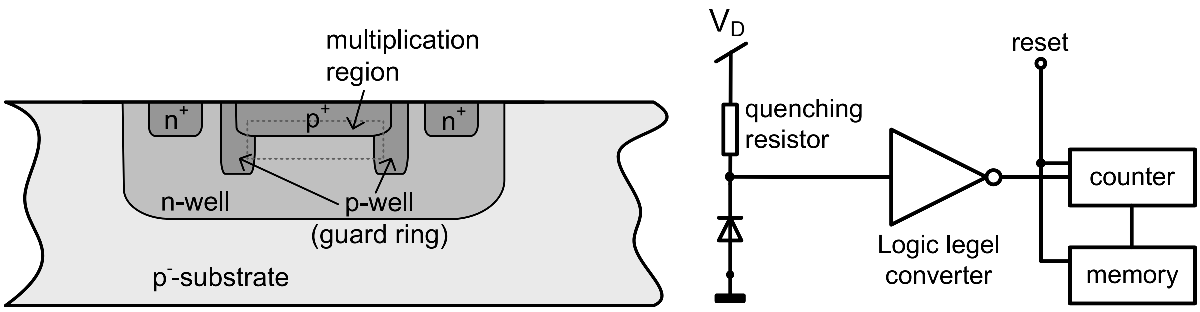
Active CMOS pixels with high-gain photodetection element: In the standard mode of operation of photodiodes, small reverse bias voltages of just a few Volts are used so that a linear behavior between optical incident power and a non-amplified photocurrent is obtained. However, the device's characteristics change when the reverse bias voltage approaches the breakdown voltage. The photocurrent starts to be amplified significantly due to impact ionization in the strong electric field within the depletion region by some of the photo-generated charge carriers. This enables the detection of even very low optical signals, whose intensities are not acquired in normal mode of operation with low reverse biasing due to the dominance of readout noise sources over the non-amplified photocurrent signal. Thus, avalanche photodiodes provide an alternative method for external gain mechanisms such as for example photon multiplier tubes or microchannel plates. [16]

The relationship between optical power and photocurrent stays linear as long as the bias voltage does not exceed the breakdown voltage, which is strongly dependent on the doping concentrations and on the geometrical design of the particular device. If the breakdown voltage is exceeded, the photocurrent experiences an almost infinite amplification (factors of 10^8 have been cited in [16]) and the current continuous to flow. This mode of operation is called Geiger mode of operation, while the sub-Geiger or linear mode of operation means a reverse bias voltage lower than the breakdown voltage but high enough to create an appreciable amplification effect of the photocurrent [16, 29]. Obviously, the Geiger mode of operation allows for detecting the arrival of single photons in a digital fashion and, hence, the device is called single photon avalanche diode (SPAD). The sub-Geiger mode enables high-gain photodetection in the analogue domain. [16]

Figure 4.7 (a) shows the cross section of an avalanche photodiode in CMOS technology

based on a slightly p-doped substrate. The high reverse bias voltage is applied to the p^+-region. In order to avoid premature breakdown of the device, which would make any photon detection impossible, a p-well guard ring has to be created. [16, 29, 31, 122, 123] It is noted that although such a pixel structure requires a high-voltage option of the particular CMOS process, it is regarded, nevertheless, as being based on standard CMOS process technology since high-voltage options are provided by many CMOS processes in recent days. [29]
In order to avoid device destruction, the high avalanche current that is continuously flowing through the SPAD once a detection event occured, has to be interrupted. The interruption is performed by a so-called "quenching circuit" that pulls the bias voltage below the breakdown voltage. Passive and active circuits are known. The simplest case of a passive circuit consists of a series resistance to the photodiode. Deeper analysis on quenching circuitries are provided in [31] as well as an accurate discrete modelling of avalanche photodiodes leading to an explanation of their time behavior in the different modes of operation.
The SPAD pixel set-up in figure 4.7 (b) uses a single resistor as quenching element. Additional electronics is needed to convert the avalanche current pulse into a voltage pulse of standard logic levels. Optionally, the pixel needs a digital or analogue counter and a memory circuitry. The pixel complexity increases accordingly. The counter can be used to count either the number of digital pulses as a measure of the light intensity or the time of photon arrival with high time resolution. Obviously, the readout of the SPADs is not trivial prohibiting SPADs in many applications to be used in a matrix style. Clever readout schemes exist but still suffer from significant reduction of the frame rates [29, 95]. In contrast to that, avalanche diodes that provide an analogue signal when being operated in the sub-Geiger mode, can use in-pixel readout electronics as shown in the figures 4.6 (a)-(d).
Beside the advantage of high timing resolution of SPADs, disadvantages include significant thermal impact on the breakdown voltage and hence on the timing characteristics, the power dissipation due to the avalanche current itself (the instantaneous pulse can attain watts of power [31]) and the optical cross talk due to optical emission in the diode structure itself. [16, 122]

CMOS sensor architecture: CMOS sensors employ a second and a third amplification stage outside the pixel in order to drive the off-chip capacitive load. Considering the analogue pixel structures, special design focus lies on the compensation of non-linearities of the in-pixel source follower.
Figure 4.8 depicts the general architecture of a CMOS active pixel sensor with its main components: the array of photo-sensitive pixels and the decoders for the row and col-



(a) Cross section through an avalanche photodiode in CMOS technology based on p^--substrate. [16, 72]

(b) SPAD digital pixel using a passive quenching circuit.

Figure 4.7: CMOS active pixel based on a high-gain photodetection element.

umn select. The pixel-wise selection for reading out the photo-generated charge enables a random access of pixels; in particular, so-called regions of interest (ROI) can be specified for readout. This flexibility is a huge advantage of CMOS sensors compared with CCD sensor concepts. However, the fact that each pixel has its own amplification stage leads to significant fixed pattern noise (FPN) due to spatial variations of oxide thicknesses, gate sizes and doping concentrations. More detailed information about process mismatches of MOS transistors, their modelling and references to process-related investigations can be found for example in [34, 83]. With regard to image sensors, the process variations appear as different offsets and gains between the pixel's source followers. FPN may necessitate the application of correction methods such as double sampling.

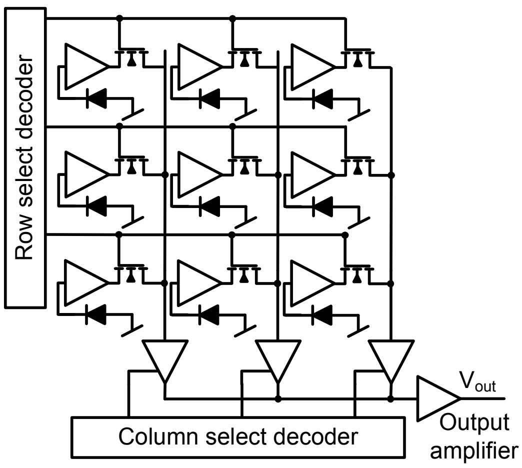


Figure 4.8: General CMOS sensor architecture.

4.2.3 The Enduring Discussion - CCD versus CMOS

There has been an on-going heated discussion during the last few years about what kind of technology will be the emerging one in solid-state image sensing: CMOS active-pixel sensors or CCDs? [40, 42, 64, 95, 119]

While CCDs have significant advantages in low-noise and high-sensitive image detection, CMOS technology is superior for solving specialized tasks that require additional on-chip or in-pixel functionality [97, 119]. Example features of CMOS sensors are high-speed sensors with frame rates up to several thousands frames per second [57, 61], high-dynamic range pixels with non-linear response [112] or a pixel-wise controlled integration time, pixels with programmable offsets and gains for a flexible adaptation of the sensor operation to the scene's optical conditions [112], global shutter transistors in the pixels, analogue-to-digital conversion already at pixel-level [120], vision pixels for human eye implants that convert an optical intensity into a sequence of electric pulses [80], in-pixel signal processing such as for example the demodulation of an optical interferometer signal [12], vision sensors performing image processing tasks such as edge detection or contrast evaluation already at pixel level [84], in-pixel noise reduction with correlated double sampling stages [68].
These additional features enable a lot of new applications in the fields of automotive, security, bio-medical, industry, and more. In particular, very cost-efficient consumer-market oriented single-chip solutions, so-called system-on-chip (SOC) or camera-on-chip [40, 42], become realizable. However, the optical fill-factor of CMOS pixels suffers from the high-density integration, and electro-magnetic interference become an important issue in the chip design. Consequently, CMOS image sensors profit from smaller feature sizes with regard to the pixel's fill-factor but at the same time smaller feature sizes reduce the performance in optical detection and make the application of pure CMOS processes with feature sizes below $0.25\mu\mathrm{m}$ [107, 119] more difficult.

While the CCD technology has experienced a straight development path, always focussing on processing methods that are optimized for electronic photon detection, CMOS technology finds itself on a "winding road" [64]. On the one hand it provides the important advantage of integrating additional functionality on-chip or in the pixel; on the other hand, the decreasing feature sizes do not automatically bring the necessary properties for optical imaging, such as for example deep submicron depletion regions [119]. CMOS processes are rather developed with regard to cost-efficiency and optimized for logic and analogue circuitry [42]. This often makes the usage of modified CMOS processes or innovative pixel structures indispensable.

The TOF pixels of the current work try to exploit the advantages of both CMOS and

CCD technologies. The applied process is a $0.8\mu m$ CMOS process with two poly-silicon and two metal layers or a $0.6\mu m$ process with two poly silicon and three metal layers. Additionally, a CCD option is provided by the foundry, meaning overlapping gate structures and an additional doping implant. The detection and demodulation are performed with charge-coupled devices in order to keep the noise during these computational steps as low as possible. The readout of the pixel is accomplished via standard source followers as it is typical for CMOS sensors. It ensures a flexible readout scheme such as for example the definition of regions of interests. The TOF pixels having background light suppression capability (see section 5.2) necessitate for example a line-by-line readout scheme in order to work properly.

4.3 Charge-Coupled Devices

Charge-coupled devices are formed by the adjacent fabrication of MOS-gates. By applying appropriate voltages to the gates, photo-generated charge can be stored. A lateral transport becomes possible due to the direct coupling between the gates. The coupling strongly benefits from the overlapping of the gate structures, which is typical for CCD processes.
The storage capability is discussed in the first subsection, where it is differentiated between two types of CCDs having either a surface or a buried transport channel. The transport of charges from one gate to another is treated in the second subsection, highlighting the principles of diffusion-limited charge transport in silicon and pointing out the resulting limitations on the charge transport speed.
For even more details about the characteristics of CCDs, the interested reader is advised to take a look at some of the many publications that exist about CCD devices. Detailed device descriptions can be found for example in [15, 52, 98, 105, 106, 110].

4.3.1 Storage of Charge Carriers

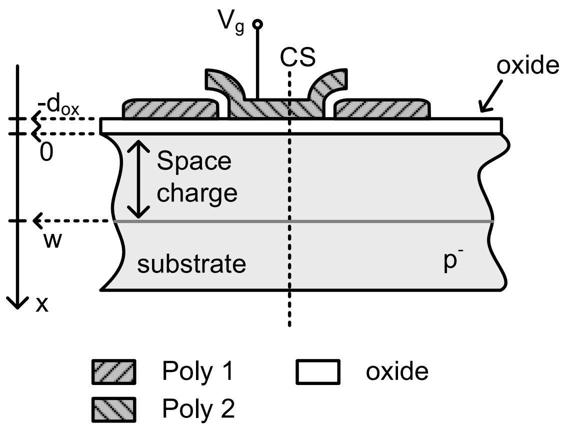
Figures 4.9 (a) and (b) show the two types of CCDs. While the gate arrangement is the same, the transport channels are different. The left figure does not have a doping implant at the surface of the substrate, so that any photo-generated charges are stored and transported at the substrate's surface. Therefore, this CCD set-up is called a surface channel CCD (S-CCD). The CCD in figure (b) has a buried channel (B-CCD) incorporated. The additional doping implant that needs to be depleted completely for proper operation leads to a shift of the storage and transport region from the surface into the substrate by a few micrometers. This has several advantages with regard to the sensitivity and the noise performance of the CCD because the photo-sensitive depletion width is enhanced into the semiconductor and the trapping noise is much less significant in the bulk than at the surface. Furthermore the fringing fields are more pronounced in buried channel devices.

In order to understand the storage of the charge carriers in the particular CCD elements, the voltage distributions within the device for given conditions, such as voltage settings at the gate and the semiconductor bulk, layer thicknesses and doping concentrations, are of essential interest. Along the vertical cross-sections CS depicted in figure 4.9, analytical expressions for the one-dimensional potential distribution can be derived for the CCD-specific mode of operation. Because the realized TOF sensors are based on p-doped silicon substrate, all considerations are made with this kind of substrate doping.

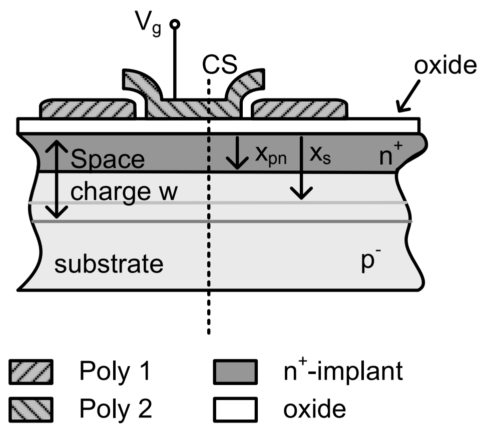
For the derivation of analytical expressions of the potential distribution in the CCD device, it is necessary to solve Poisson's equation and reasonably determine the boundary conditions. The Poisson equation describes the relationship between the potential distribution and the charge density in three-dimensional space. Its one-dimensional simplified form looks as follows:

$$\frac{\partial^2 \Psi(x)}{\partial x^2} = -\frac{\rho_q(x)}{\epsilon\epsilon_0}, \tag{4.9}$$

where $\Psi(x)$ is the potential at the location x, $\rho_q(x)$ denotes the charge density, ϵ is the dielectric constant of the material and ϵ_0 is the permittivity of free space. The definition of the x-axes is depicted in figure 4.9. The origin is situated at the semiconductor



(a) Surface channel CCD.



(b) Buried channel CCD.

Figure 4.9: Two major types of CCDs.

insulator interface.

For convenience, all following considerations are related to the effective gate voltage V_{eff}, which is defined as the difference between the gate voltage V_g and the flat band voltage V_{FB}:

$$V_{eff} = V_g - V_{FB}. \tag{4.10}$$

Here, the flat band voltage describes the voltage that is needed to obtain a flat band condition when considering the band diagram of the CCD [106]. V_{FB} is generally unequal zero due to two effects: 1. There is a difference in the work functions Ψ_{ms} between the polycrystalline-silicon gate material and the monocrystalline-silicon bulk material. 2. There are positive and fixed oxide charges Q_{ox}, generally represented at the silicon-silicon oxide interface. The amount of voltage drop for compensating the oxide charge is $-Q_{ox}/C_{ox}$, where C_{ox} denotes the oxide capacitance. The flat band voltage results in

$$V_{FB} = \Psi_{ms} - \frac{Q_{ox}}{C_{ox}}, \tag{4.11}$$

with

$$C_{ox} = \frac{\epsilon_{ox}\epsilon_0}{d_{ox}}. \tag{4.12}$$

Both, the charge and capacitance parameters are related to unit area. d_{ox} is the thickness of the insulator region between the poly-silicon gate and the semiconductor bulk (see figure 4.9).

Surface Channel

In order to collect minority carriers below the CCD gate, the device must be operated with positive effective gate voltages $V_{eff} > 0$. If this were not the case, the device would accumulate majority carriers at the semiconductor surface.
Positive effective gate voltages create a depletion region of width w, reaching from the semiconductor surface into the bulk. According to the model of a one-sided abrupt pn-junction, the depletion width is similarly expressed as [106]

$$w = \sqrt{\frac{2\epsilon_0\epsilon_{si}\Psi_s}{qN_A}}, \tag{4.13}$$

based on the surface potential Ψ_s. N_A denotes the acceptor concentration of the substrate. The surface potential is given by [15, 106]

$$\Psi_s = V_{eff} + \frac{Q_n}{C_{ox}} + \left[\frac{\epsilon_0\epsilon_{si}qN_A}{C_{ox}^2} \cdot \left(1 - \sqrt{1 + \frac{2C_{ox}^2 \cdot \left(V_{eff} + \frac{Q_n}{C_{ox}}\right)}{\epsilon_0\epsilon_{si}qN_A}}\right)\right]. \tag{4.14}$$

The number of collected electrons is denoted by Q_n. If there are no free electrons present for the compensation of the positive gate voltage, the depletion region contains only fixed negative acceptor ions. This mode is called deep depletion mode, which is a non-equilibrium state. As soon as free electrons are available, for example due to optical or thermal generation, these minority carriers will create an inversion layer at the surface. According to the equations 4.13 and 4.14, the amount of collected charge directly affects both the surface potential as well as the depletion region. The larger the number of collected minority charge carriers, the lower will be the surface potential and the more decreases the depletion region. The last relation is obvious, because fewer acceptor ions are necessary to compensate the positive charge on the gate. In strong inversion, the minority carriers are stored in a very thin layer just at the surface of the semiconductor bulk. The width of this layer typically ranges from 1 to 10nm [105].

The one-dimensional potential distribution through the cross-section CS (see figure 4.9) is split into three regions:

1. *Oxide region* ($-d_{ox} < x < 0$): The potential falls linearly from V_{eff} down to the surface potential Ψ_s, corresponding to

$$\Psi(x) = V_{eff} - \frac{V_{eff} - \Psi_s}{d_{ox}} \cdot (x + d_{ox}). \tag{4.15}$$

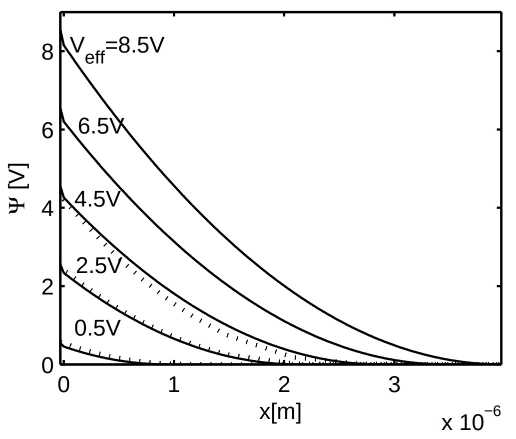
2. *Depletion region* ($0 < x < w$): Since a constant charge distribution is present, the potential decreases quadratically, according to the solution of Poisson's equation. The distribution is given as

$$\Psi(x) = \frac{qN_A}{\epsilon_0 \epsilon_{si}} \cdot \left(\frac{x^2}{2} - wx \right) + \Psi_s. \tag{4.16}$$

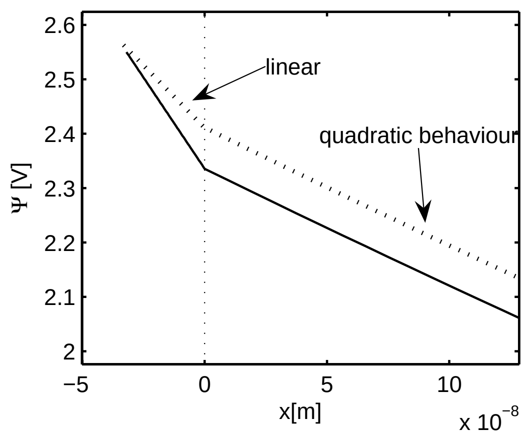
3. *Remaining semiconductor bulk region* ($w < x$): Since there is no space charge in the remaining bulk region, the potential corresponds to the externally applied bulk voltage, which is generally zero:

$$\Psi(x) = 0. \tag{4.17}$$

Figure 4.10 shows the potential distributions of a MOS-gate structure for a set of different effective gate voltages from 0.5 to 8.5V. Two curves have been plotted, where the solid lines correspond to the analytically calculated distributions according to the equations 4.15, 4.16 and 4.17 and the dashed lines correspond to the result of a numerical simulation. The simulation is a more realistic computation because both the process and the device simulations have been performed in a two-dimensional space.



(a) Comparison of theory with numerical simulation. The voltage distribution is shown up to a depth of $4\mu\mathrm{m}$ in the semiconductor bulk. The solid lines correspond to the analytical characteristics and the dashed lines are based on the numerical simulation.

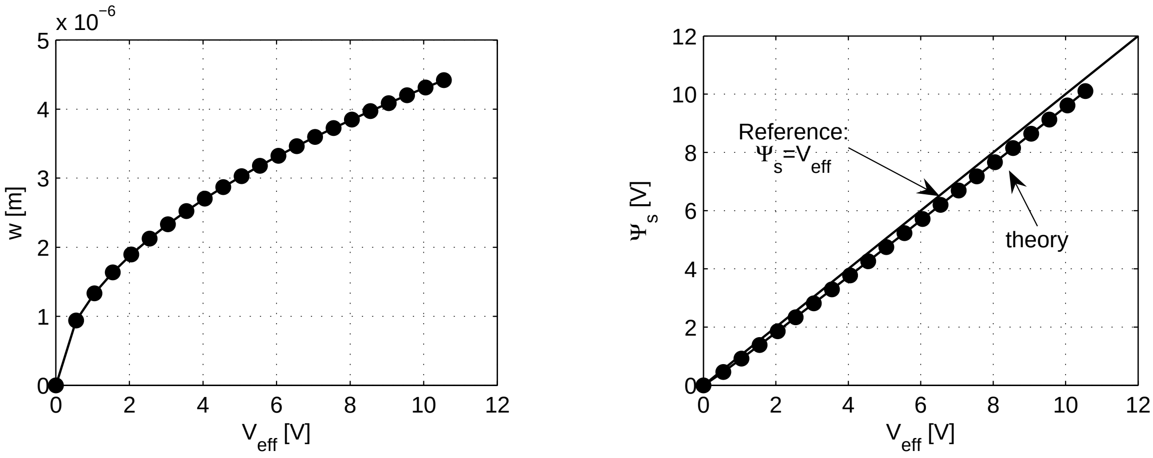


(b) This figure is a zoom of the voltage characteristics at $V_{eff} = 2.5\mathrm{V}$, depicting the linear increase in the oxide region ($x < 0$). In the silicon material ($x > 0$), the potential decreases quadratically down to $0\mathrm{V}$.

Figure 4.10: Voltage distribution through the vertical cross section of a surface channel CCD device.

It is shown that the mathematical model, for example the assumption that the depletion region behaves like a one-sided abrupt junction, is a good approximation to reality.
Following from the equations 4.13 and 4.14, the oxide thickness and substrate doping have a direct impact on the potential distribution and the depletion width. A higher doping concentration, for example, implies a shorter depletion region. The analytical computations presented in this work use an average substrate doping of about $10^{15}\mathrm{cm}^{-3}$ and an oxide thickness of $32\mathrm{nm}$, which are typical values for a combined CCD/CMOS process.
While figure 4.10 (a) shows a good agreement between the analytical and numerical approaches, figure 4.10 (b) is a zoom into the graph for $V_{eff} = 2.5\mathrm{V}$, making visible the linear decreasing behaviour in the oxide region. The dependencies of the depletion width and the surface potential on the effective gate voltage are shown in figures 4.11 (a) and (b) for the particular doping concentration and oxide thickness mentioned above. Obviously, the surface potential of the surface channel CCD can be linearly approximated by [106]

$$\Psi_s \approx V_{eff}. \tag{4.18}$$



(a) Depletion width versus effective gate voltage.

(b) Surface channel potential versus effective gate voltage.

Figure 4.11: Theoretical effect of the effective gate voltage $V_{eff} = V_g - V_{FB}$ on the depletion width and the surface potential.

Buried Channel

The potential distribution of the buried channel is characterized by the additional n^+-implant in the semiconductor bulk region. The doping and depth of the n^+-implant are chosen in such a way that a depletion region is created below the gate when a gate voltage of V_g greater than a threshold voltage is applied. This threshold voltage should be less than or at least equal to 0V, so that the semiconductor is also depleted when no voltage is applied to the gates.
Having a depletion region below the gate that ranges from $x = 0$ to $x = w$ (see figure 4.9), the device has to be split into four regions for solving Poisson's equation. In addition to the boundaries of the insulator and space charge region as given in the surface channel CCD, two parameters are defined: x_{pn} denotes the depth of the pn-junction and x_s means the position of the maximum potential in the buried channel. In the following, the solution of the Poisson equation is given, while the derivation and additional information about the buried channel device operation can be found for example in [103] and [110], respectively. The doping concentrations of the n^+-implant and the semiconductor bulk are denoted by N_D and N_A, respectively. An abrupt pn-junction is assumed, implying a uniform doping concentration N_A and N_D. Furthermore, the boundary conditions are $\Psi(-d_{ox}) = V_{eff}$ and $\Psi(w) = 0$.

1. *Oxide region* ($-d_{ox} < x < 0$): The potential increases linearly in the silicon dioxide between the gate and the bulk, and it is described by

$$\Phi(x) = V_{eff} - E_{ox} \cdot (x + d_{ox}), \tag{4.19}$$

   where the electric field E_{ox} is expressed as

$$E_{ox} = -\frac{qN_D x_{pn}}{\epsilon_{ox}\epsilon_0} - \frac{qN_A x_{pn}}{\epsilon_{ox}\epsilon_0} \cdot \left(1 + \frac{\epsilon_{si}\epsilon_0 d_{ox}}{\epsilon_{ox}\epsilon_0 x_{pn}}\right) + \frac{\epsilon_{si}\epsilon_0}{\epsilon_{ox}\epsilon_0}\sqrt{C_1}. \tag{4.20}$$

   The parameter C_1 is given after this itemization.

2. *n-side depletion region* ($0 < x < x_{pn}$): The potential increases quadratically according to

$$\Psi(x) = \Psi_{max} - \frac{qN_D}{2\epsilon_{si}\epsilon_0} \cdot (x - x_s)^2, \tag{4.21}$$

   where Ψ_{max} denotes the maximum potential in the depletion region and is expressed as

$$\Psi_{max} = \frac{qN_A}{2\epsilon_{si}\epsilon_0}\left(1 + \frac{N_A}{N_D}\right)\left[-\left(1 + \frac{\epsilon_{si}\epsilon_0 x_{ox}}{\epsilon_{ox}\epsilon_0 x_{pn}}\right)x_{pn} + \frac{\epsilon_{si}\epsilon_0}{qN_A}\sqrt{C_1}\right]^2. \tag{4.22}$$

   The maximum channel potential is located at

$$x_s = x_{pn} + \frac{N_A x_{pn}}{N_D} \cdot \left(1 + \frac{\epsilon_0\epsilon_{si} d_{ox}}{\epsilon_0\epsilon_{ox} x_{pn}}\right) - \frac{\epsilon_0\epsilon_{si}}{qN_D} \cdot \sqrt{C_1}. \tag{4.23}$$

3. *p-side depletion region* ($x_{pn} < x < w$): The potential decreases quadratically down to 0V.

$$\Psi(x) = \frac{qN_A}{2\epsilon_{si}\epsilon_0}(w - x)^2, \tag{4.24}$$

   where the depletion width is

$$w = -\frac{\epsilon_0\epsilon_{si} d_{ox}}{\epsilon_0\epsilon_{ox}} + \frac{\epsilon_0\epsilon_{si}}{qN_A} \cdot \sqrt{C_1}. \tag{4.25}$$

4. *Remaining bulk region* ($x > w$): According to the boundary condition, the potential in the remaining bulk region is zero.

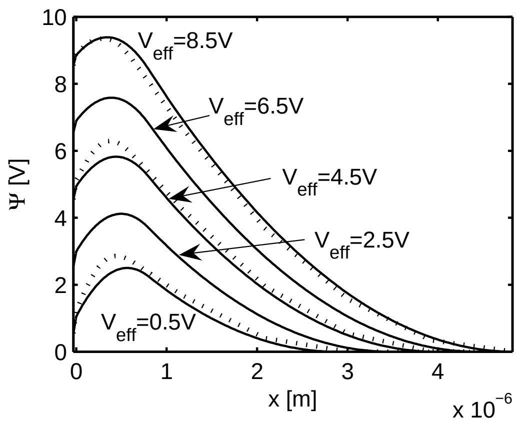
$$\Psi(x) = 0 \tag{4.26}$$

The parameter C_1 is given by

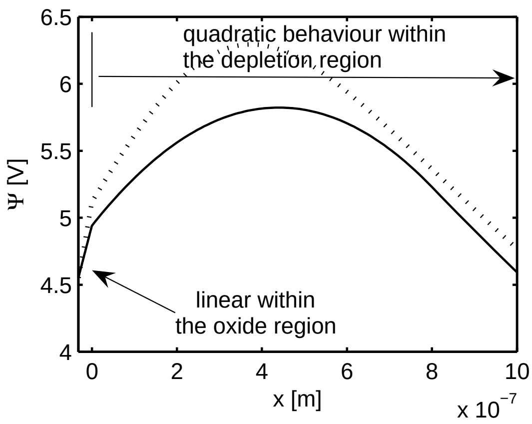
$$C_1(V_{eff}) = \left(\frac{qN_A x_{pn}}{\epsilon_{si}\epsilon_0}\right)^2 \left(1 + \frac{\epsilon_{si}\epsilon_0 x_{ox}}{\epsilon_{ox}\epsilon_0 x_{pn}}\right)^2 + 2\frac{qN_A}{\epsilon_{si}\epsilon_0} \cdot \left(V_{eff} + \frac{qN_D x_{pn}^2}{2\epsilon_{si}\epsilon_0}\left(1 + 2\frac{\epsilon_{si}\epsilon_0 x_{ox}}{\epsilon_{ox}\epsilon_0 x_{pn}}\right)\right). \tag{4.27}$$

Figure 4.12 shows the potential distribution in the buried channel CCD for several effective gate voltages from 0.5 to $8.5\mathrm{V}$. The comparison of the analytically calculated graphs with the numerically simulated data, which has been extracted from a two-dimensional process and device simulation, shows again that the basic physical model (abrupt pn-junction, uniform doping concentrations) describes the device quite well for practical purposes.

The dependencies of the depletion region and the maximum channel potential on the effective gate voltage are illustrated in figure 4.13, using the device parameters of the realized TOF sensors. It is obvious that the depletion width is bigger than that of surface channel CCDs.
In general, the maximum potential is greater than the gate voltage, leading to the con-

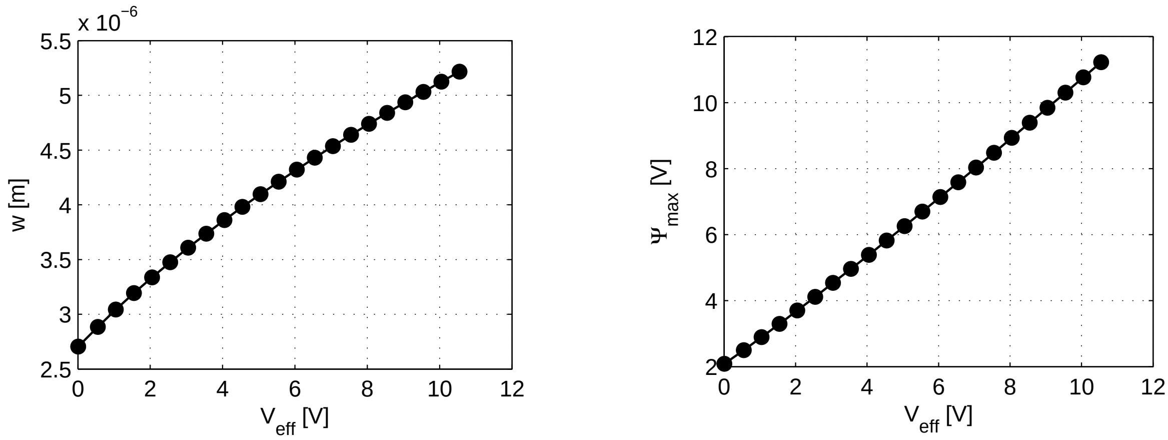


(a) Comparison of theory with numerical simulation. The voltage distribution is shown up to a depth of $5\mu\mathrm{m}$ within the semiconductor bulk. The solid lines correspond to the analytical characteristics, and the dashed lines are based on the numerical simulation.



(b) This figure is a zoom of the voltage characteristics at $V_{eff} = 4.5\mathrm{V}$, depicting the linear increase in the oxide region ($x < 0$).

Figure 4.12: Voltage distribution through the vertical cross section of a buried channel CCD device.



(a) Depletion width versus effective gate voltage.

(b) Maximum channel potential versus effective gate voltage.

Figure 4.13: Theoretical effect of the effective gate voltage $V_{eff} = V_g - V_{FB}$ on the depletion width and the channel potential.

dition

$$\Psi_{max} > V_{eff}. \tag{4.28}$$

Approximately, the maximum channel potential is directly proportional to the effective gate voltage plus an additional constant voltage offset V_0, so that

$$\Psi_{max} \approx V_{eff} + V_0. \tag{4.29}$$

This approximation is an important relationship because it leads to the conclusion that the electric fringing fields in the channel are approximately the same as between the gates due to the elimination of the constant offset voltage in the first derivative $\partial\Psi/\partial x$. Since the maximum potential is located inside the depletion region and photo-generated electrons move to the point of highest potential, the charge storage and transport is performed in the bulk of the semiconductor instead of at its surface. The buried channel is located a few hundred nanometers inside the semiconductor (see figure 4.12). Surface effects such as short lifetimes and significant trapping noise are mitigated.

4.3.2 Charge Transport

The appropriate gate biasing of the CCD gate arrangement enables the lateral transport of photo-generated and collected electrons. Figure 4.14 illustrates the basic idea of charge transport in CCDs, where no differentiation between surface and buried channel has been made. In the case of a surface channel, the surface potential Ψ_s is depicted below the gate structure. In this figure, the positive potential is plotted downwards. If a buried channel CCD is used, the depicted potential corresponds to the maximum potential Ψ_{max}. The potential distribution can be considered as potential wells, which are able to store a specific amount of minority carriers.

In figure 4.14 (a) the situation is depicted when charges have been collected below gate 1. While this gate is set to high potential, the low potentials on the neighbouring gates avoid charge leakage to these gates. By applying subsequently appropriate low and high voltages to the gates, as shown in figures (b) and (c), the charge moves to the right, always following the direction of highest channel potential, which corresponds to minimum energy for electrons.

Charge Transport Methods: The speed of the lateral charge transport in the y-direction is limited by the dominating transport mechanism. We distinguish between three transport phenomena of free charge carriers, whose principles are roughly sketched in figure 4.15: [28, 60, 105]

1. *Self induced drift*: Electrons experience the Coulomb force due to the electric field induced by the charges themselves. The electric field is denoted by E. It is directly proportional to the density gradient in the y-direction [28]

$$E \propto q \cdot \frac{\partial n}{\partial y}, \tag{4.30}$$

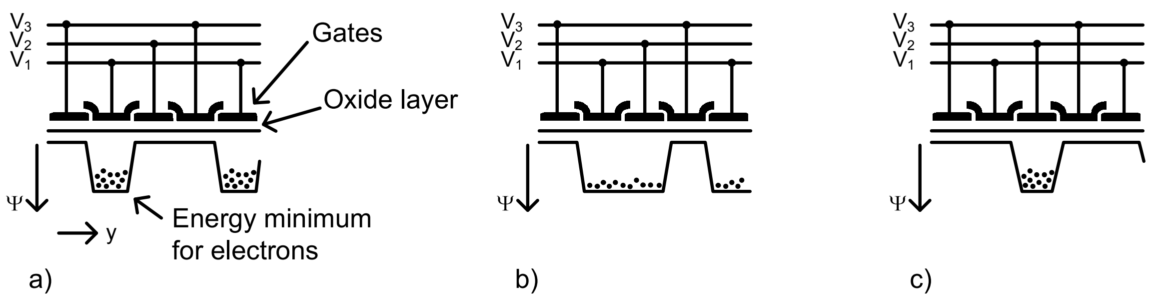
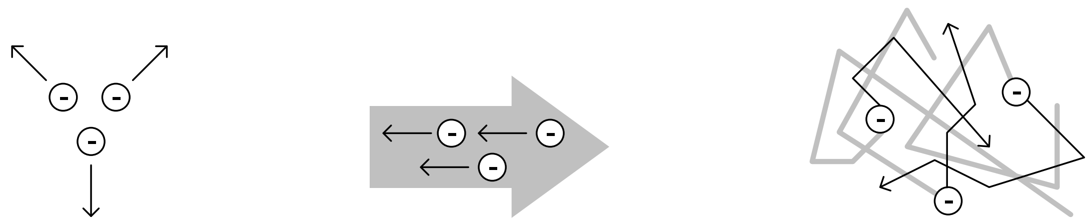


Figure 4.14: Lateral charge transport in CCDs.



(a) Self induced drift (Coulomb repulsion).

(b) Electron movement by drift fields.

(c) Thermal diffusion.

Figure 4.15: Transport of free charge carriers.

where n is the density of free electrons. The proportionality factor depends on the effective capacitance where the charge has been collected.

2. *Drift by external electric fields*: Externally applied fields such as the fringing fields in CCDs, lead to a significant drift component as well. Based on the approximations of the equations 4.18 and 4.29 for the surface and buried channel CCD respectively, the electric fields in the channel correspond to

$$E = \frac{\Delta V_g}{\Delta y}, \tag{4.31}$$

where Δ denotes the difference of the particular parameters.

3. *Thermal diffusion*: In any case free electrons will be moving around also due to thermal agitation. Accordingly, thermal diffusion processes are strongly temperature dependent.

Because the TOF pixels deal with the transportation of small numbers of electrons [60], the gradient $\partial n/\partial x$ and hence the self-induced field component can be neglected for the pixel's speed estimation. [28]

The transportation time of the photo-generated charge carriers is rather dominated by thermal diffusion and drift processes due to externally applied electric fields. As seen in section 4.1.2, the average transport time over a path length s by thermal diffusion is given by

$$t_{diff} = \frac{s^2}{D}. \tag{4.32}$$

If electric fields are present, the transport time (without taking diffusion into account) is

$$t_{drift} = \frac{s}{v}, \tag{4.33}$$

where v is the drift velocity of the charge carriers. The drift velocity depends on several parameters, namely the silicon crystal orientation, the electric field, the doping concentration and the temperature. The drift velocity is illustrated for different electric fields and doping concentrations in figure 4.16, where a room temperature of 300K and a crystal orientation $< 111 >$ have been assumed. The underlying model of these graphs is the model of Scharfetter and Gummel [51]. It is noted that the drift velocity for doping concentrations $N < 10^{16}/\text{cm}^3$ reaches a maximum value. This is a reason why the buried channel should be weakly n-doped.
The drift velocity for weakly doped silicon material can be expressed by the phenomenological expression

$$v(E(s)) = \frac{\mu_0 E(s)}{1 + \frac{\mu_0 E(s)}{v_{th}}}, \tag{4.34}$$

where the mobility μ_0 is defined as the ratio between the maximum drift velocity v_{th}, which is the thermal velocity of charge carriers, and the electric field E_0 at which 3dB of the maximum drift velocity is reached. For silicon material, μ_0 is about $1400\text{cm}^2/(\text{Vs})$, E_0 is about $7 \cdot 10^3\text{V/cm}$, and $v_{th} \approx 10^7\text{m/s}$ at 300K.

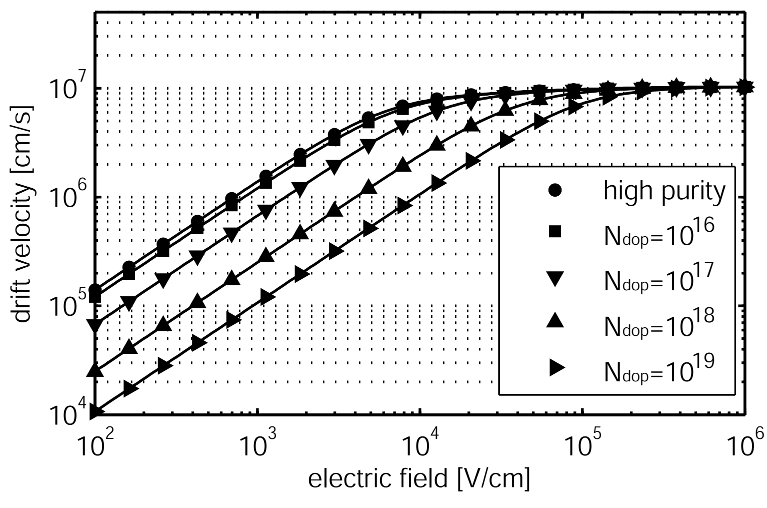


Figure 4.16: Drift velocity of electrons in silicon material with different doping concentrations N_{dop}. Room temperature is assumed and the silicon has a crystal orientation of $< 111 >$.

Charge transport by electric fields can be accomplished much faster than the transport based on thermal diffusion processes. This is shown by a comparison of the transport time that is needed for the transportation of one electron over the path s. Assuming that no high-field effects affect the drift velocity, equation 4.34 can be linearized as

$$v(E(s)) = \mu_0 E(s) \text{ with } E(s) < E_0. \tag{4.35}$$

Using the relation $v = s/t_{drift}$ and assuming further a constant electric field generated by a voltage drop of ΔV over the path length s, the drift time results in $t_{drift} = s^2/(\mu_0 \Delta V)$. The ratio between the diffusion time and the drift time describes the gain when using electric fields for the charge transport, and it is given as

$$G_{tr} = \frac{t_{diff}}{t_{drift}} = \frac{\mu_0}{D} \cdot \Delta V \approx 38 \cdot \Delta V, \text{ with } E = \frac{\Delta V}{s} < E_0. \tag{4.36}$$

Here, the Einstein relation has been applied relating the charge's mobility and diffusivity:

$$\begin{aligned} D &= \tfrac{k\vartheta}{q}\mu \\ &\approx \tfrac{1}{38} \cdot \mu \text{ for } \vartheta = 300\text{K}. \end{aligned} \tag{4.37}$$

Depending on the applied voltage drop and the electric field over the transport path s, the transport time can be reduced by several orders of magnitude by applying electric fields instead of exploiting thermal-diffusion-based charge transport.
The charge transport in surface channel CCDs is dominated by diffusion processes, while the fringing fields in buried channel CCDs have a more significant impact on the transport speed. This is the case because the lateral potential distribution is smoothed out deeper in the silicon bulk, as is illustrated in figure 4.17.

While the standard TOF pixels are mainly limited in speed by relatively slow diffusion processes, the enhanced pixel structures, developed for the demodulation of very high frequencies, are based on lateral electric drift fields in order to reduce pixel-inherent signal delays. However, the physical principles of charge transport mechanisms as described above, are essential for the design and analysis of these pixel structures.

Transport efficiency: Due to trapping states in the silicon material, the lateral transport of charges is not perfectly accomplished, which means a detectable loss of charges occurs during the transport. Following this, the transport efficiency is defined as the ratio between the amount of charge carriers that are stored after and before the lateral shift by one CCD gate.

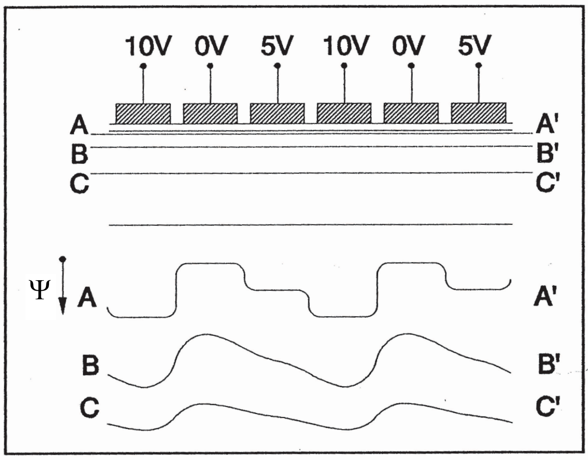


Figure 4.17: Illustration of the fringing fields at the Si-SiO2 interface (A-A′), at a shallow depth (B-B′), and deeper (C-C′) in the silicon bulk. [106]

In common CCD sensors, the efficiency parameter is an important figure of merit because even high efficiency values lead to significant loss of information when shifts over several thousands of gates have to be performed. For example, a transport efficiency of 0.99999 leads to a loss of up to 8% of the original amount of charge when the charge packet is transported through 4×2048 CCD gates of a 1024×1024 frame transfer CCD sensor that is based on a four phase CCD transport system [40,106]. Therefore, recent standard CCD processes are optimized to have very high transport efficiencies. In particular, buried channels help to improve the transport efficiency because the density of trapping states is lower inside the semiconductor bulk area.

With regard to the TOF pixels, the transport efficiency is not a critical parameter because the charge needs to be transported over only a few gates within the pixel. No additional shift over the whole sensor area is required for read out because the collected charge carriers are already detected within the pixel using standard source follower stages.

4.3.3 Charge Detection

To detect the photo-generated charge means its conversion into a voltage that can subsequently be processed by analogue or digital electronics. In the TOF pixels, the charge detection already takes place in the pixel as is typical for active pixel sensors (see section 4.2). The detection is done using a standard source follower according to the illustration in figure 4.18. By transferring the charge packet onto the sense diffusion node, a variation ΔV of the voltage at the output node depending on the signal charge Q is given by [103]

$$\Delta V = \frac{Q}{C_{dep} + C_{gd}}, \tag{4.38}$$

where C_{dep} is the depletion capacitance of the sense diffusion node and C_{gd} is the capacitance between the gate and drain of the MOS transistor.
The design of the output stage capacitances requires a trade off: On the one hand, a small capacitance leads to a high conversion gain and low contribution of kTC-noise. On the other hand, small capacitances restrict the collection of charges to small amounts, which would in turn inhibit high-precision distance measurements.

4.4 Quantum Efficiency of the MOS-Gate Structure

Following the theoretical derivations of the achievable stochastic accuracies for PN or sinusoidally modulated signals, the standard deviation of the distance measurement is determined by the total number of electrons generated by the light signal. The number of electrons N_{el}, in turn, is related to the impinging number of photons N_{ph} by the quantum efficiency

$$QE = \frac{N_{el}}{N_{ph}}. \tag{4.39}$$

A higher quantum efficiency allows for increasing the measurement accuracy or, when keeping the accuracy constant, for reducing the integration time and hence, for increasing the frame rate of the sensor. Thus, the quantum efficiency is an important system parameter, which for a given wavelength strongly depends on the particular fabrica-

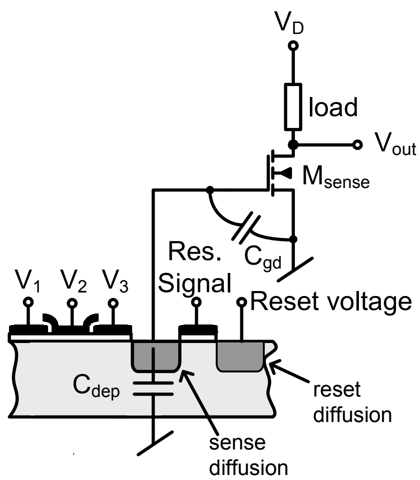


Figure 4.18: Charge detection.

tion steps and the pixel design parameters. While the process influences are difficult to estimate, the impact of the pixel design with regard to its realisation with several layers of different materials can be more efficiently predicted by theory. The basic theoretical considerations and the measurement of a CCD gate structure are discussed in the following subsections.

4.4.1 Theoretical Analysis of the QE

Due to several effects, the number of detected photons depends on the particular wavelength considered. Figure 4.19 shows what can happen when an incident photon propagates through some layers of different materials towards the depletion region in the bulk area of the semiconductor device [95]. Only photons generating charge carriers that reach the depletion region, will be detected by the pixel.

Six different effects, adopted in principle from [95], prevent photons from detection:

1. Reflections of the incident photons at the top surface due to differences between the refractive indices of the air and the top layer material, which is generally silicon nitride (Si_3N_4).

2. Multiple reflections in the covering thin layers, again due to differences between the refractive indices of the layers.

3. Absorption of photons already in the covering layers, since the covering layers are not perfectly transparent.

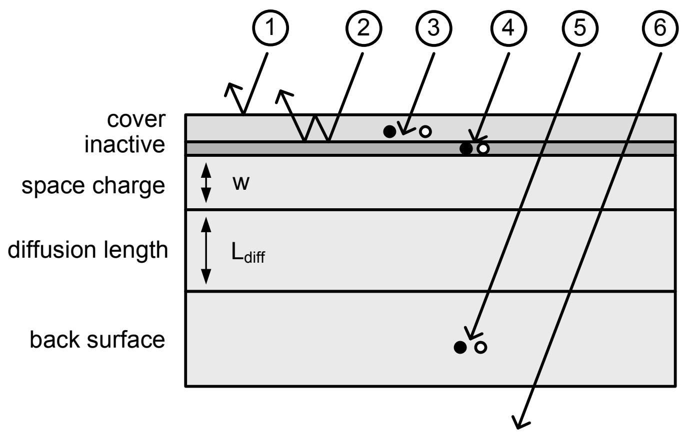


Figure 4.19: Propagation of light through a pixel structure composed of stacked layers of different materials. [95]

4. A very short lifetime of photo-generated charge carriers in the inactive regions near the surface of the semiconductor prevent the separation and collection of these carriers. Such inactive regions are very thin (less than 1nm) because they are mainly the result of material defects and dangling bonds at the interface, or they are the result of high doping concentrations near contacts [95,105].

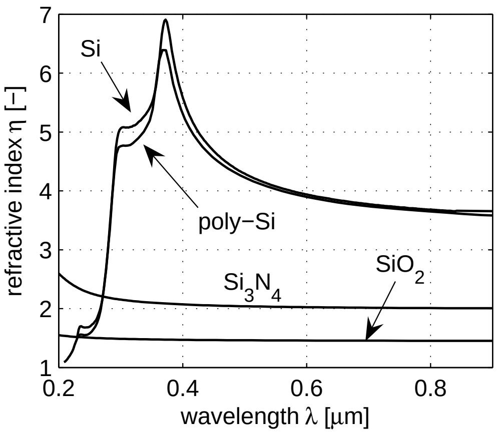
5. Any charge that is generated deeply in the substrate might not have a chance to reach the depletion width by thermal diffusion.

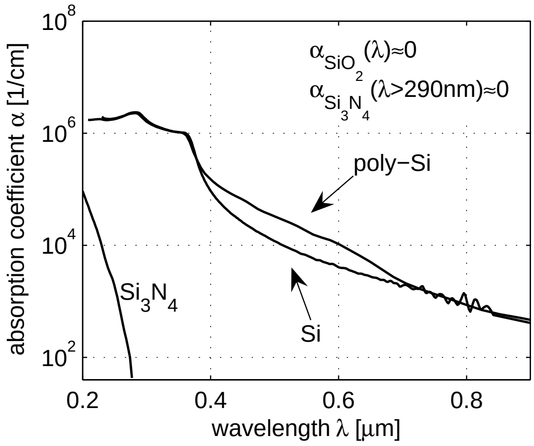
6. A fraction of the impinging photons might travel through the image sensor.

All the effects mentioned are based on two physical phenomena: Firstly, the treatment of the photons as an electromagnetic wave describes the reflections or transmissions at the interfaces between two medias of different refractive indices. Secondly, the transmission of the photon energy to the material's valence electrons according to the internal photo-effect describes the absorption of the optical energy when propagating through a medium.
The material-specific refractive indices η as well as the absorption coefficients α depend strongly on the optical wavelength and material properties such as the crystal's purity, its orientation, the fabrication parameters and the temperature [86]. Some literature values of material coefficients at 300K are shown in figure 4.20. The selected materials are those used in the CCD gate structure of the ZMD process [7]: silicon (Si), silicon dioxide (SiO_2), silicon nitride (Si_3N_4) and poly-silicon. The data for silicon are taken from [63] and the data for the other materials have been extracted from the Sopra database [2]. The fundamentals of optical reflection, transmission and absorption can be found for example in [11,86,105,113,126].

Theoretical description of the multilayer system of a CCD gate structure: With regard to quantum efficiency, the number of photons reaching the substrate needs to be calculated. Figure 4.21 depicts the multilayer system with the materials from top to bottom: silicon nitride (Si_3N_4) as a passivation layer, silicon dioxide (SiO_2), poly-silicon gate and the silicon substrate. The infinite multiple reflections of the impinging light wave between the layers lead to interference and absorptions, both of which can theoretically be described by a matrix theory that is presented in detail for example in [11]. The transmission coefficient of the multilayer system is expressed as

$$\mathrm{t} = \frac{2\eta_{air}}{\eta_{air} E_{air} + H_{air}}, \tag{4.40}$$





(a) Refractive index versus optical wavelength.

(b) Absorption coefficient versus optical wavelength. The coefficient of SiO_2 is missing in this graph because it is almost equal to zero.

Figure 4.20: Refraction and absorption parameters of silicon (Si), silicon dioxide (SiO_2), silicon nitride (Si_3N_4) and poly-silicon.

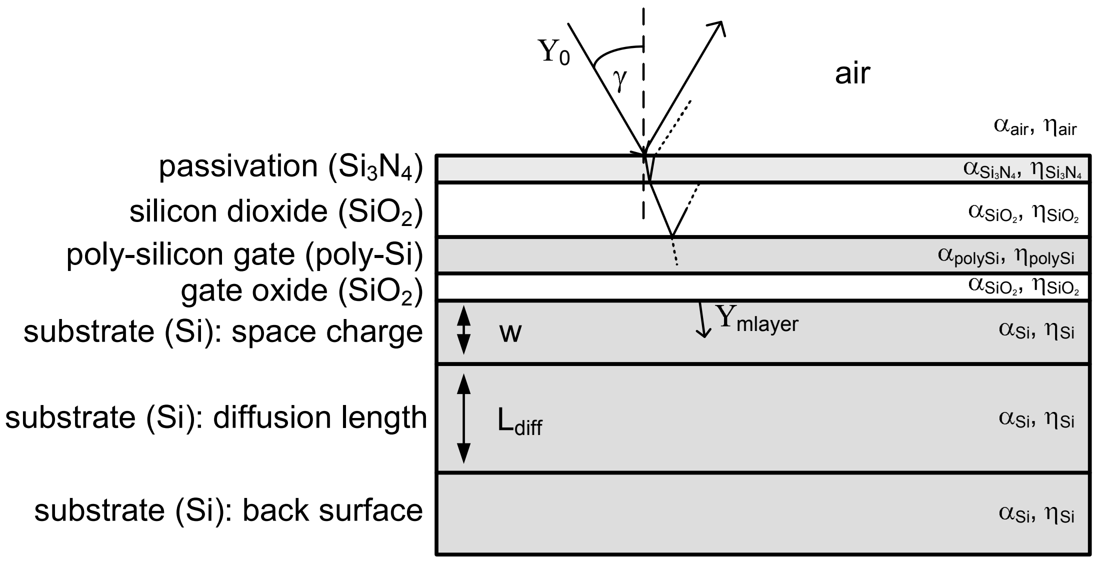


Figure 4.21: Layer composition of a pixel with a CCD gate structure.

where E_{air} and H_{air} are the electric and magnetic vectors in air, respectively. These vectors are computed by a matrix multiplication

$$\begin{pmatrix} E_{air} \\ H_{air} \end{pmatrix} = \mathfrak{M} \cdot \begin{pmatrix} 1 \\ \eta_{sub} \end{pmatrix}, \tag{4.41}$$

where $\mathfrak{M}$ describes the complete multilayer system, including the refraction and absorption of the nitride, dioxide and gate layers (see [11]). The photon flux reaching the substrate results in

$$\Upsilon_{mlayer}(\lambda) = \frac{\eta_{sub}}{\eta_{air}} |\mathfrak{t}|^2 \cdot \Upsilon_0, \tag{4.42}$$

where η_{sub} and η_{air} are the refractive indices of air and the substrate, respectively. The refractive index of the substrate is given by the refractive index of silicon material. However, because the transmission coefficient $\mathfrak{t}$, which has been defined in equation 4.40, refers to the propagation of the light wave through the non-substrate surface layers, this model has to be enhanced by the absorption in the substrate. Finally, the detected photon flux Υ_{det} can be expressed as [65, 103]

$$\Upsilon_{det}(\lambda) = \Upsilon_{mlayer}(\lambda) \cdot \left(1 - \frac{e^{-\alpha w}}{1 + L_{diff}\alpha_{Si}(\lambda)}\right), \tag{4.43}$$

and the quantum efficiency is the ratio between the detected photons and the incident photons

$$QE(\lambda) = \frac{\Upsilon_{mlayer}(\lambda)}{\Upsilon_0} \cdot \left(1 - \frac{e^{-\alpha w}}{1 + L_{diff}\alpha_{Si}(\lambda)}\right). \tag{4.44}$$

The multiplication factor includes the absorption in the depletion region of width w according to the Lambert law [105, 113, 126] and a second part that takes into account the electrons that have been generated below the depletion region but which are still collected due to their diffusion into the depletion region. The detailed derivation of this diffusion-based model can be found in [103].

4.4.2 QE measurement

The measured and theoretically computed quantum efficiency of a poly-silicon gate structure with a buried channel of doping concentration such as used in the TOF pixels is shown in figure 4.22. The layer design of the test structure corresponds to the schematic in figure 4.21.

As the 3D imaging system normally operates with non-visible light in the near-infrared region, the quantum efficiency should be optimized for long wavelengths around

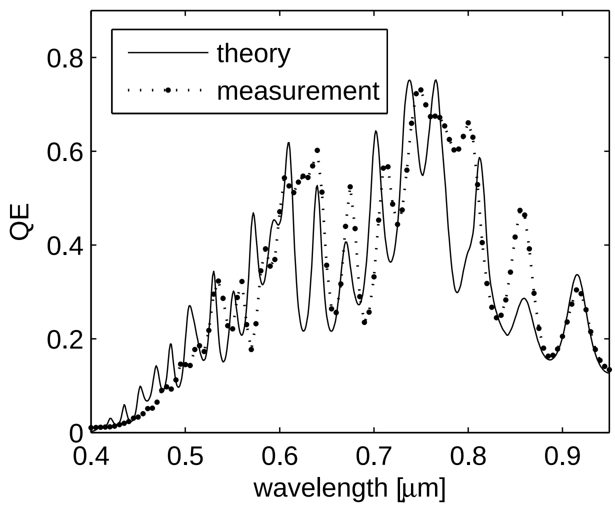


Figure 4.22: Quantum efficiency - measurement and theory.

850nm. Such an optimization includes the adaptation of the processing steps in order to reach suitable refractive indices, and the adaptation of the layer thicknesses. It is obvious that small variations of the quantum efficiency curve can result in large variations of the performance of the distance measurement system. The measured quantum efficiency around the optical wavelength of $\lambda = 850\text{nm}$, which is interesting for the TOF sensors, is larger than 40%.

The theoretical curve in figure 4.22 is based on the theory explained one subsection above. A fit of the layer thicknesses was performed within the confidence limits provided by the foundry of ZMD [7]. This fit results in the following thicknesses: silicon nitride 750nm, silicon dioxide $3.105\mu\text{m}$, poly-silicon of the gate 302nm and the gate oxide 33nm. The diffusion length that was used in equation 4.44 was found to be about $25\mu m$, which corresponds to a charge carrier lifetime of $\tau = L^2/D \approx 0.2\mu s$ ($D \approx 37\text{cm}^2/s$). The depletion width was set to $3\mu m$. As refraction and absorption coefficients, the data of figure 4.20 were used.

The basic characteristics could be well reestablished by the theoretical computations. However, some remarkable distortions occur nevertheless at some particular wavelengths. The following list of potential error sources clarifies the complexity of the physical model:

- The refractive indices and absorption coefficients are taken from literature. However, they can vary depending on the purity of the crystalline silicon material. Different indices would result in different peak frequencies at small wavelengths.

- The bandwidth of the light source is not infinitely small, which has a low pass effect on the quantum efficiency curve.
- Though the measurement apparatus is calibrated, the wavelengths have a small uncertainty.
- The simulation assumes an incident angle of 0 degree, which is an idealisation of the system. However, a wider angle would have a low-pass filtering effect on the simulated characteristics of the QE.
- The measurement of the layer thicknesses has an uncertainty.
- The lifetime of the charge carriers and the depletion width are approximations. They have a direct impact on the behaviour at large wavelengths.

It is noted that the significant decreases of the quantum efficiency for short and long wavelengths are the result of absorption, particularly in the poly-silicon gate, and the almost lossless transmission through the depletion region, respectively. Absorption by silicon nitride would be even more significant at even shorter wavelengths $\lambda < 290$nm, so that the top-layers of nitride and silicon dioxide do not have a significant impact with regard to the absorption of photons in the depicted range of wavelengths.

Chapter 5

Demodulation Pixel in CCD/CMOS-Technology

Three-dimensional time-of-flight imaging sensors exploit the possibility of modern semiconductor processes to integrate a high density of smart pixels on one chip. Many approaches exist for the realisation of the necessary in-pixel signal processing for measuring the optical signal's delay. Some of the state-of-the-art pixel architectures are reported in [9, 23, 26, 37, 71–74, 82, 94, 108, 111], and the interested reader is advised to make use of those references to obtain a widespread overview of the available pixel technologies.

In this chapter, first a brief overview of the fundamental lock-in pixel structure is given before we present the architecture of the CCD/CMOS demodulation pixels developed in the framework of the present thesis. We introduce enhanced pixel structures for enabling the application of the pixels in scenes of high-dynamic range or with high background illumination. Another emphasis lies on new pixel concepts for the demodulation of very high frequencies up to several hundreds of Megahertz. Practical verifications of the concepts conclude this chapter.

5.1 CCD/CMOS Lock-In Pixel

The CCD/CMOS lock-in pixels are designed according to the abstract pixel model introduced in chapter 2 concerning the time-of-flight distance measurement. This means that each pixel is composed of a photo-sensitive area and a region for directing the photo-generated electrons into a bucket that accumulates the electrons during a specific period of integration. This integration process is repeated to increase the signal to noise ratio. At the end of the whole integration process a pixel-specific number of charge packets is available for read-out. Figure 5.1 illustrates the four essential physical functions of a demodulation pixel: photo detection, charge separation, repeated charge addition and the storage of a specific number of charge packets.

The first subsection briefly reviews the history of CCD/CMOS lock-in pixels before the second subsection sketches the state-of-the-art pixels currently being incorporated into the solid-state 3D-camera. The drawbacks of these structures are pointed out, explaining the motivation for the development of the enhanced demodulation pixels.

5.1.1 Review

The first demodulation pixels were reported by Spirig [103, 104]. Various implementations of demodulation pixels were described, mainly differing in the number of samples of the modulation signal that are acquired within one integration period at a time. The charge packets stored in the pixel are also called taps. The concept of a multitap lock-in CCD [102] was already successfully proven for application to optical phase measurements in the early 90s. [103]
The architecture of such a multitap pixel is shown in figure 5.2. The upper part of the pixel consists of a photo-gate, a dump gate and dump diffusion and a 4-phase BCCD pipeline register. The photo-gate forms the photo-sensitive region because it is not covered by any metal. The dump gate and dump diffusion are designed for electronic shutter purposes, for example photo-generated charges can be dumped during the read-out process. The bottom part of the pixel is composed of a 4-phase BCCD storage and read-out register. Additionally, a transfer gate links the upper and bottom parts.
Depending on the number of storage gates (pixels with up to 16 taps had been realised), this particular number of samples is first stored in the upper pipeline register before the charge packets are subsequently transferred in parallel mode to the bottom storage gates by employing the transfer gate. This procedure is repeated for many modulation

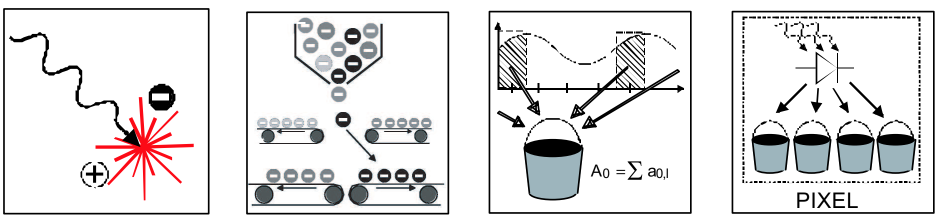


Figure 5.1: Physical functions of a demodulation pixel. Extracted and modified from [60].

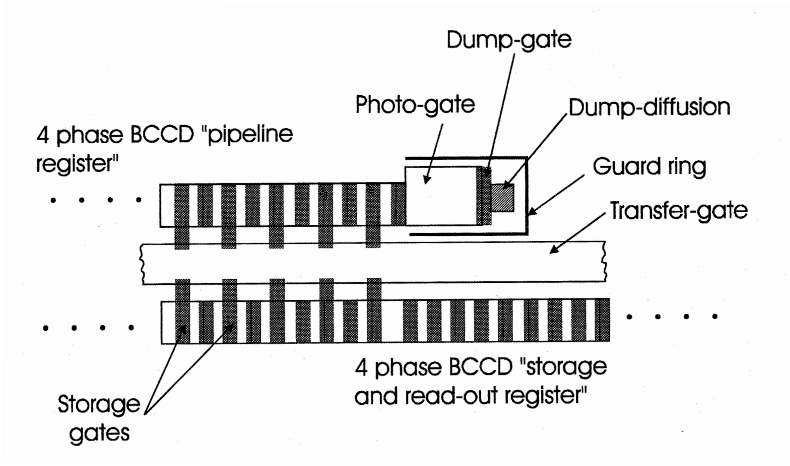
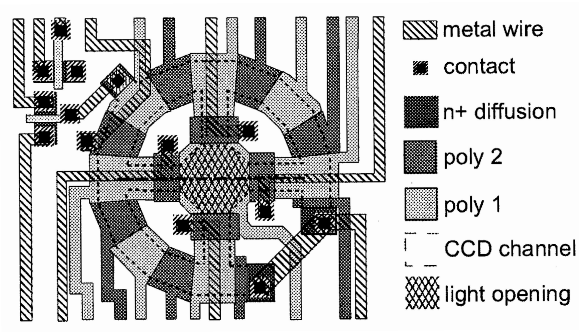


Figure 5.2: Schematic layout of the multitap lock-in CCD. Five taps are shown. [103]

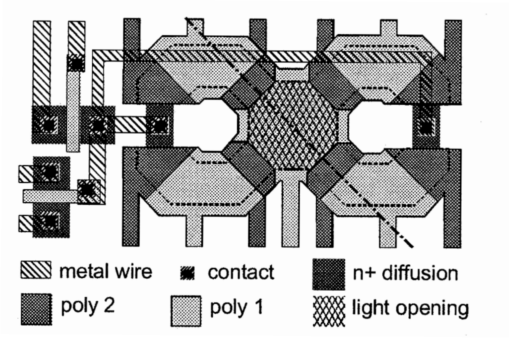
periods in order to increase the signal to noise ratio, before the charge packets are read out.
The main drawbacks of this multitap pixel approach are a low fill-factor (usually $\leq 5\%$ in a $2\mu m$-process) and the slowness in the demodulation and read-out processes because the design-rule-limited large CCD structures mainly employ thermal diffusion processes for the charge transport into the storage gates. This limits the pixel to a maximum modulation frequency of about 1MHz [60], which is not suitable for high-precision range imaging.

Since four samples are already more than necessary for the distance measurement with sinusoidal modulation, Lange [60] compared different pixel architectures having four or less taps and an in-pixel amplifier. Basically, if signals with less than four taps are sampled, the phase measurement is still based on the four-sample approach but a serial acquisition of the the four samples is performed.

In [60] two major designs of 4-tap structures are presented, which are shown in figure 5.3. The first design is a wheel shaped CCD structure, where the four storage gates, also referred to as integration gates, are located around the photo-gate in the middle of the pixel. The read-out of the four samples is subsequently performed by one amplification stage using the n^+-diffusion node. The second 4-tap structure employs two separated output diffusion nodes for the read-out. However, the integration gates are still located around the photo-gate. This design is more symmetrical and hence, it reduces inhomogeneities between the taps. The pixels were realised in a 2-μm-process.
Lange [60] points out that the major drawbacks are low fill-factors due to the fact that each sample needs some significant storage space (fill factors of 2.2% and 6.6% had been achieved for the wheel shaped four-tap pixel and the two-side readout version, respectively). Furthermore, the pixels suffer from the inhomogeneous response of the



(a) One diffusion node.

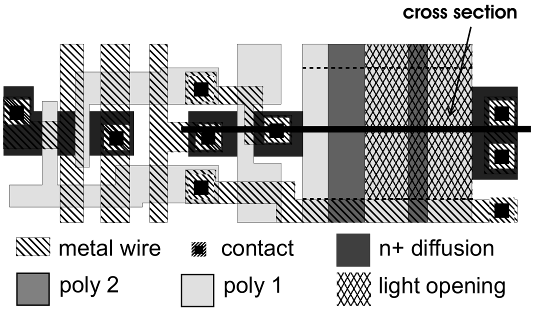


(b) Two diffusion nodes (on the left and right sides).

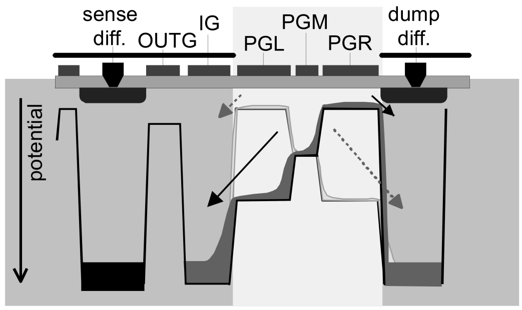
Figure 5.3: Four-tap lock-in pixel. [60]

four sampling points, either due to unsymmetrical design or due to process mismatch between the transfer gates.

As a consequence, the first compact solid-state 3D TOF camera [59] was set up with neither of the above mentioned multitap pixel architectures. The sensor of the first 3D TOF camera was rather based on the 1-tap lock-in pixel.
By offering just one storage site, this kind of demodulation pixel maximizes the fill-factor (more than 22.3% had been achieved), and inhomogeneities between the taps are avoided because all samples use the same physical storage and amplification stages. Figure 5.4 (a) sketches a design example of the 1-tap pixel in a $2\mu\mathrm{m}$-process. The cross-sectional view in (b), including the schematised potential distribution in the semiconductor bulk for different settings of the gate voltages, depicts the operation principle of the pixel. During the sampling of the impinging optical signal, all photo-generated charges are conducted to the integration gate IG. Between two sampling intervals, all electrons are dumped by the dump diffusion node. The operation requires appropriate voltage settings at the middle, left and right photo-gates (PGM, PGL and PGR) according to figure 5.4 (b). In order to read out the sampled charge packets, the charges are transferred from the integration gate via the out gate OUTG to the diffusion node, which is connected to the source follower stage.
The 1-tap approach has two major disadvantages: 1. A certain amount of the photo-generated charge is wasted instead of being exploited for the distance measurement. 2. The serial acquisition of all four samples negatively affects the frame rate of the camera system and makes the 3D images very susceptible to fast movements of the objects. However, these drawbacks were accepted in the design of the first 3D camera with the argument that the effects of lower frame rate, stronger susceptibility to fast object



(a) Pixel layout. Design published in [60].



(b) Cross-sectional view and illustration for the potential distribution in the semiconductor. [60]

Figure 5.4: One-tap pixel architecture.

movements and the waste of some part of the optical information is more or less compensated by the significant increase of the fill factor compared to multitap pixel architectures and by using sampling intervals equal to half the modulation period. Higher fill factor means shorter integration times and thus less susceptibility to object movements. A longer sampling interval results in wasting fewer photo-generated electrons and living with a measured signal amplitude that is reduced by about 36%. In particular, the high uniform tap response was another decisive factor for the implementation and incorporation of this pixel architecture into the 3D camera.

5.1.2 State-of-the-Art

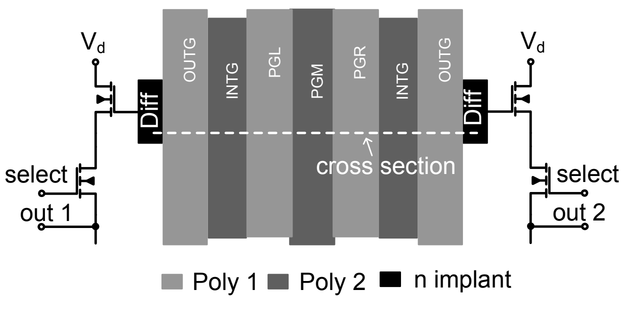
Today, the palette of efficiently usable demodulation pixels is not restricted exclusively to 1-tap structures. Improved technologies allow the reliable design of 2 or 4-tap pixel structures. It is common to all current pixels that they implement several diffusion nodes and amplifiers according to the number of taps, so that the read-out speed can be significantly increased by parallel read-out.
Mismatch problems are strongly reduced by the increased processing homogeneity over time and space. Remaining offset variations have to be eliminated by the accomplishment of a calibration measurement. The fill-factor benefits from even smaller process feature sizes.

The latest generation of 3D TOF cameras that are based on CCD/CMOS lock-in pixels (see for example [24, 55, 75, 77, 78]), apply a 2-tap pixel architecture, such as sketched in figure 5.5. It has a completely symmetric architecture, consisting of the photo-gates

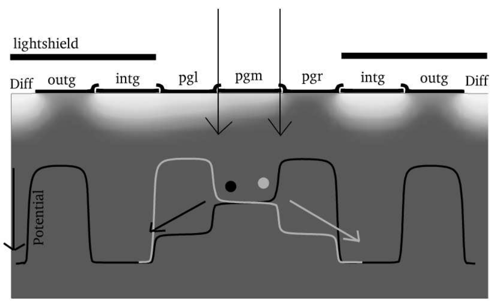
PGL, PGR and PGM in the middle, and on both sides an integration gate INTG for the charge accumulation, an out gate OUTG for decoupling purposes and a diffusion node, which is connected with a source follower circuit. Fill-factors of more than 20% have been obtained with pixel sizes down to $40 \times 40\mu m^2$. For many applications this architecture provides the best compromise between fill-factor, inhomogeneity, acquisition speed and distance measurement accuracy (no charge is wasted during the integration process). Instead of dumping the charge carriers after one sampling interval, these carriers are stored below the second integration gate so that they can also be considered as another sample of the signal.
The potential distribution shown in figure 5.5 (b) for two scenarios of gate voltage settings is the result of the numerical simulation of the device. It is the potential characteristics of the maximum potential Ψ_{max} in the buried channel, which is located at a depth of about $0.8\mu m$ below the semiconductor surface. Here, two points are worth mentioning:

1. Pixel structures with several taps can also be operated in a one-tap mode by reading out the charge packet of just one integration gate. Of course, the timing has to be adapted to an acquisition of four consecutive samples.

2. As illustrated in figure 5.5, the photo-generated charges are separated immediately after their generation and transported to the particular integration gate. Any accumulation of electrons is avoided below the photo-gates so that these CCDs can always be considered to be working in deep depletion mode. This ensures



(a) Illustration of the CCD gate arrangement viewed from the top with the indication of the four transistors within the pixel used for the amplification of the diffusion voltages.



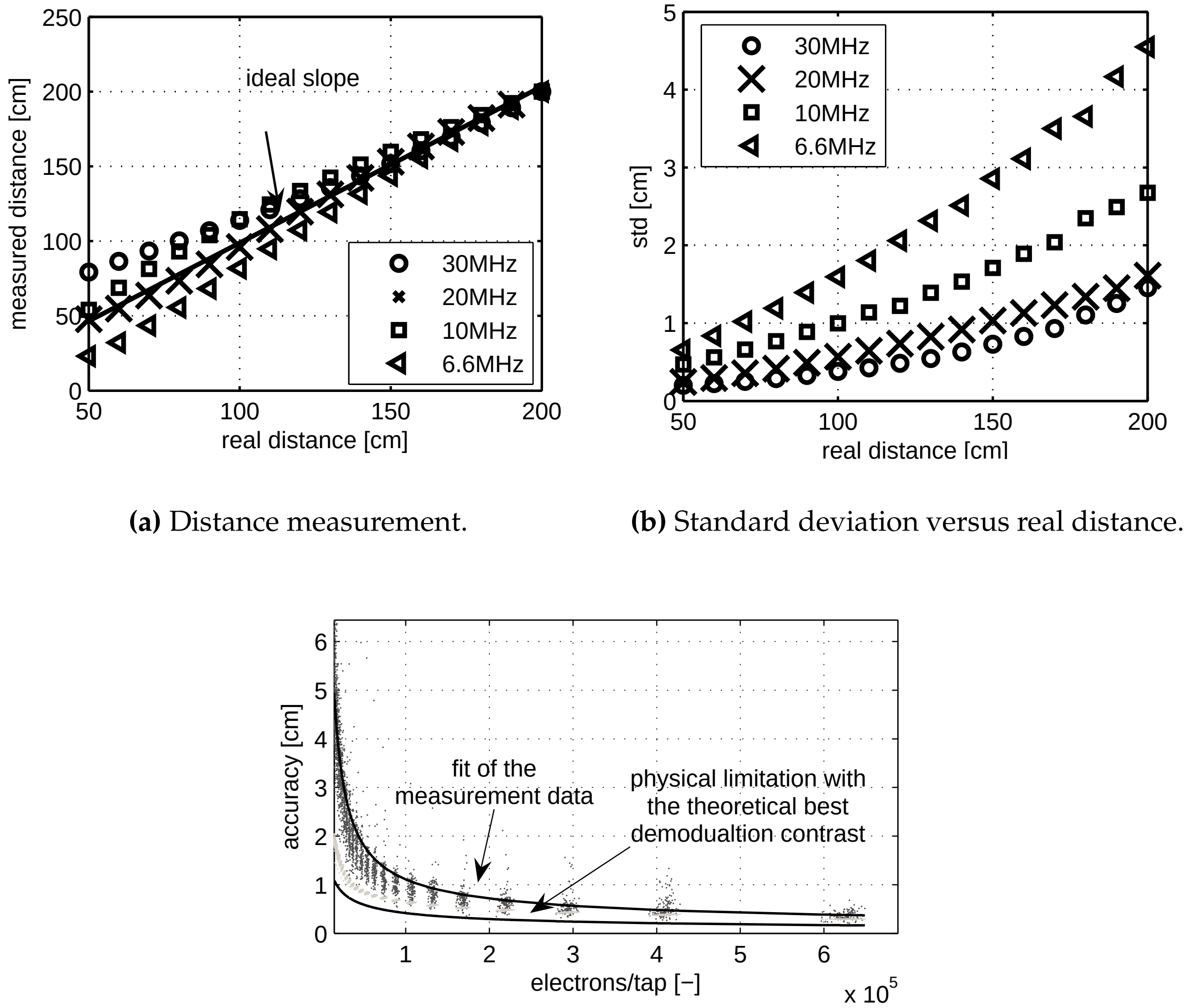
(b) Cross section of the two-tap pixel, showing the CCD gate arrangement and the potential distribution for two scenarios of the gate voltage settings.

Figure 5.5: Two-tap pixel.

the achievement of large quantum efficiency values that are independent on the integration status.

The distance measurement with a 2-tap pixel is subject to changes in the lighting conditions due to the requirement of two consecutive integration cycles. Therefore, one future challenge in 3D-TOF imaging will be the realisation of 4-tap pixels that have a large fill-factor and do not suffer from any mismatch problems. Estimations of the practicality of 4-tap structures in a $0.6\mu\mathrm{m}$ technology have shown that 4-tap pixels with a fill-factor of 1% less than that of a comparable 2-tap pixel with similar demodulation speed become possible [55], making the 4-tap approach even more attractive for the future.

Practical verification: A distance measurement from $50\mathrm{cm}$ to $200\mathrm{cm}$ with a 2-tap pixel sensor that is integrated into the SwissRanger 3D TOF camera [75], is shown in figure 5.6 (a) for the demonstration of the phase measurement principle's functionality. Modulation frequencies of 6.6, 10, 20 and $30\mathrm{MHz}$ have been employed, whereas a theoretical cut-off frequency of about $40\mathrm{MHz}$ is expected due to the dominance of slow diffusion processes in the charge transport [24]. The theoretical assessment of the cut-off frequency is discussed in detail in the next subsection. The different optical signal characteristics when using different frequencies explain the offsets of the graphs (the illumination unit was optimized for a $20\mathrm{MHz}$ signal). These offsets have to be calibrated out.
By depicting the 1-σ standard deviation of the same distance measurement versus the real distance (see figure 5.6 (b)), the basic relationships between the measurement accuracy, the effect of decreasing optical power at longer distances and the positive influence of high modulation frequencies becomes visible (see chapter 2 on the theory of measurement accuracy and the power budget).
Figure 5.6 (c) shows a direct comparison between the distance measurement and the theory by depicting the standard deviation versus the mean number of generated electrons per tap. This graph refers to the distance measurement with $20\mathrm{MHz}$ optical modulation signal. For a wide range of optical power, the CCD/CMOS pixels operate close to the photon shot noise limit. All measurements are accomplished with a sinusoidally modulated signal, and hence, equation 2.30 is used for the comparisons with theory. A mean demodulation contrast of $c_d \approx 40\%$ has been measured, while the theoretically best demodulation contrast is $c_{d,max} \simeq 64\%$ for a sampling interval that is half the modulation period.



(a) Distance measurement.

(b) Standard deviation versus real distance.

(c) Standard deviation versus mean number of photo-generated electrons per tap for the 20MHz measurement. Black dots are the measured data points and the gray dots correspond to the physical limitation based on the measured demodulation contrast.

Figure 5.6: Distance measurement with a two-tap pixel sensor. [24] The integration time was set to 4ms and the object had a reflectivity of 90% with properties of a Lambert reflector.

Limitations: By considering the simulation of the potential distribution in figure 5.5 (b), one can recognize three important aspects limiting this 2-tap demodulation pixel:

1. *Dynamic range:* The capacity of charge storage in the integration gate is limited by the gate size and voltage applied, so that highly reflective objects can lead to saturation and, hence, inhibit the distance measurement.

2. *Background light:* Charges generated by the optical modulation signal as well as by background illumination are stored below the integration gate, so that high background levels quickly lead to pixel saturation.

3. *Demodulation speed:* The potential characteristic is rather step-like instead of linearly decreasing, which leads to a dominating charge transport by thermal diffusion. Thus, the photo-gate is limited in its size with respect to being able to demodulate high frequencies; or large photo-gates reduce the maximum modulation frequency remarkably, making the pixel unattractive for high-precision distance measurement.

These limitations of the "standard 2-tap lock-in pixel" are overcome by the enhanced lock-in pixel structures, which are the topic of the following sections. We start this part of the thesis by introducing a generic model for estimating the demodulation speed of a pixel.

5.1.3 Demodulation Speed

The demodulation speed is determined by the 3dB cut-off frequency of the pixel. An assessment of the transit time of the electrons through the detection and demodulation regions becomes necessary. Based on that, a simplified model for the estimation of the cut-off frequency is discussed that can be applied in a general way to all pixel structures.

Transit Time

The transit time is defined as the the time that is needed for an electron to pass through the detection and demodulation regions to the storage node. Hence, the transit time is a composite of a drift and a diffusion component. Assuming a strict dominance of just one of the two charge transport methods within certain regions of the semiconductor, the transit time is expressed as

$$t_{tr} = t_{drift} + t_{diff}, \tag{5.1}$$

where t_{drift} and t_{diff} denote the total drift or diffusion times through N_{drift} drift regions and N_{diff} diffusion regions, respectively. The corresponding definitions of t_{drift} and t_{diff} are given below in equations 5.2 and 5.3. Induced drift fields are neglected because, during the operation of the demodulation pixels, the photo-generated electrons are moved immediately to the storage nodes so that an accumulation in the photosensitive region is avoided.

Following from the considerations of the transport methods of free charge carriers in semiconductors in section 4.3.2, the drift time is described by the following integral expression

$$t_{drift} = \sum_{g=1}^{N_{drift}} \int_0^{s_{max,g}} \frac{1}{v(E(s))} ds = \sum_{g=1}^{N_{drift}} \int_0^{s_{max,g}} \frac{v_{th} + \mu_0 E(s)}{v_{th} \mu_0 E(s)} ds, \tag{5.2}$$

and the diffusion time is given by

$$t_{diff} = \sum_{g=1}^{N_{diff}} \frac{s_{max,g}^2}{D}, \tag{5.3}$$

where s is the path and $s_{max,g}$ is the total path length within the g-th transport region.

The following aspects summarise the possibilities for the minimization of the transit time:

- The path lengths s_{max} could be reduced. However, in general the specifications of the required pixel size do not allow such a reduction of the sensitive area.
- The number of diffusion regions N_{diff} needs to be as small as possible or, in other words, the drift fields should dominate in most parts of the transportation-regions.
- The drift velocities should be maximized.

The last item, concerning the operation with maximum drift velocities, can be obtained with high lateral electric drift fields and by using semiconductor material with low doping concentrations. As shown in section 4.3.2, the drift velocity in electric fields reaches a maximum for doping concentrations less than $N_{dop} = 10^{16}/\mathrm{cm}^3$. Thus, the buried channel should be weakly doped.

Cut-Off Frequency

The model for the assessment of the 3dB cut-off frequency is based on the finite maximum transit time of the photo-generated electrons through the pixel. The impulse response function $h(t)$ describes the reaction of the pixel to an ideal delta-function (Dirac-pulse) of infinite optical energy. It is modeled as an ideal rectangular function as depicted in figure 5.7. The mathematical description is given by

$$h(t) = \frac{1}{t_{tr}} \cdot \mathrm{rect}\left(\frac{t - t_{tr}/2}{t_{tr}}\right) = \begin{cases} 1/t_{tr} & \text{if } 0 \leq t \leq t_{tr} \\ 0 & \text{if } t < 0 \vee t > t_{tr} \end{cases},$$

where the following condition for conserving the system's energy has been taken into account:

$$\int_{-\infty}^{\infty} h(t)dt := 1. \tag{5.4}$$

The real behaviour of the impulse response function depends on the pixel structure. In general, the rectangular function is rather deformed, as illustrated in figure 5.7, approaching more a Dirac-pulse like shape. Thus, the realistic behaviour of the impulse response function has less impact on the amplitude transfer characteristics, so that the theoretical model could be considered in the following rather as a tool for the worst-case estimation of the amplitude transfer characteristics.
The resulting photo-current corresponds to the convolution of the detected optical power signal with the response function and a subsequent multiplication by the responsivity R_{ph} of the pixel:

$$i(t) = R_{ph} p_r(t) * h(t). \tag{5.5}$$

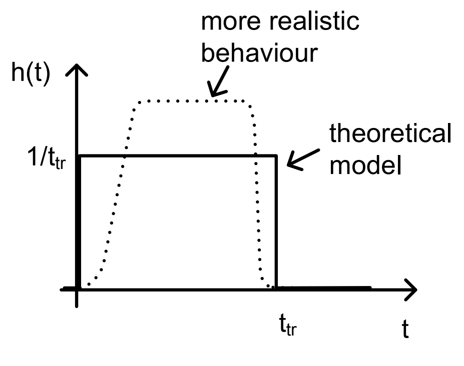


Figure 5.7: Impulse response function of a demodulation pixel depicting a simplified idealized model and a probably more realistic behaviour.

The convolution in the time-domain means a multiplication in the frequency domain so that

$$I(f) = R_{ph} P_r(f) \cdot H(f). \tag{5.6}$$

The frequency spectrum of the response function is expressed by the complex sinc function according to

$$H(f) = \mathrm{sinc}(\pi f t_{tr}) \cdot e^{-j\pi f t_{tr}}, \tag{5.7}$$

where the exponential term describes the phase shift. By defining a 3dB-cut-off frequency f_{cut} with the utilisation of the transit time t_{tr}, the following relationship can be derived from equation 5.7:

$$\mathrm{sinc}(\pi f_{cut} t_{tr}) \approx 0.71 \quad \Longleftrightarrow \quad f_{cut} \approx \frac{7}{16 t_{tr}}. \tag{5.8}$$

It is noted that the phase shift $\varphi = \pi f t_{tr}$ of the modulation signal is seen as an additional offset in the distance measurement, which can be compensated by calibration.

5.2 Enhanced CCD/CMOS Lock-In Pixels for High Dynamic Range

As shown in section 2.4, the dynamic range of the scene can be very demanding on the demodulation pixel. Firstly, there is the quadratic relationship between the object's distance and the detected optical power. Secondly, additional background light generates electrons that are stored in the integration gates as well. Both effects can lead to undesired saturation and hence, strongly limit the dynamic range of the imaged scene.

It is obvious that a reduction of the integration time can avoid saturation in both cases. However, since a shorter integration time affects all demodulation pixels of the sensor, it can happen that some pixels will not detect the minimum number of photons required for a specific measurement accuracy. Thus, the integration time has a lower boundary, which is given by the minimum photon threshold.
A second possibility for the avoidance of pixel saturation is the design of larger output capacitances. However, two drawbacks are coupled with this approach. 1. Larger capacitances need more space, reducing the fill-factor of the pixel. 2. The noise floor is higher with larger capacitances, shown for example by the equation that describes the thermal noise during reset:

$$\sigma_Q = \sqrt{k\vartheta (C_{dep} + C_{gd})}, \tag{5.9}$$

where the output capacitance corresponds to the sum of the diffusion node's depletion capacitance C_{dep} and the gate-drain capacitance C_{gd} of the source follower's transistor (see also figure 4.18). A negative impact on the distance measurement accuracy is implied by larger output capacitances. Thus, it is concluded that this approach does not provide a suitable technique for the creation of high dynamic range TOF pixels.

In the following two subsections, one new pixel concept particularly designed for the improved handling of significant changes in signal light power within the scene and another very efficient concept particularly developed for active in-pixel background light suppression are presented. Both concepts make use of analogue circuitries that control the process of charge transfer from the integration gates to the diffusion nodes, also called shift-process.
Figure 5.8 depicts the pixel's cross section through the CCD gate arrangement including an illustration of the fundamental processing steps of 1. the signal demodulation and accumulation and 2. the shift of the stored charge carriers to the sensing diffusion nodes. The standard lock-in pixel accumulates photo-generated charges for a given integration time, which is the same for all pixels of the sensor, and the full number of accumulated charges is shifted from the integration gates to the sensing diffusion nodes for read-out, though only the difference of the two charge packets is of interest for the distance measurement (this is valid for both sinusoidally as well as pseudo-noise modulated optical signals, as seen in chapters 2 and 3). The different operations of the newly implemented concepts are explained below.

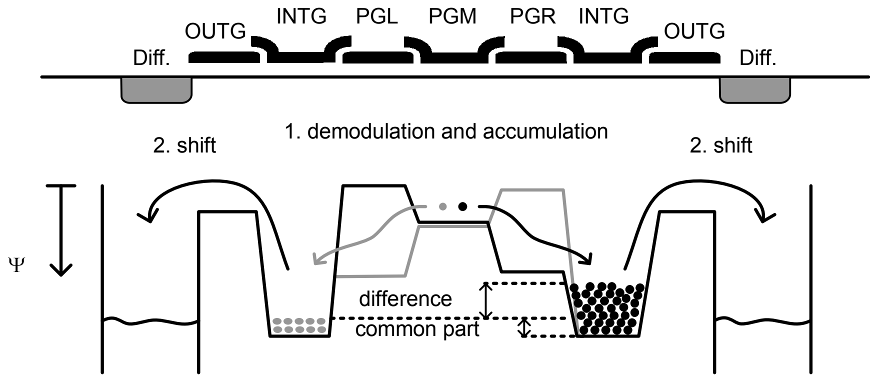


Figure 5.8: Lock-in pixel and its functional principle. [62]

5.2.1 Pixel with Individual Integration Time Control

A common method for increasing the dynamic range of 2D-imagers is the pixel-wise control of the integration time. This concept of pixel-wise integration (PWI) has been adopted to the demodulation pixels. It is worth mentioning that this approach of PWI only works with 4-tap pixels in order to ensure that all four samples are acquired with exactly the same integration time.

Principle of Pixel-Wise Integration

Due to the high-frequency coupling between the photo-gates and the read-out structure, no continuous observation of the status of integration is possible. Consequently, the totally specified integration time has to be divided into several sub-intervals. After each sub-interval a shift process is performed by setting the integration gate voltage from high to low level, enabling the comparison of the integrated charge at the sense diffusion node with an externally supplied threshold level V_{th}. This comparison is accomplished after all charge has been shifted by setting the threshold input of a comparator from low level to the threshold level. If this threshold level has been reached by at least one of the four taps, no more shift processes may be performed in order to avoid saturation of the diffusion nodes. Figure 5.9 illustrates the idea of step-wise integrating and shifting the photo-generated charges as long as the externally specified threshold level has not been reached. The in-pixel comparator activates a lock signal that prevents further shift processes. Electrons that are generated after the activation of the shift-lock signal are kept below the integration gate and the photo-gates until the reset signal also effects a shift of those charges and sets the diffusion node to the reset potential.
The threshold voltage can be set for example to half the level of saturation. Each time when the threshold voltage has not been reached, the next interval of integration could be doubled. Implementing such a control sequence minimizes the number of integration intervals.

Pixel Architecture

The enhanced pixel architecture is illustrated in figure 5.10. In this implementation, one pixel consists of two 2-tap structures. By controlling the photo-gates of the second structure, with the modulation signal shifted in time by a quarter of the modulation period, the outputs out1 to out4 correspond to the four samples. This 4-tap pixel structure is al-

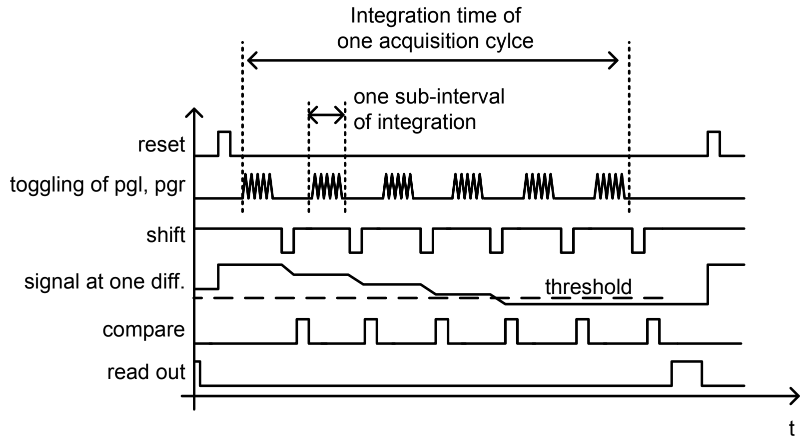


Figure 5.9: Timing diagram of the pixel-wise integration pixel. [62]

ternatively called a double 2-tap pixel because two spatially separated photo-sensitive regions are used.
The sensing diffusions are conventionally connected with the source follower circuits for read-out purposes but, in addition to that, they are feedback connected to an in-pixel comparator. This comparator sets a flip-flop *FF* as soon as the threshold voltage V_{th} is reached by at least one of the four taps. When the flip-flop is active, the integration gate is held at high voltage V_{high} so that no more shift processes can be externally forced by the control signal, denoted by *intg*.
The in-pixel integration of the comparator and the flip-flop requires as simple circuitries as possible with just a few transistor elements in order to still provide a tolerable optical fill factor. The fill-factor of the pixel realized in a $0.6\mu\mathrm{m}$ technology and having a pixel pitch of $60 \times 90\mu\mathrm{m}^2$, was approximately 8%. [62]

Experimental Results without Background Light

The concept of pixel-wise integrating demodulation pixels is verified by the distance measurement in figure 5.11, where the distance to a white target has been measured over a range from $70\mathrm{cm}$ to $4\mathrm{m}$. A test chip consisting of a 12×10 pixel matrix was used for the measurement. The PWI pixel is compared with the standard lock-in pixel by operating the same pixel structure either in PWI mode or in normal mode without the dynamic control of the shift process. The maximum integration time in the case of the PWI mode corresponds to the integration time specified in normal mode. A fixed number of five sub-intervals has been chosen for the PWI operation according to the principle of doubling the integration time after each sub-interval. The following relationships between the number of shifts and the realised integration times are

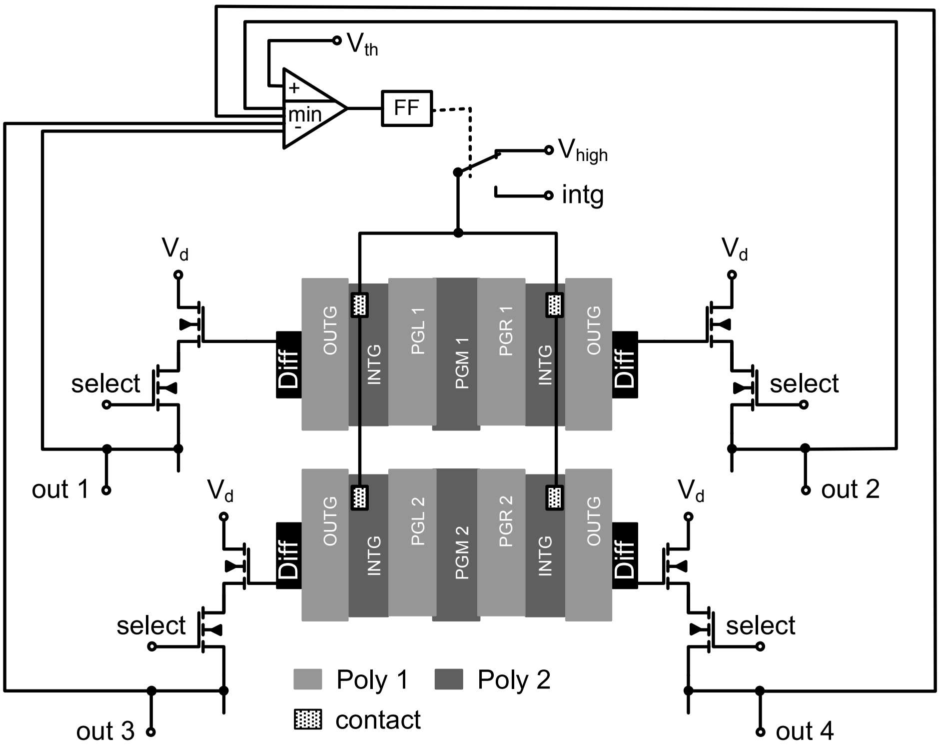


Figure 5.10: Illustration of the pixel architecture incorporating dynamic control of the integration time.

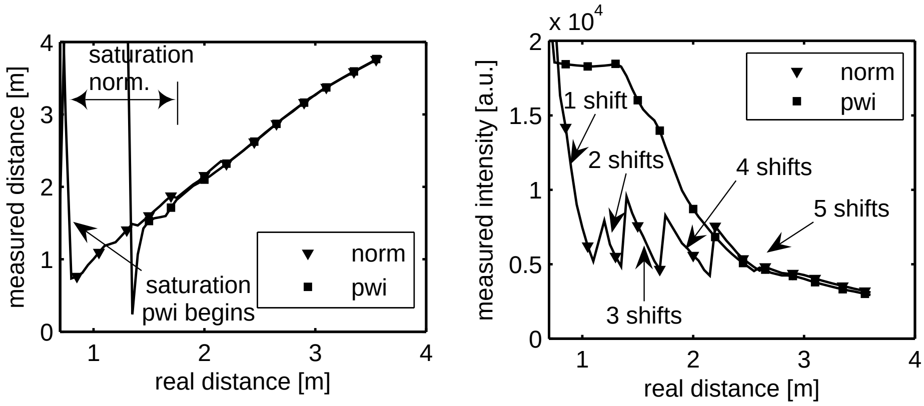


Figure 5.11: Distance measurement. The increased dynamic range of the PWI pixel compared to the standard lock-in pixel is demonstrated.

valid: 1 shift $\rightarrow$ 0.255ms, 2 shifts $\rightarrow$ 0.510ms, 3 shifts $\rightarrow$ 1.02ms, 4 shifts $\rightarrow$ 2.04ms and 5 shifts $\rightarrow$ 4.08ms. In the case of the PWI mode, the distance is correctly measured down to 70cm. The standard lock-in pixel already saturates at approximately 1.7m. The ratio between the maximum and minimum integration periods corresponds to a gain of 16 with regard to the number of electrons that can be stored in the pixel. Referring to equation 2.38, describing the quadratic relationship between optical power and object's distance, implies that the target can be moved towards the camera 4 times closer in PWI mode than in the normal mode of operation. However, we measured a factor 2.5, which is smaller than that expected due to the changes of the the optical conditions for different distances. These changes of optical conditions have two reasons: 1. No perfect coaxial test set-up was realised. 2. Different durations of integration imply different mean optical power emitted by the illumination module because the module reacts sensitively to temperature variations.
The depiction of the measured intensity in the right hand graph, of figure 5.11, based on the same distance measurement, shows the automatic switching between the number of shift processes in the pixel. At short distances the highest amount of optical power is detected, so just one shift process is performed.

Due to the fact that shorter integration times are automatically applied to each pixel individually, the full well capacity of the pixels is effectively increased by a factor 16. While the pixel operating in normal mode can provide 1 million electrons for read out, the mode of pixel-wise integration allows for reading out a number corresponding to 16 million electrons. Based on equation 2.39 the increase in the dynamic range is calculated as

$$DB_{se}(\text{PWI mode}) - DB_{se}(\text{norm. mode}) = 20 \cdot \log_{10}\left[\frac{FW(\text{PWI mode})}{FW(\text{norm. mode})}\right] \approx 24\text{dB}. \quad (5.10)$$

Table 5.1 shows the dynamic range of the sensor for distance resolutions from 1cm to 20cm. While the dynamic range varies from 24dB to 76dB when the sensor is operating in normal mode, the PWI sensor operation allows for covering dynamic ranges from 48dB to 100dB at distance resolutions from 20cm down to 1cm, respectively.

Table 5.1: Dynamic range of the TOF sensor operating either in PWI- or in normal mode.

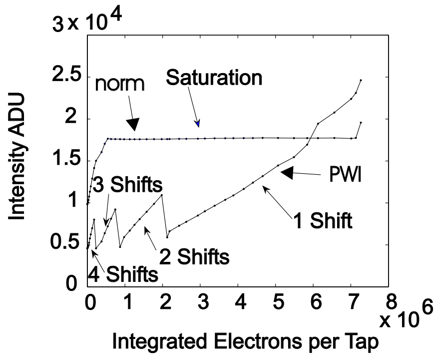
Distance resolution σ_R	Minimum number of stored electrons	DR_{se}(PWI mode) at σ_R	DR_{se}(norm. mode) at σ_R
20cm	150	100dB	76dB
15cm	280	95dB	71dB
10cm	620	88dB	64dB
5cm	2500	76dB	52dB
1cm	62000	48dB	24dB

Experimental Results with Background Light

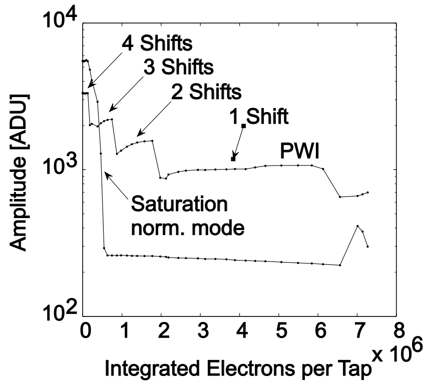
Another investigation includes the behaviour of the PWI pixel under background illumination. Obviously, background light implies shorter integration times in order to avoid saturation of the sensing diffusion, assuming that the full amount of integrated charge is transferred to the diffusion node for each shift process. If the signal light power is kept constant, two aspects will lead to reduced measurement accuracies:
1. The shorter integration time means detection of fewer photo-generated electrons.
2. The background light itself introduces an additional noise source.

The results of a distance measurement with a target at constant distance and with a fixed optical signal power are shown in figure 5.12. Background light is increased linearly. Four evaluations of the PWI and normal mode measurements show their different behaviour with respect to the detected intensity, amplitude, distance and standard deviation of the distance, depending on the total number of electrons generated by the background light per tap at longest integration time; this is a measure of the background light power. Again, the comparison between the PWI and normal mode of operation is based on the same pixel device.
The switching from one integration time to the other can be clearly seen in the intensity and amplitude plots (figures 5.12 (a) and (b)). Figure (c) shows a distance plot. The early saturation of the pixel in normal mode leads to completely erroneous distances. Though the PWI mode enables the distance measurement under light conditions generating up to 6 times more electrons, the distance acquisition is strongly affected by inhomogeneities between the four taps. A background light calibration becomes necessary. Such a calibration, however, can only be accomplished with difficulty because a calibration routine necessitates the same integration times between two measurements,

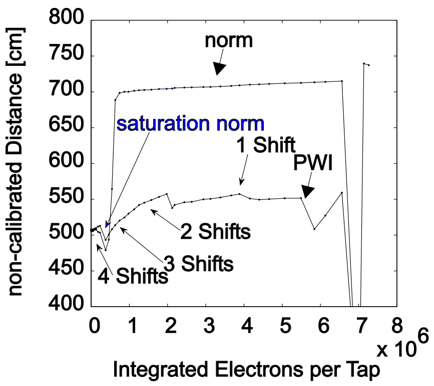
which is not automatically ensured by the PWI pixel.
Figure 5.12 (d) closes the summary of this measurement by depicting the standard deviation of the distance versus the intensity.



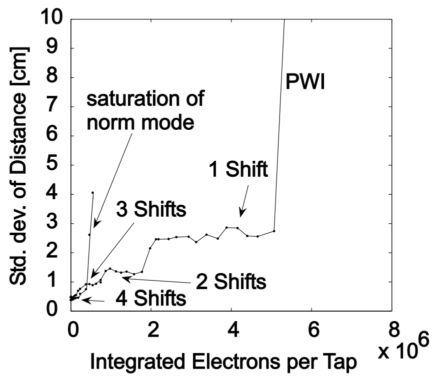
(a) Detected intensity versus the number of background electrons.



(b) Detected amplitude versus the number of background electrons.



(c) Measured distance versus the number of background electrons.



(d) Measured standard deviation of the distance versus the number of background electrons.

Figure 5.12: Experimental result obtained with the PWI pixel in the presence of background light and with constant signal light power and a fixed distance of the target.

5.2.2 Pixel Based on the Minimum Charge Transfer Principle

The second technique for increasing the dynamic range of 3D TOF sensors exploits the fact that only the difference between two charge packets is needed for the distance computation; this is the case for both sinusoidal and pseudo-noise modulation techniques (see chapters 2 and 3). As illustrated in figure 5.8, a significant common part of electrons due to background light and imperfect demodulation (the demodulation contrast is usually less than 1) is stored in both integration gates. This part of the charge packet does not need to be transferred to the sensing diffusion for read-out purposes. Thus, the storage capacity of the integration gate can be designed to be as large as possible, enabling the storage of a huge number of particularly background-light generated electrons, and the sensing diffusion node can be optimized in its size to the maximum power of the optical modulation signal and low noise contribution.
It should be mentioned at this point that an arbitrary size of the CCD gates does not decrease the pixel's noise performance because the operation is still in the charge domain. Thus, the drawback of a higher noise floor, as described in the introduction of the current section for the enlargement of the sensing diffusions, is not relevant for the scaling of the CCD gates.

The suppression of shifting the common part of electrons to the sense node implies the principle's name of "Minimum Charge Transfer" (MCT). The method can be applied to either 2-tap or 4-tap structures since no dynamic adaptation of the integration time between two acquisitions is performed in the pixel. The realised pixel makes use of a double 2-tap structure, having a fill-factor of 11% at a pixel pitch of $60 \times 70\mu\text{m}^2$ [62].
The basic idea was mainly motivated by the in-pixel suppression of background light. The principle is not well suited to the enhancement of the dynamic range based exclusively on the signal light power without any background illumination. In this case the PWI pixel should rather be used.

Principle of Minimum Charge Transfer

Figure 5.13 sketches the situation of electron and potential distribution in the demodulation pixel after the shift process has finished. The common part of electrons in both integration gates is kept in the integration gate area.
In order to arrive at this pixel status, the shift process is accomplished by decreasing the integration gate voltage with a given slew rate. As soon as all sensing diffusions of the pixel structure have obtained some electrons, the reduction of the integration gate voltage is stopped. The observation of the sensing diffusions is done by a comparator

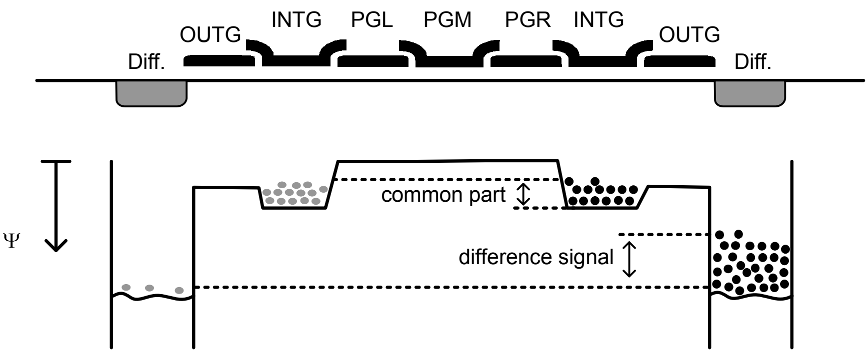


Figure 5.13: Illustration of the MCT principle after the shift process. As soon as both sensing diffusions have obtained electrons during the shift process, the integration gate potential is frozen so that the common part of electrons is kept in the CCDs.

circuit that deactivates the dynamic integration gate signal when all diffusion nodes cross an externally specified threshold voltage V_{th} very close to the reset voltage. The corresponding timing diagram is shown in figure 5.14, illustrating the procedure just described. The hard shift process during the reset sets the integration gate to $0\mathrm{V}$. It is necessary for dumping the common-mode electrons that are still kept in the CCD region.

Pixel Architecture

The realised double 2-tap pixel architecture is schematically depicted in figure 5.15. In comparison to the PWI pixel architecture, the required comparator circuitry of the MCT pixel does not have to be included in each pixel because there is no individual-

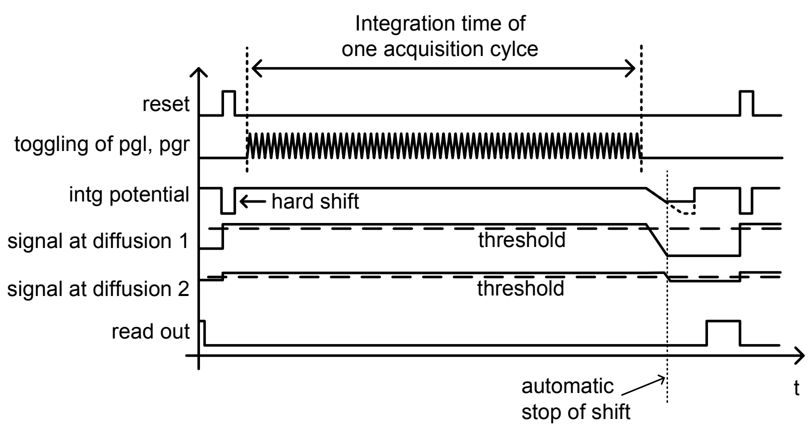


Figure 5.14: Timing diagram of the MCT pixel.

pixel control. Here, the monitor circuit can be applied on a column basis, enabling a line by line read-out scheme. Two additional transistors have to be added within the standard lock-in pixel for switching between the fixed high-level gate voltage V_{high} or the integration gate signal *intg* externally supplied by the comparator circuit. This ensures a high optical fill-factor of the pixels.
The comparator acts as a monitor of the four taps. The shift process takes place when the pixel is selected. During the process, the integration gates are connected to the external signal that shows a linearly decreasing behaviour over time as sketched in the timing diagram in figure 5.14. As soon as all taps begin to cross a certain threshold voltage, which is specified from the external electronics and is close to the reset voltage, the monitor circuit holds the integration gate voltage on a constant level. This means the abrupt interruption of the shift process according to the idea of the minimum charge transfer.

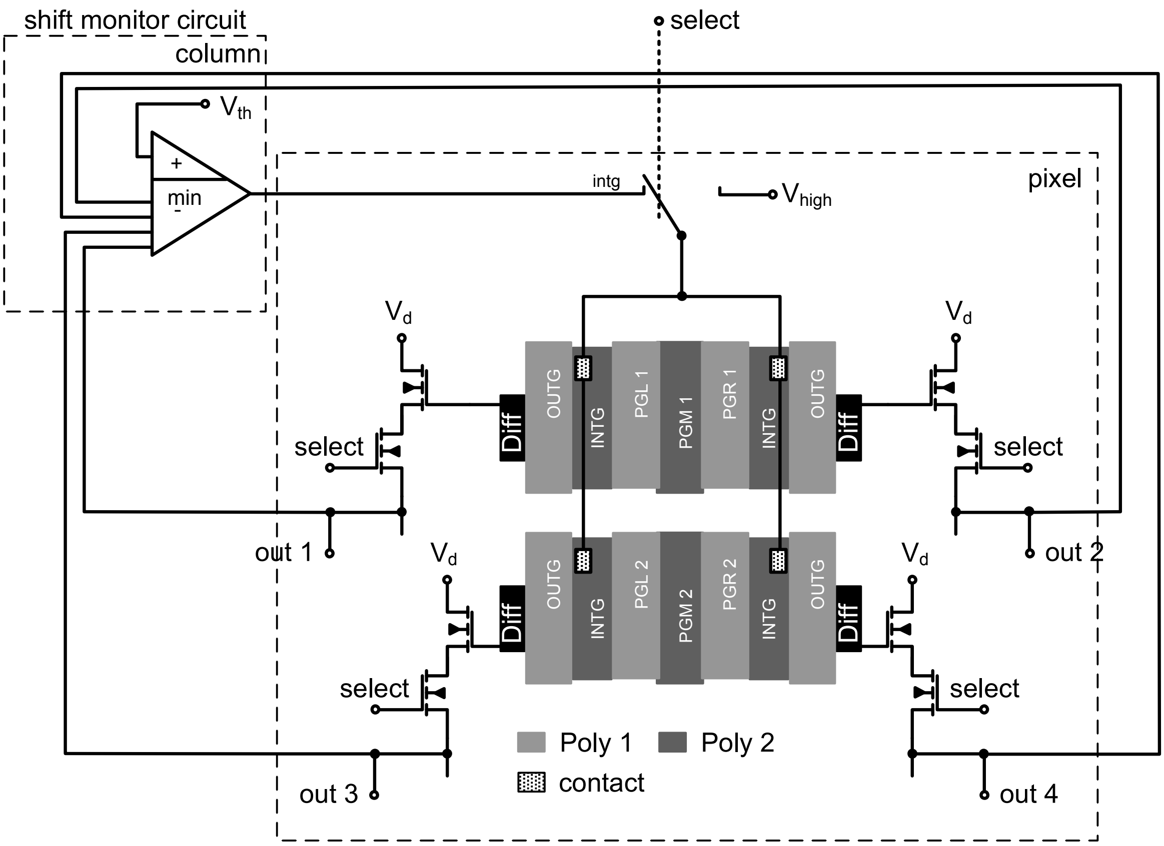


Figure 5.15: Illustration of the MCT pixel structure having a direct connection to the column-wise comparators that monitor the shift process.

Experimental Results

A test chip of 12×13 demodulation pixels with an additional row of shift monitor circuits at the bottom of the pixel matrix was used for the concept's practical verification. The achievable gain of the MCT pixel is best quantified by comparing the storage capacity of the diffusion node with the integration gate capacity. While the diffusion node limits the standard pixel operation when background light is present, the MCT pixel is rather limited by the integration gate capacity. The gain G_{mct} is defined as the following ratio

$$G_{mct} = \frac{Q_{sense}}{Q_{gate}}, \tag{5.11}$$

where Q_{sense} and Q_{gate} describe the maximum amounts of stored charge in the sensing diffusion and the integration gate, respectively.

The first measurement, of which the results are graphically shown in figure 5.16, provides the comparison between the capacities of the diffusion node and the integration gate. In figure (a), the pixel, operating in normal mode, detects the background light intensity while increasing the background light power continuously. The pixel saturates when one tap has accumulated approximately $950,000$ electrons in the sensing diffusion.

Figure (b) illustrates the accuracy of a distance measurement where the MCT function-

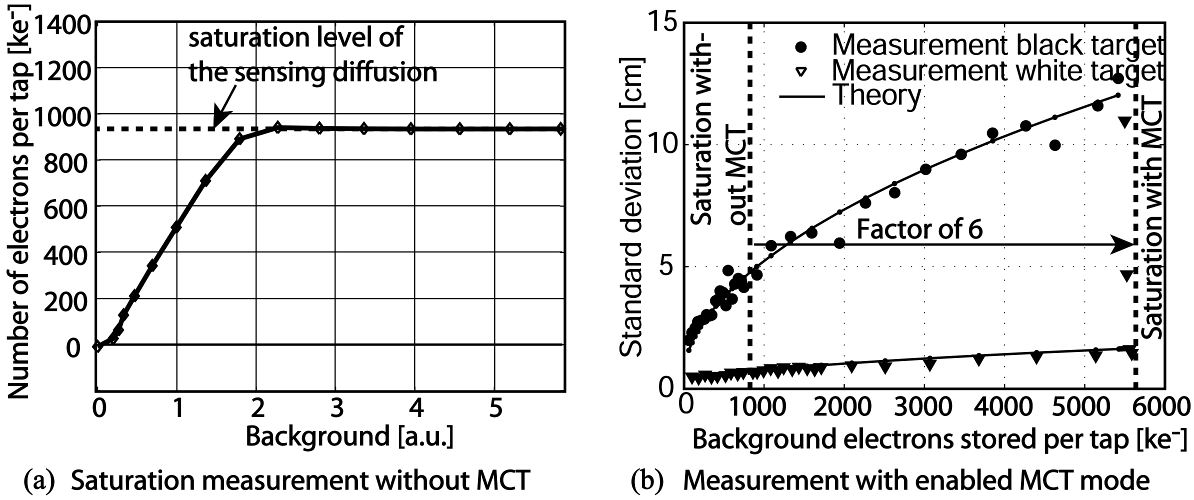


Figure 5.16: Measurement result with and without MCT functionality. Firstly, the gain in capacity for storing electrons which are generated by background light, is illustrated. Secondly, the operation of the distance measurement at the physical limitation given by photon shot noise is verified by comparison with the theoretical expectation.

ality has been switched on. Two measurement series are included: one with a highly-reflective target (white target) and one with a strongly absorbing target (black target), having reflectivities of 90% and 5%, respectively. The high-reflectivity target leads to the generation of approximately $285,000$ electrons, while the optical signal reflected by the black target generates approximately $34,000$ electrons per tap (it is noted that those values correspond to the specified integration time of this measurement sequence).
The measurement is firstly used to point out the significantly higher level of accumulated electrons per tap where the pixel saturates due to the detection of background light. By considering the measurement series with the black target, the number of electrons generated by the signal component is negligibly small compared to the background electrons. Thus, one can learn from this graph that the pixel saturation becomes effective at a number of approximately 5.8 million accumulated background electrons. Hence, this corresponds to a gain in background light rejection capability of the MCT pixel of at least a factor $G_{mct} = 6$ compared to the normal mode of operation.
Secondly, the accuracy measurement shows a good agreement with the theoretical expectation when the photon shot noise is the dominating noise source. The theoretical curves have been generated with the equation 2.30 (page 21), where the number of amplitude electrons and background electrons were measured and the constant noise floor was set to zero.

A more pragmatic or application-specific analysis of the MCT pixel is given in [76]. It is shown that this test pixel structure can efficiently be used for suppressing the background light power of indirect incident sunlight under practical circumstances. This is particularly interesting for the application of a MCT-based 3D camera in outdoor environments.
The following assumptions are made: the full sunlight power is impinging on a diffusively reflecting object; the integration time is $2.5\mathrm{ms}$ or less; an optical bandpass filter of $40\mathrm{nm}$ width is used. By employing the relationships of the system's power budget as described in section 2.3 and the spectral solar irradiance (see figure 2.12), the measured number of background electrons can be attributed to an equivalent ratio of background-light-to-sunlight power on the object with respect to the before-mentioned assumptions. The result shows that the maximum intensity of background illumination with which the pixel can deal is approximately 1.3 times larger than the natural sunlight intensity.

The design of the MCT pixel can obviously be adapted to the particular application. Depending on the required fill-factor and pixel pitch, even more efficient background suppression capability could be easily obtained by enlarging the size of the integration gate.

5.2.3 Comparison of PWI and MCT Pixel

Both pixel approaches for the increase of the dynamic range have their advantages and disadvantages. Whereas the MCT pixel is extremely robust against background illumination, the PWI pixel is preferably used for scenes with less background light such as in indoor applications. There the optimized integration time for each pixel results in improved distance accuracy.
The clear superiority of the MCT pixel in background light suppression is illustrated in figure 5.17. A measurement of the distance standard deviation is plotted for all three operation modes (normal, PWI and MCT operation) depending on the background light normalized to the incident sunlight power. While the standard demodulation pixel saturates already at background light power of 0.3 times the sun's brightness, the PWI and MCT pixel still deliver reliable distance data. In particular the MCT pixel operates well up to background light conditions corresponding to 1.3 times the sun's brightness.

A broader comparison between the PWI, MCT and standard pixels is shown in the following table 5.2, mainly summarising again the results discussed above.

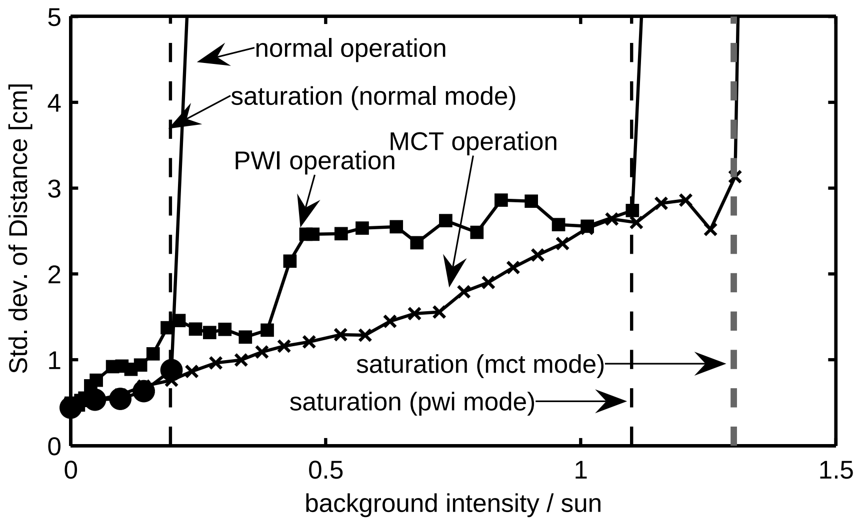


Figure 5.17: Comparison of the background light suppression capabilities between the different operation modes: PWI, MCT and normal mode.

Table 5.2: Comparison of the MCT and PWI pixels. (a.s. means application specific.)

	MCT test pixel	**PWI test pixel**	**MCT general**	**PWI general**	**Standard pixel**
Pixel structure	double 2-tap	double 2-tap	2-tap, double 2-tap, 4-tap	double 2-tap, 4-tap	N_t taps
Pixel pitch	$60 \times 70\mu\mathrm{m}$	$60 \times 90\mu\mathrm{m}$	a.s.	a.s.	a.s.
Comparator	column-wise	pixel-wise	column-wise	pixel-wise	none
Fill factor	11%	8%	=	↓↓	↑
Full well at sense node	$0.95 \cdot 10^6$ e$^-$	$1 \cdot 10^6$ e$^-$	a.s.	a.s.	a.s.
Charge handling at integration gate	$5.8 \cdot 10^6$ e$^-$	$> 1 \cdot 10^6$ e$^-$	>> full well at sense node	≥ full well at sense node	≥ full well at sense node
Dynamic range of the signal	improved	strongly improved	↑	↑↑	=
Background suppression	very efficient	efficient	↑↑	↑	=

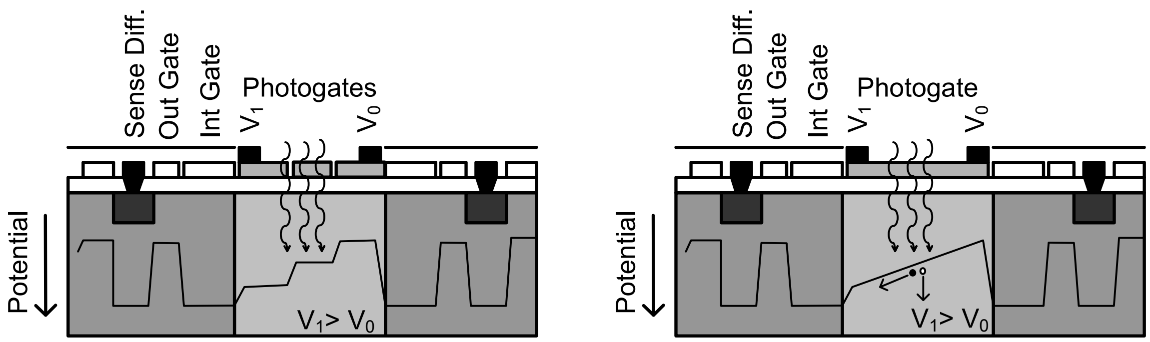
5.3 Pixels for High-Speed Demodulation

Two major concepts of pixel architectures for the demodulation of high optical modulation frequencies up to several hundreds of Megahertz are introduced in the following two subsections. They represent a major contribution of this thesis work to the problem of high-frequency demodulation. In order to achieve high-speed functionality in the pixel, the signal delay due to the time-consuming charge transport needs to be minimized. This can be accomplished by pixel structures that use lateral electric drift fields instead of thermal diffusion processes (see section 4.3.2).

The first approach uses modulated lateral electric drift fields for a more efficient and quick charge transport to the integration gates. The second architecture, in turn, is based on static drift fields, where a clear separation between the detection and the demodulation region has been made. In particular the last concept is very promising due to several advantages such as high speed demodulation, low power consumption and low requirements to the external driving electronics due to minimized in-pixel capacitive loads.

5.3.1 Pixel Using Modulated Electric Drift Fields

Fast charge transportation from the detection region to the storage nodes requires lateral electric drift fields. The creation of potential steps in the semiconductor such as is the case in the three-gate demodulation pixel, have to be avoided. This leads to the idea of replacing the discrete three-gate structure by an one-gate structure that is formed by a very highly resistive material.
The conceptual idea is shown in figure 5.18, where the left hand figure highlights again the disadvantageous step-wise potential distribution of the three-gate demodulation pixel. This three-gate arrangement is replaced by the high-resistive gate as shown in figure (b), which enables the generation of a constantly decreasing potential in the semiconductor. By applying a potential difference at the left and right corner of the high-resistive gate of $V_1 - V_0$, a current begins to flow through the gate implying a constant potential decrease. As in conventional CCDs, the gate potential is seen at the semiconductor's surface, generating there the desired lateral drift field.



(a) Schematic of the three-gate structure of the standard lock-in pixel design.

(b) Principle of the high-resistive one-gate structure. [22]

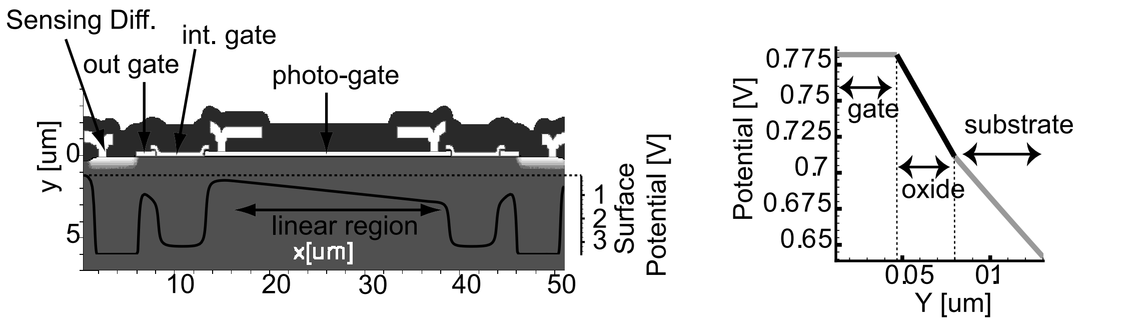
Figure 5.18: Comparison between the standard pixel consisting of three photo-gates and the drift field pixel having incorporated just one large photo-gate of high-resistivity material. While the standard pixel generates a step-wise potential distribution in the semiconductor, the drift field pixel creates a characteristic of constantly decreasing potential.

The result of a process and device simulation in figure 5.19, using high-resistivity poly-silicon gate material of $30\mathrm{k}\Omega/\square$, confirms the concept of generating lateral electric drift fields in the semiconductor material. The simulation assumes a charge transport in the surface channel, determined by the maximum potential depicted in the figure, which is located at the semiconductor insulator interface (cross section at $y = 0.1\mu\mathrm{m}$). In that simulation the integration gates are set to $3\mathrm{V}$ and the gate contacts are set to $0\mathrm{V}$ and $1\mathrm{V}$, respectively.
The potential decrease through the vertical cross section at $x = 20\mu\mathrm{m}$ is shown in figure 5.19 (b). It confirms the appropriateness of the approximation of neglecting the potential drop in the oxide layer so that the surface channel potential can be set equal to the gate potential. Consequently, a lateral drift field at the semiconductor insulator surface can approximately be computed as

$$E(x) = -\mathrm{grad}(V_g(x)), \tag{5.12}$$

where $V_g(x)$ is the gate potential. Based on the knowledge of the potential distribution in the semiconductor, a transit time and a cut-off frequency can be estimated according to the descriptions in section 5.1.3, which is done for the particular pixel implementation below.

Similar gate devices, based on high-resistivity gates for the generation of constant drift fields and a static charge transport, had been developed and tested already by Hoffmann [48], who intended to use such gate arrangements in shift registers for example to build very fast transmission lines in memory devices. However, the demodulation pixels presented here exploit the drift field in a dynamic fashion, modulating the drift field according to the reference signal of the sampling or the correlation process. Thus,



(a) Process and device simulation; potential gradient at y = 0.1um.

(b) Potential gradient at x = 20um.

Figure 5.19: Process and device simulation of the 1-dimensional cross-section of a drift field pixel, formed by a high-resistivity poly-crystalline silicon material. [22]

the gate voltages V_0 and V_1 are toggled as the left and right photo-gates of the standard lock-in pixels. If sinusoidal modulation is used, this means a modulation of the drift field synchronously with the optical modulation frequency.
The storage and read-out of the photo-generated charges are realised in the same fashion as in the standard lock-in pixels. The integration gates adjacent to the photo-gate contacts are set to high potential during the accumulation process and to low potential for shifting the charge via the out gate to the sensing diffusion node. Each sensing diffusion node has its own source follower circuit.

Gains and Losses of the Modulated Drift Field Pixel

Compared with the standard lock-in pixel architecture, the idea of using modulated drift fields promises some important advantages, but it also incorporates some significant performance penalties. The following summary provides a brief overview of the benefits and costs similarly described in [22].

Advantages:

1. The dominating factor of the charge transportation method is the drift phenomenon. That allows high modulation frequencies and high distance measurement accuracies.

2. Due to the fast charge transportation, the photo-sensitive area can be designed to be large, so that splitting into several photo-sensitive areas does not become essential for large-area demodulation pixels.

3. The reduction of the photo-gate structure to just one high-resistivity gate simplifies the pixel design.

4. Large sensitive areas can be combined with small conversion capacities.

5. Since the fast charge transport enables stretching of the photo-gate in both directions (seen from the top), square-shaped photo-gates can be realised, which allow the efficient use of micro-lenses for improved fill factors.

Disadvantages:

1. The realisation of high-resistivity poly-silicon photo-gates is not necessarily a standard option of the applied CMOS process, which might increase the processing costs.

2. The permanent current flow through the photo-gate leads to elevated ohmic on-chip power consumption. Furthermore, depending on the strength of the current flow, a warming of the sensor has probably to be taken into account in the evaluation of the dark current.

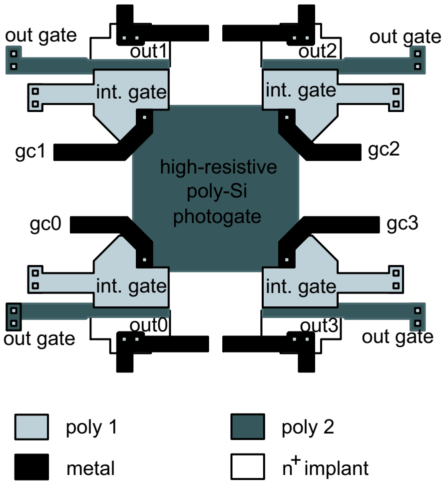
3. The need for very high-resistivity material for reducing the current flow through the photo-gate competes with the need for low resistivities for keeping the RC-times low and hence for ensuring high-speed switching performance of the pixel.

4. For large array sensors, the requirements of the driver electronics are high due to the capacitive and resistive loads. Sensors of high lateral pixel resolution become difficult to realise.

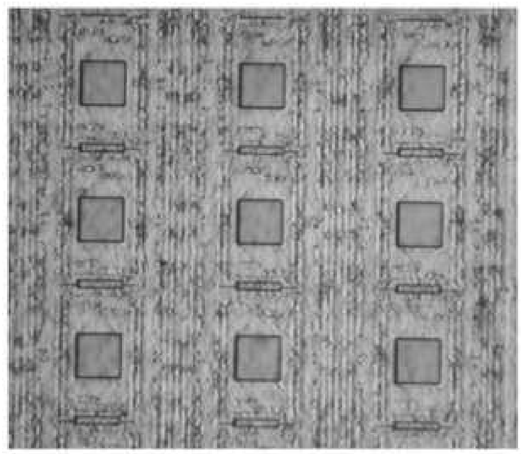
Realisation of a 4-Tap Drift Field Pixel and Estimation of the Demodulation Performance

For the verification of the pixel concept, a 4-tap drift field pixel has been realised in a $0.8\mu\mathrm{m}$ CMOS process offering an additional CCD option and the growth of a very highly-resistive poly-silicon gate with a resistivity of approximately $30\mathrm{k}\Omega/\square$ [7]. The first pixel implementation makes use of a charge transport in a surface channel. With a pixel pitch of $85 \times 97\mu\mathrm{m}^2$, a fill factor of approximately 13% could be achieved, which is rather high for a 4-pixel structure. The quadratic-shaped photo-gate has a side length of $25\mu\mathrm{m}$. The design is shown in figure 5.20. Four gate contacts, denoted by $\mathrm{gc0}...\mathrm{gc3}$ are placed at the corners of the photo-gate. The appropriate control of the gate contact signal as shown in figure 5.21, enables the sampling of the optical sinusoidally modulated signal.

The 4-tap pixel requires a simulation of the two-dimensional potential distribution on the photo-gate, which gives a different result from the ideal constant drift field in the case of a one-dimensional simulation in figure 5.19. The two-dimensional simulation of the rectangular shaped high-resistivity photo-gate is based on the resistor network model, shown in figure 5.22 (a). Standard mesh analysis [44] results in the potential distribution on the gate. The desired distribution in the semiconductor is obtained by



(a) Design of one drift field pixel.



(b) Micrograph of the 3×3 drift field pixel matrix.

Figure 5.20: Implementation of the 4-tap pixel, which is used for the generation of modulated lateral electric drift fields inside the semiconductor bulk.

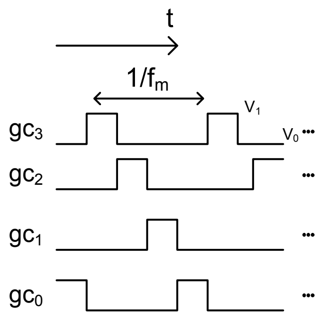
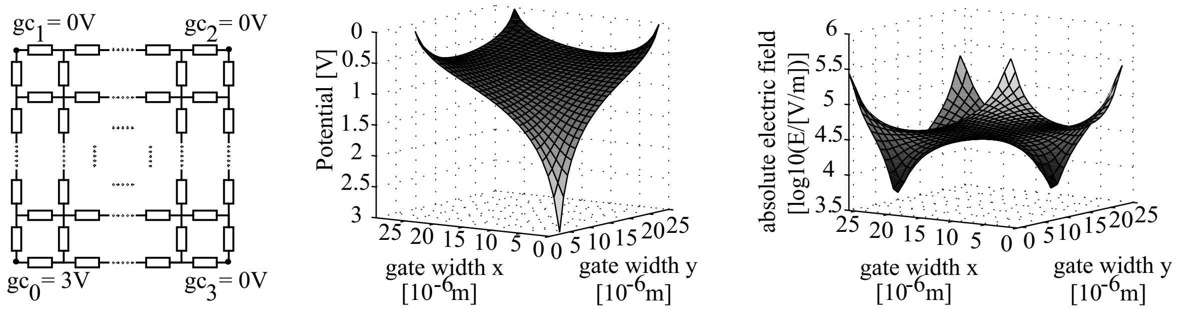


Figure 5.21: Control of the gate contact voltages by switching them synchronously with the modulation frequency between high level V_1 and low level V_0.

a smooth attenuation by a factor of 0.9, which is the simulated ratio between the surface and gate potentials (extracted from figure 5.19 (b)). The distribution of the surface potential is illustrated in figure 5.22 (b). The corresponding distribution of the absolute lateral electric drift fields, numerically calculated with equation 5.12 and depicted in figure 5.22 (c), clarifies that the expected fields are no longer constant over space. Thus, the photo-generated electrons experience in some regions a stronger and in some other regions a weaker electric force. The fields are of the order of $10^3 ... 10^6 \mathrm{V/m}$. Hence, in certain areas at the corners of the photo-sensitive region, the electric fields cause almost saturation velocity, which is reached for electric fields of about $10^6 \mathrm{V/m}$ at room



(a) Resistor network.

(b) Characteristics of the gate potential.

(c) Characteristics of the absolute electric field.

Figure 5.22: Drift field photo-gate: discrete component model and simulation result. [22]

temperature (see figure 4.16).

In order to obtain an estimate on the theoretical performance of the pixel, the cut-off frequency is modeled by following the approach of section 5.1.3. The upper bound of the electron transit time is said to be approximated by the diagonal path from a corner of low potential to the opposite corner, which is set to high potential (from gate contact $gc2$ to $gc0$ in figure 5.22). Based on the numerical simulation performed with gate voltages of $0\mathrm{V}$ and $3\mathrm{V}$, the transit times of electrons generated at a certain point on the diagonal connection between the high-level and opposite low-level corner are extracted and illustrated in figure 5.23, where the calculation of these data points is based on equation 5.2. While the highest lateral electric fields near the gate contact of high potential also produce the shortest transition times, the longest time of approximately $t_{tr} \approx 7\mathrm{ns}$ is given over the complete diagonal path. This transit time corresponds rather to a worst-case estimation of the cut-off frequency, which is computed to be

$$f_{cut} \geq \frac{7}{16 \cdot 7.5\mathrm{ns}} \approx 60\mathrm{MHz}.$$

A zero-order approximation of the cut-off frequency would use the average transit time of $t_{tr} \approx 3.2\mathrm{ns}$ (mean of the data in figure 5.23), which means that the pixel supports demodulation frequencies of more than $135\mathrm{MHz}$.

The demodulation speed is strongly dependent on the particular size of the photo-gate and the applied gate contact voltages. As an example, a quadratic-shaped photo-gate of $15\mu\mathrm{m}$ side length and gate contact voltages toggling between 0 and $6\mathrm{V}$ would lead to an average transit time of $600\mathrm{ps}$ or cut-off frequencies beyond $700\mathrm{MHz}$. Thus, it is worth

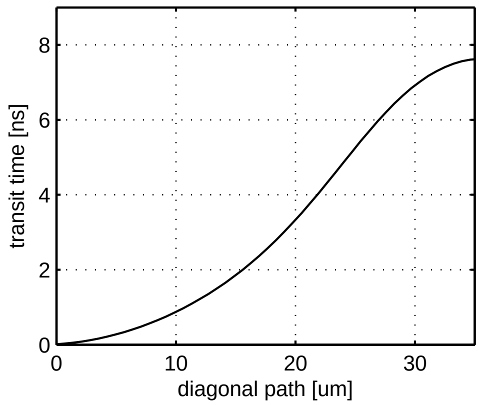


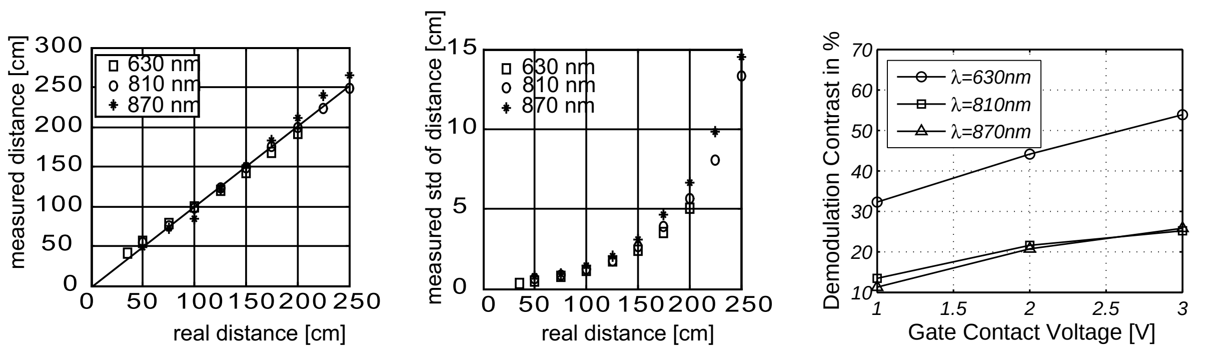
Figure 5.23: Transit time of an electron that has been generated at a certain point on the diagonal connection between $\mathrm{gc0}$ and $\mathrm{gc2}$ and that will drift to the high-level set gate contact $\mathrm{gc0}$. This figure applies to the simulation of a quadratic shaped photo-gate with $25\mu\mathrm{m}$ side length and toggling voltages between 0 and $3\mathrm{V}$.

remembering that the considerations made above as well as the following experimental results belong to the particular pixel implementation.

Another aspect of pixel performance should briefly be mentioned at this point: the dark current, in stochastic terms modeled as shot noise, has a direct impact on the distance measurement accuracy. Basically, the dark current is rather a measure of process quality, which is why it is not investigated in more detail in this work. However, the special structure as well as the operation principle of the drift field pixel makes an evaluation of the dark current necessary. Firstly, pixel geometries with pronounced edges like the rectangular geometry of the drift field pixel, show a lot of traps, increasing the possibility of random carrier generation, which implies higher dark currents [22, 99]. Secondly, any warming of the sensor due to for example the ohmic power dissipation in the photo-gate of the drift field pixel increases the dark current. In [22] a dark current model distinguishing between two components, one of the sensing diffusions and one dark current generated in the CCD channel beneath the photo-gate, is applied in practice. Measurements show that firstly the dark current does not vary significantly compared to the dark currents of standard CCDs in the same technology. Thus, no significant increase due to any sensor warming could be seen. Secondly the measured strength of the dark current allows this parameter to be neglected in the estimation of the distance measurement if integration times of milliseconds or shorter are used.

Experimental Results

Practical verification of the modulated drift field concept has been tested by measuring real distances from $35\mathrm{cm}$ to $250\mathrm{cm}$ using continuous-wave optical modulation of $20\mathrm{MHz}$. The results are shown in figure 5.24, where the linearity is depicted in (a), (b) shows the corresponding accuracy and (c) illustrates the result of the evaluation of the demodulation contrast. The measurement is performed with a set of three different optical wavelengths: $630\mathrm{nm}$, $810\mathrm{nm}$ and $870\mathrm{nm}$. Furthermore, the demodulation contrast is compared for different swings of the gate contact voltages.
While the systematic errors are explained by systematic distortions of the optical modulation signal (see section 2.2.5), the measurement accuracy follows a quadratically increasing characteristic, which is due to reduction of detected optical power for distant objects. By considering equation 2.30 and recalling the inverse proportionality between the number of signal electrons and the square of the object's distance, it is obvious that parasitic noise sources of the external electronics dominate the measurement accuracy This is a very unsatisfactory effect, which originates from the test set-up.
Following from figure 5.24 (c), the demodulation contrast seems to attain a maximum value for gate contact voltages that are slightly greater than $3\mathrm{V}$. That is the reason why at least $3\mathrm{V}$ have been used in the distance measurement. On the other hand, even higher gate contact voltages have been avoided in order to minimize the power consumption of the sensor, decreasing the requirements on the driving electronics. The dependence of the demodulation contrast on the gate contact voltage is described by the fact that the fringing fields at the edges of the electrode counteract the lateral field



(a) Linearity of the distance measurement and ideal slope.

(b) Stochastic error of the distance measurement (over 200 iterations).

(c) Demodulation contrast at 20MHz modulation.

Figure 5.24: Measurements with the modulated drift field pixel.

under the electrode.
The imperfect shutter efficiency and the limited transfer time of the charges within the substrate are one reason for the reduced demodulation contrast compared to the theoretical best contrast of 90%. The dependence on the optical wavelength is explained by the fact that photons of larger wavelengths reach deeper into the semiconductor, where they are not so strongly affected by the high electric lateral fields near the Si/SiO_2 interface. Hence the separation is less efficient. [22]

5.3.2 High-Speed Pixel Based on Static Drift Fields

Inspired by the modulated drift field pixel with its great advantages of generating lateral electric drift fields for fast charge transportation, but also the coexistent drawback of high power consumption, a second pixel structure has been developed. This second drift field pixel is rather based on static lateral electric drift fields and overcomes the problems of high in-pixel power consumption and low-pass filtering effects of the gate signals due to the huge resistive and capacitive loads given by the photo-gate. At the same time, very high demodulation frequencies are supported.

Pixel Concept

The basic idea of the pixel concept is the strict separation of the different functional blocks of photon detection, signal demodulation, signal integration and output amplification. A block diagram of the pixel architecture is illustrated in figure 5.25. In comparison with the aforementioned pixel structures, the difference is the additional spatial separation of the detection and demodulation regions, which were previously combined.

The photon-electron conversion is accomplished in the detection region, which is designed to be relatively large area. Lateral electric drift fields enable the transport of the photo-generated charge carriers into the direction of the demodulation region at high speed. This process introduces only a very short delay to the signal, generally less than 1ns. The photo-current arriving at the demodulation region is sampled or correlated by switching appropriately between the in-pixel storage capacitances. The demodulation region is designed to be as small as possible in order to minimize the diffusion-based charge transport regions and hence the response time. Standard source followers are used for the charge amplification.

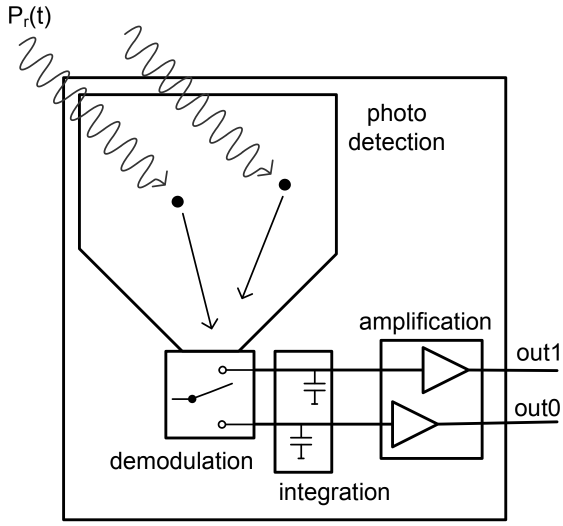
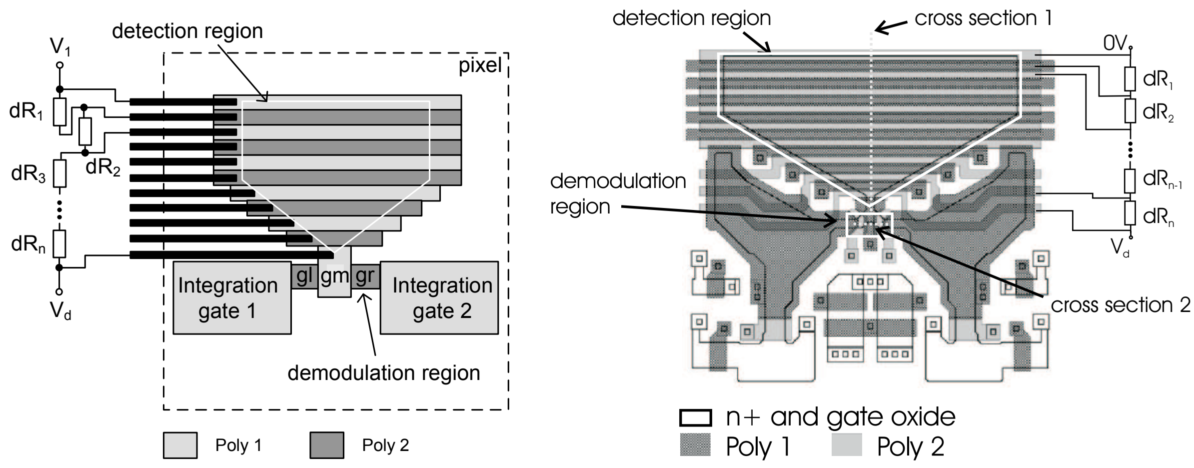


Figure 5.25: Block diagram describing the pixel's demodulation concept.

Realisation of a Static Drift Field Pixel

While the demodulation stage can be realised as a gate arrangement similar to that in the standard demodulation pixels but shrunken in size, the detection region needs a special gate arrangement that allows for the generation of a static drift field. One possibility is to use a one-gate structure of high-resistive poly-silicon material. However, the desire for low in-pixel power consumption led to another gate arrangement, such as that depicted in figure 5.26 (a). The corresponding layout is shown in figure 5.26 (b).

The lateral electric drift field in the semiconductor channel is generated by 18 adjacent photo-transparent poly-silicon gates of minimal width. The gates are set to constant voltages, which linearly increase from 0V to $V_d = 7$V. The generation of these voltages is performed outside the pixel by a resistor line. According to the CCD principle, the potential characteristic on the gates is duplicated at the semiconductor-oxide interface due to the capacitive coupling between the gates and the bulk. The smearing of the potential distribution in the CCD channel, particularly if buried channel CCDs are used, enables the creation of constant lateral drift fields.
According to the depicted cross sections through the detection region and the demodulation region, both a process and a device simulation have been performed. The results are shown in figure 5.27 (a) and (b). The potential distribution in the buried channel of the detection region confirms the creation of an almost ideal constant lateral drift field. The second cross section in figure 5.27 (b) shows the typical potential distribution of the three-gate demodulation region. Since the toggle gates (pgl and pgr) are switched



(a) Schematic of the pixel's gate arrangement.

(b) Layout.

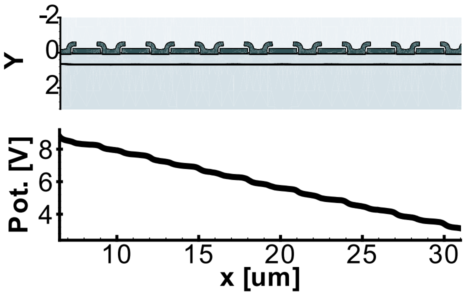
Figure 5.26: Implementation of the static drift field pixel.

between 6 and 8V during operation, the simulation made use of these static voltages (pgm was set to 7V). Figure 5.27 (c) shows a SEM photo of the cross section through the photo-sensitive region, making the different layers visible: silicon substrate, very thin field oxide, overlapping poly-silicon gates, silicon-dioxide, a passivation layer of silicon nitride and the three metal layers.

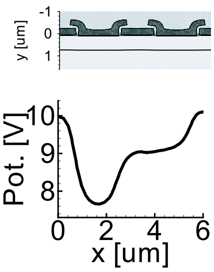
The static drift field pixel was fabricated in a $0.6\mu\mathrm{m}$ CMOS process offering an additional CCD option with buried channel implant and overlapping poly-silicon gates [7]. A fill factor of 25% has been realised with a 2-tap pixel structure of $40 \times 40\mu\mathrm{m}^2$. The pixel was incorporated in a 3D-TOF sensor having a lateral resolution of 176×144 pixel. Two micrographs, one of the sensor and one of a small part of the pixel matrix, are shown in figures 5.28 (a) and (b), respectively.

Performance Estimation

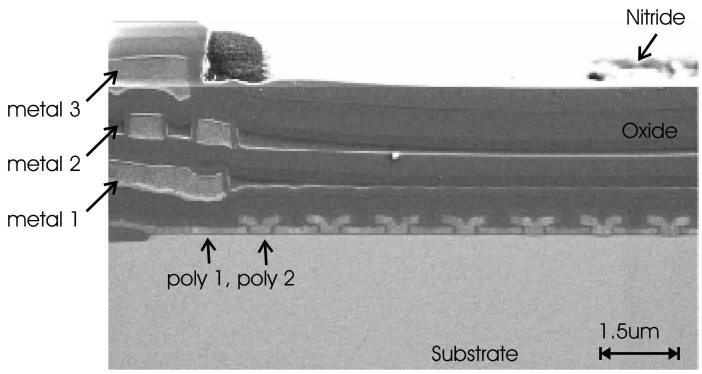
The estimation of the transit time is based on the decomposition of the electron transit path through the detection region into two regions of different drift fields. As sketched in figure 5.29, the upper region has a drift field of $E = V_d/(b_1 + b_2)$ and the bottom region, which is bounded by the tapering shape, creates a drift field at its border of $E * \sin(\gamma)$, where the geometrical relationships between b_1, b_2 and γ can be seen in the figure. Applying equation 5.2 to the static drift field pixel, the maximum drift time



(a) Simulation of channel potential through cross section 1.



(b) Simulation of potential through cross section 2.



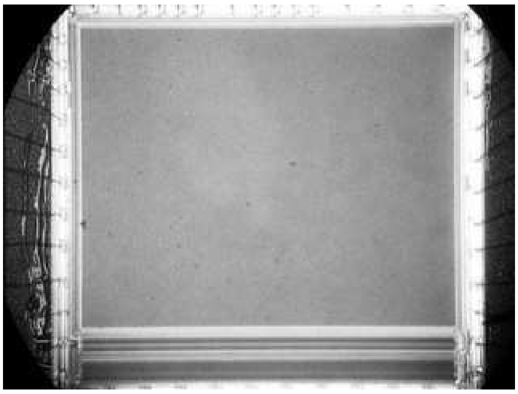
(c) SEM photo of the cross section through the photo-sensitive area of the static drift pixel.

Figure 5.27: Cross sections through the detection and demodulation regions of the static drift field pixel.

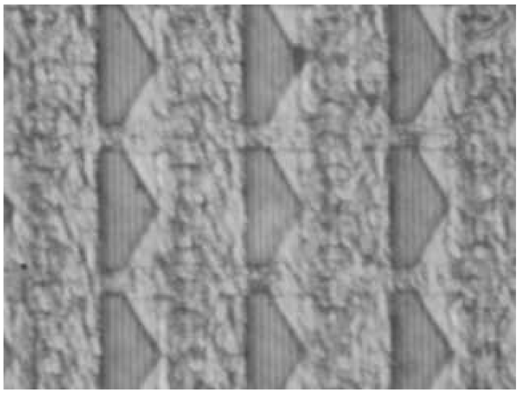
results in the following expression:

$$t_{drift} = \frac{b_1(b_1+b_2)}{\mu_0 V_d} + \frac{b_1}{v_{th}} + \frac{b_2(b_1+b_2)}{\mu_0 V_d \sin^2(\gamma)} + \frac{b_2}{v_{th}\sin(\gamma)}. \tag{5.13}$$

Based on the geometrical pixel parameters of $b_1 = 7.3\mu\text{m}$, $b_2 = 9.6\mu\text{m}$ and $\gamma = 30°$ and a total voltage drop of 7V over the 18 gates of the detection region, the maximum drift time is estimated to be $t_{drift} = 1.17\text{ns}$. In addition to that, the charge transport through the demodulation region adds another delay, which is in the order of approximately $t_{diff} \approx 3\text{ns}$ assuming a charge transportation exclusively by thermal diffusion. Thus, according to equation 5.1, a maximum transit time of approximately 4.17ns is estimated,



(a) Sensor consisting of 176×144 pixels.



(b) Detailed view on the top of a 3×3 sub-array of the pixel matrix.

Figure 5.28: Micrographs of the static drift field sensor.

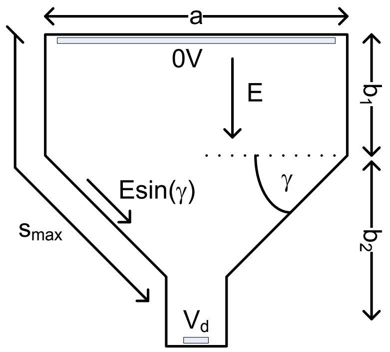


Figure 5.29: Geometrical relations in the detection region.

and, with equation 5.8, a minimum 3dB-cut-off frequency of

$$f_{cut} \geq 105\text{MHz} \tag{5.14}$$

is found.
The following worst-case assumptions made in this assessment are recalled: 1. The longest transition path has been selected for the assessment, though the largest part of the photo-generated electrons is generated inside the detection area, experiencing a shorter transport path. 2. The transport through the demodulation region is strongly influenced by the fringing fields, extending widely below the neighbouring gates and, hence, accelerating the charge transport even more.

Design Aspects - Advantages and Disadvantages

In the following, some important design aspects are summarized exhibiting the advantages and disadvantages of the new demodulation pixel.

Gate capacitances: The gate capacitances that have to be charged and discharged with the speed of the modulation frequency are minimized thanks to the small sizes. Compared to standard lock-in pixels of similar size and fill-factor, a factor of 10 or more can theoretically be saved. This means lower power consumption of the driver electronics or higher modulation frequencies can be achieved. In practice a reduction by a factor of about 4 was measured, which is less than the theoretical value due to the impact of parasitic capacitances.

Transit time: The fast charge transport through the detection region by lateral electric drift fields and furthermore, the short paths through the demodulation region, enable the utilization of the pixel in high-speed applications leading to increased distance accuracies.

Optical fill factor: Due to the decoupling of the high-speed demodulation site from the detection site, the photon-collection region can be designed with a large area without compromising the demodulation frequency. This enables the design of time-of-flight pixels that have high optical fill factors and support high modulation frequencies at the same time.

Microlenses: The geometry of the pixel is more appropriate for the utilization of micro lenses than a stripe-like structure as present in a standard lock-in pixel having three adjacent photo-gates [23].

Noise: The pixel operates entirely in the charge domain, where the accumulation of photo-charge is virtually noise-free. Therefore the pixel is expected to exhibit the same noise-limitations as standard CCDs.

Low power consumption: The application of linearly increasing but temporally constant voltages to the successive arrangement of the gates in the detection region ensures low power consumption in this non dissipative arrangement.

Process options: In comparison to the drift field pixel presented in section 5.3.1, which necessitates the deposition of high-resistivity poly-silicon material for creating the photogate, the requirements of the process in the case of the static drift field pixel are relaxed. No unconventional high-resistivity material is required. The ladder of consecutively arranged resistors for the generation of the constant voltages makes use of standard integrated process components. Thus, the same process options as for the standard lock-in pixel are needed, meaning a buried channel implant and overlapping poly-silicon gates. However, this pixel will strongly benefit from the ongoing reduction of the process feature size: 1. Narrower gates allow for approximating the discretized potential decrease in the detection region to the ideal linear behaviour. 2. The requirement of overlapping gates will disappear as soon as the gaps between two adjacent gates no longer have a significant impact on the potential distribution in the semiconductor. The problem of gap spacing in CCDs is addressed for example in [49, 114, 115].

Pixel complexity: The density of the design has been increased compared to the standard lock-in pixel.

Enhancement to multi-tap structures: Since the demodulation region has to be designed to be as small as possible, the enhancement to real multi-tap structures is difficult to accomplish. The wiring and mismatch between the several channels are important issues in this context.

Experimental Results

The practical verification of the static drift field concept has been accomplished in two steps. First, evidence of high speed charge transport through the detection region thanks to lateral drift processes is obtained. Secondly, the demodulation capability is verified, which allows the utilisation of the pixel in a real-time 3D time-of-flight imaging system.

Figure 5.30 sketches the measurement set-up used for the verification of the effectiveness of the drift field. A small region of the detection region is illuminated with a small-sized temporally modulated laser spot first at the top side and then at the bottom

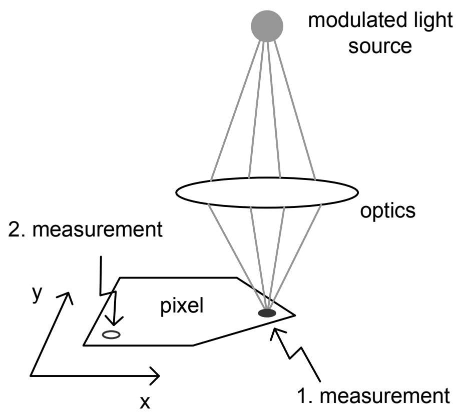


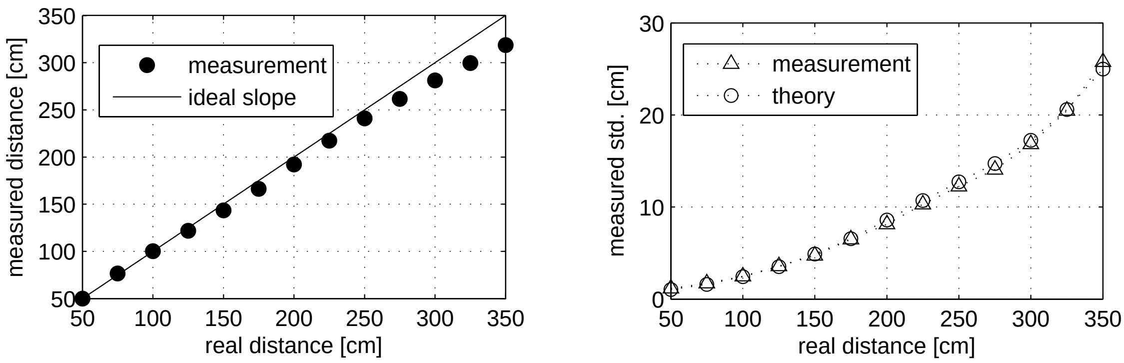
Figure 5.30: Simplified illustration of the microscope set-up used for the measurement of the electron transit time through the detection region.

side. By demodulating the photo-signal, the difference of the two phase measurements gives an indication of the drift time. We measured time differences of about

$$t_{drift} \approx 0.5\text{ns}.$$

It is worth noting that the measured time is shorter than the theoretically estimated one of about 1.17ns. Two reasons explain this behaviour: 1. Self-induced drift fields are completely neglected in the theoretical model. However, those fields become effective if the electron generation rate is high, which is the case in the measurement set-up used. 2. The theoretical model provides a worst-case estimation assuming that the electrons take the longest path and the electric fields are reduced by a factor of $\sin(\gamma)$ in the tapering region. Both are obviously not the case in practice. In particular, the optical set-up does not allow the perfect adjustment of the laser spot in the right top corner of the detection region.
As a consequence of the measurement, it is concluded that the lateral charge transport is very effectively accomplished by electric drift fields.

The results of the distance measurement, both the linearity measurement as well as the accuracy measurement, are shown in figure 5.31, where a 20MHz sinusoidal modulation signal has been used. The measured distance and accuracy are depicted versus the real distance of an object with a reflectivity of approximately 10%.
The linearity measurement in figure 5.31 (a) confirms the suitability of the pixel for the



(a) Linearity: measured distance versus real distance.

(b) Accuracy: measured standard deviation versus real distance. Comparison with photon shot noise limited measurement precision.

Figure 5.31: Distance measurement with the static drift field sensor at 20MHz modulation frequency.

demodulation of modulated light signals. Slight deviations of the measured distance values from the ideal slope are mainly due to harmonic distortions of the emitted optical signal.
The corresponding standard deviation in figure 5.31 (b) shows a quadratic behaviour as a function of the distance, implying the dominance of the measurement by the noise sources of the external electronics as a result of a less than ideal measurement set-up. Based on a measured demodulation contrast of more than 40%, the theoretical values, computed with equation 2.30, are depicted in the figure as well, demonstrating the good agreement between practice and theoretical system model. Moreover, it is worth mentioning that millimeter resolution is obtained with this pixel structure under strong optical power conditions.

To demonstrate the functionality of the full sensor with its 176×144 pixels, a 3D-image of a human hand in front of a wall is shown in figure 5.32. The distance data is gray coded according to the chart on the right side of the figure.

Figure 5.32: 3D-image taken with the TOF sensor based on the static drift field pixel.

5.4 Summary

The demodulation pixels have experienced a continuous improvement during the last few years. First multi-tap structures showed important drawbacks with regard to fill-factor and demodulation speed. The pixel structures recently implemented in 3D-TOF cameras are based on a 2-tap architecture that provide the best compromise between fill-factor, demodulation speed and number of taps. The pixels enable recent cameras to be used in a wide range of applications. Depth resolutions in the millimeter range are achieved for a wide range of optical power conditions. However, the steady increase in the demands for high-precision measurements even below millimeter range and for the utilisation of 3D-cameras in arbitrary environments with strong optical conditions, makes the development of new pixel concepts became necessary.

The enhanced pixel structures that the are subject of this research project, were shown to improve the flexibility and applicability of the 3D-TOF camera system. Pixels with individually controlled integration time and minimum charge transfer make it possible, for example, to use 3D-TOF cameras in outdoor applications. Furthermore, for the first time demodulation pixels based on fast lateral charge transport due to electric drift fields were realised. Two concepts of drift field pixels, using either static or modulated drift fields, have been verified by measuring distances. While the static drift field structure exhibits important advantages in power consumption, high-speed and constant charge transport through the whole detection region, the modulated drift field pixel has a particular advantage in its design simplicity. The achievable sub-nanosecond transit times of electrons should make it possible to implement GHz demodulation pixels, once electronic driver circuits are available with the required performance. This opens the way to optical 3D TOF measurements with a precision below $100\mu\mathrm{m}$.

Chapter 6

Summary and Perspective

The first chapter of the present work summarises the system-related design-aspects of a 3D-TOF camera concerning the power budget, the requirements for the dynamic range of the imager and the limitations of the sinusoidal modulation scheme. Based on those considerations, the need for new optical modulation techniques and enhanced pixel concepts becomes obvious in order to enable the 3D-TOF to operate under various demanding environmental conditions. The main aspects, which are all covered in this work, are the following:

- multi-camera operation without intra-camera interference,
- the handling of high dynamic range scenes with regard to both signal- and background-light,
- high-speed demodulation to achieve precise distance measurements in the sub-millimeter range.

We have realised the capability of multi-camera operation with an optical intensity-modulation technique that is based on maximal-length pseudo-noise sequences. Extensive simulations for predicting the performance in a multi-camera environment are accomplished and theoretical analysis for single-camera operation are given. The analysis covers the achievable stochastic accuracy, which is determined by photon shot noise, non-linearities of the output channels, multi-path propagation in the scene or multiple reflections in the optics and the effect of low-pass filtered modulation signals. We have checked the correctness of our theoretical models by performing real measurements with state-of-the-art 3D-TOF cameras. Measurements with up to five cameras have shown that distance variations of less than $10\mathrm{cm}$ can be expected, which is in good agreement with theory. This result means an important advancement of the

TOF-technology. The sinusoidal approach would be completely inappropriate when five cameras are concurrently running with the same frequencies. However, the single-camera performance with pseudo-noise modulation suffers, which necessitates a careful consideration of the operation conditions when deciding which modulation technique to use.
While the modulation with maximal-length sequences demonstrates the possibility of successfully suppressing the optical signals of concurrently running cameras, the reduced single-camera performance makes future research necessary, where the focus lies on even more efficient modulation signals. Combinations between m-sequences and sinusoidal signals, new constructions of binary bit sequences that show orthogonal properties and enable a phase measurement or heterodyne phase measurement with chirp signals could be topics for future exploration.

The present research results will form an important contribution to 3D-TOF cameras in the future, since many new applications will arise where several cameras are concurrently imaging the same scene and no synchronisation between them is possible. This includes primarily applications in the field of robotics, autonomously driving vehicles and gaming. There, the research can profit from the enhancement of the pixel structures when pixels with larger numbers of taps are efficiently realisable. Thus, the theory of modulation and the hardware aspects have to be adjusted to each other and they can profit from each other.

We have presented two new pixel concepts that enhance the 3D-TOF range imaging to much larger dynamic ranges than previously possible. The concepts use additional electronic circuits, which are implemented either pixel- or column-wise. However, in either case the pixels always operate in the almost noisefree charge domain, allowing distances to be measured with high precision, only limited by photon shot noise.
One pixel structure extends the signal range significantly by controlling the integration time of each pixel individually. A second pixel structure handles background light up to 1.3-times that of sunlight when certain system aspects such as an appropriate bandpass filter and integration time are considered.

For the first time real outdoor applications are realisable with lock-in pixels that are based on a combined CCD/CMOS process technology. In order to address even more application fields with this technology, the combination of both pixel concepts would be desirable but still suffers from low fill-factors. Future decrease of the process feature sizes will allow the efficient combination of these approaches.

The high-speed demodulation problem was successfully solved by development of two novel concepts of demodulation pixels, both based on lateral electric drift fields. The

first pixel uses a one-gate structure of very high-resistivity poly-silicon material. The dynamic control of the current flowing through the gate itself enables the modulation of the lateral electric fields in the semiconductor bulk. The principle of fast charge transportation based on electric drift fields was demonstrated with large-area pixels having a $25 \times 25 \mu\mathrm{m}^2$ quadratic shape. Distance measurements at $20\mathrm{MHz}$ modulation proved the functionality of the pixels, though higher modulation frequencies could not be tested due to driving problems with the external electronics. The performance of standard pixels with similar size would have been insufficient already at the demonstrated modulation speed, so that the gain of the novel pixel is obvious. The main drawbacks of this pixel structure, inhibiting the utilisation of the pixel in dense matrix sensors, are the in-pixel power consumption and the large gate-capacitances that have to be switched at the modulation frequency.
The second demodulation pixel based on lateral electric drift fields is even more promising because it overcomes the afore-mentioned disadvantages of the first drift field pixel. A static drift field is generated by a suitable arrangement of CCD elements avoiding any additional power consumption. The exploitation of static drift fields within a sufficiently large detection region enables fast charge transportation, while the small demodulation region ensures small switching capacitances, reducing the requirements of the driving electronics tremendously. Distance measurements with a 176×144 pixel sensor verify the concept of the static drift field pixel. We measured signal delays, due to the finite transit time of electrons through the detection region, to be in the sub-nanosecond range, showing good agreement with theory.

The drift field pixels represent a new generation of demodulation pixels, to which the basic theory of 3D-TOF range imaging (sampling and correlation, stochastic accuracy and dynamic range) can be applied in the same way as for the standard CCD lock-in pixels. This in turn allows the use of these pixels in combination with the concepts for increased high-dynamic range and background light suppression. Thus, the drift field pixels are suggested to be the basic pixel structures for future 3D-TOF cameras, enabling high-precision range imaging in various fields of applications.

Future process-related research will emphasise the development of the drift field pixels, aiming at transferring the technology to standard CMOS processes of small feature size. This reduces process costs, allows for implementing the pixels in various CMOS processes and enables high fill-factors.
The static drift field pixel will be the preferred pixel structure to be realised in a standard CMOS process because it ensures high-speed demodulation and it benefits from smaller feature sizes in the detection region. Smaller gate widths enable the creation of even smoother drift fields. However, the transfer to standard CMOS processes with

small feature sizes will not be straightforward because the changes of the doping concentrations of the bulk and the drain/source diffusions have a direct impact on the amount of stored charge and, hence, on the achievable measurement accuracy.

All these considerations show that optical 3D-TOF range imaging has been enhanced in the course of this thesis work to various extreme operating conditions concerning multi-camera environments, high-dynamic range scenes, extreme background light conditions or demanding high-precision applications. However, the universal real-time range camera that deals with all problems at the same level of perfection is still missing. Some of the solutions presented are just intermediate steps and they still need further improvement, or alternative concepts may be required.

To summarise, we have shown that optical time-of-flight 3D image sensing is a very promising imaging modality, capable of acquiring real-time imagery of our environment in its full three-dimensional richness. The state-of-the-art in 3D-TOF severely limits its application to situations with extreme illumination conditions, multi-camera operation and where sub-millimeter precision is required. In the framework of the present thesis all three problems have been investigated thoroughly, new theoretical approaches have been proposed, these concepts have been realised technically and their predicted performance has been demonstrated successfully in practice. In this way, the results of this thesis contribute relevant elements to optical 3D TOF imaging, making this exciting technology even more widely applicable and opening new and more demanding domains of application for it.

Appendix A

Signal Theory

A.1 Summation of Two Sinusoids

The summation of two sinusoids having the same frequency but different phases φ_1 and φ_2 and different amplitudes P_{A1} and P_{A2}, respectively, is depicted in figure A.1. The amplitude of phasor two is smaller than that of phasor one. The resulting phasor has an amplitude P_Σ and a phase φ_Σ. The phase difference φ_Δ describes the phase measurement error due to the superposition of the two sinusoids. Its general form is derived in the following and the maximum phase error is calculated.

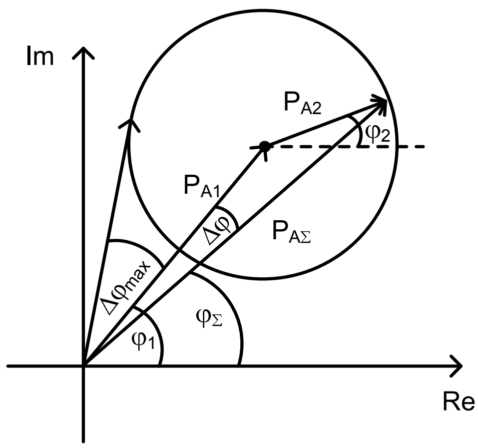


Figure A.1: Phasor diagram depicting the summation of two sinusoids with the same frequency but different phases.

The summation of the two original phasors is expressed as

$$\begin{aligned} P_\Sigma &= P_{A\Sigma} \cdot e^{j\varphi_\Sigma} \\ &= P_{A\Sigma} \cdot \cos(\varphi_\Sigma) + jP_{A\Sigma} \cdot \sin(\varphi_\Sigma) \\ &= P_{A1} \cdot (\cos(\varphi_1) + \tfrac{P_{A2}}{P_{A1}} \cdot \cos(\varphi_2)) + jP_{A1} \cdot (\sin(\varphi_1) + \tfrac{P_{A2}}{P_{A1}} \cdot \sin(\varphi_2)). \end{aligned} \tag{A.1}$$

The correspondences between the real and imaginary parts lead to the two relationships

$$P_{A\Sigma} \cdot \cos(\varphi_\Sigma) = P_{A1} \cdot (\cos(\varphi_1) + \frac{P_{A2}}{P_{A1}} \cos(\varphi_2)) \tag{A.2}$$

$$P_{A\Sigma} \cdot \sin(\varphi_\Sigma) = P_{A1} \cdot (\sin(\varphi_1) + \frac{P_{A2}}{P_{A1}} \sin(\varphi_2)). \tag{A.3}$$

By dividing equations A.2 and A.3, the phase of the resulting phasor φ_Σ is computed as

$$\varphi_\Sigma = \arctan\left[\frac{\sin(\varphi_1) + \frac{P_{A2}}{P_{A1}}\sin(\varphi_2)}{\cos(\varphi_1) + \frac{P_{A2}}{P_{A1}}\cos(\varphi_2)}\right]. \tag{A.4}$$

The phase error $\Delta\varphi$ is

$$\Delta\varphi = \arctan\left[\frac{\sin(\varphi_1) + \frac{P_{A2}}{P_{A1}}\sin(\varphi_2)}{\cos(\varphi_1) + \frac{P_{A2}}{P_{A1}}\cos(\varphi_2)}\right] - \varphi_1. \tag{A.5}$$

By performing the substitution

$$\bar{\varphi_2} = \varphi_2 - \varphi_1, \tag{A.6}$$

the geometrical relations become independent on φ_1, so that φ_1 can be set to 0 for simplification. The last equation is rewritten as

$$\Delta\varphi = \arctan\left[\frac{\frac{P_{A2}}{P_{A1}}\sin(\bar{\varphi_2})}{1 + \frac{P_{A2}}{P_{A1}}\cos(\bar{\varphi_2})}\right]. \tag{A.7}$$

By computing the derivative

$$\frac{\partial\Delta\varphi}{\partial\bar{\varphi_2}} = 0, \tag{A.8}$$

the maximum phase error is found for $\bar{\varphi_2} = \arccos(-P_{A2}/P_{A1})$. Following from that, the maximum phase error $\Delta\varphi_{max}$ is given by

$$\Delta\varphi_{max} = \arctan\left[\frac{\frac{P_{A2}}{P_{A1}}}{\sqrt{1 - \left(\frac{P_{A2}}{P_{A1}}\right)^2}}\right]. \tag{A.9}$$

A.2 Expectation of the Cross-Correlation between m-Sequences

The expectation value of the ratio between the cross-correlation values of m-sequences and the autocorrelation's peak value is defined as

$$\Theta_{avg} := \text{mean}\left\{\frac{\Phi_c(t)}{\Phi_a(0)}\right\}, \tag{A.10}$$

where $\Phi_c(t)$ denotes the cross-correlation between the m-sequences and $\Phi_a(0)$ is the peak value of the autocorrelation function. The cross-correlation function is related to the correlation between one unipolar m-sequence $M_{u,1}(t)$ and another m-sequence $M_{b,2}(t)$ of bipolar nature, so that

$$\Phi_c(t) = (M_{u,1} \otimes M_{b,2})(t), \tag{A.11}$$

where u and b denote unipolar or bipolar, respectively. Based on the fundamental property of the cross-correlation between two bipolar m-sequences [87]

$$\sum_{g=0}^{n_p-1} \Phi_{M_{b,1},M_{b,2}}(g) = +1, \tag{A.12}$$

and the relation between unipolar and bipolar m-sequences

$$M_u = \frac{1}{2}\left(M_b + 1\right), \tag{A.13}$$

the following computation can be performed:

$$\begin{aligned}
\text{mean}\left(\Phi_c\right) &= \tfrac{1}{n_p}\left\{\sum_{g=0}^{n_p-1} \Phi_c(g)\right\} \\
&= \tfrac{1}{n_p}\cdot \sum_{g=0}^{n_p-1} \left[\left\{\tfrac{1}{2}(M_{b,1}(g)+1)\right\} \otimes M_{b,2}(g)\right] \\
&= \tfrac{1}{n_p}\cdot \sum_{g=0}^{n_p-1} \left[\tfrac{1}{2}M_{b,1} \otimes M_{b,2} + \tfrac{1}{2}M_{b,1} \otimes 1\right] \\
&= \tfrac{1}{n_p}\cdot \sum_{g=0}^{n_p-1} \left[\tfrac{1}{2}M_{b,1} \otimes M_{b,2} + \tfrac{1}{2}\right] \\
&= \tfrac{n_p+1}{2n_p}
\end{aligned} \tag{A.14}$$

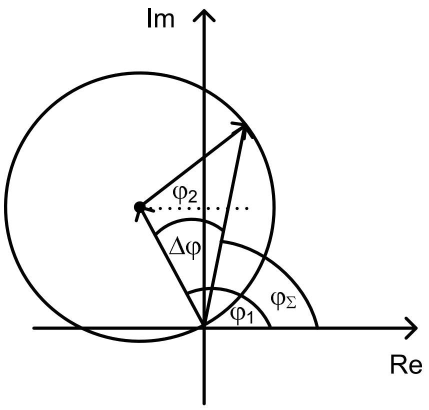
Since the peak of the autocorrelation corresponds to $(n_p+1)/2$, Θ_{avg} can be expressed as

$$\begin{aligned}\Theta_{avg} &= \tfrac{n_p+1}{2n_p} \cdot \tfrac{2}{n_p+1} \\ &= \tfrac{1}{n_p}.\end{aligned} \tag{A.15}$$

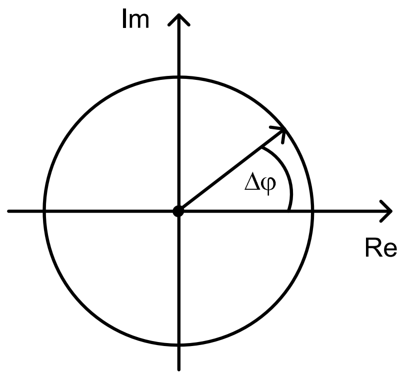
The conversion between discrete and continuous time domains does not have to be inserted into the calculations because all cases are completely covered by the discrete sequences. The continuous-time domain means just an interpolation between the discrete correlation values [87], so that the mean values are not affected at all.

A.3 Distance Error for Sinusoidal Modulation

The addition of two sinusoids, depicted in figure A.2 (a), results in a sinusoid with a phase denoted by φ_Σ. Compared with the original phase φ_1, which is sought by the



(a) Phasor diagram for the addition of two sinusoids of the same frequency and amplitude.



(b) Phasor diagram showing the circle of the resulting phasor. When a large number of cameras N_c is operating simultaneously, the origin of the circle coincides with the origin of the coordinate system.

Figure A.2: Phasor diagram for the addition of two cameras and a very large number of cameras.

distance measurement, the measured phase error is given as

$$\Delta\varphi = \arctan\left[\frac{\sin(\varphi_1) + \frac{P_{A2}}{P_{A1}}\sin(\varphi_2)}{\cos(\varphi_1) + \frac{P_{A2}}{P_{A1}}\cos(\varphi_2)}\right] - \varphi_1, \tag{A.16}$$

where φ_2 describes the phase of the second sinusoid. By substituting $\bar{\varphi_2} = \varphi_2 - \varphi_1$ and assuming the same amplitudes for both sinusoids $P_{A2}/P_{A1} = 1$, the phase error is found

$$\Delta\varphi = \frac{\bar{\varphi_2}}{2}. \tag{A.17}$$

Based on this consideration, the expectation value of the systematic relative distance error and its standard deviation can be estimated. The systematic error is described as

$$\langle\delta\rangle = \frac{1}{2\pi}\langle\Delta\varphi\rangle, \tag{A.18}$$

and the standard deviation is expressed as

$$\sigma_\delta = \sqrt{\frac{1}{4\pi^2}\cdot\langle\Delta\varphi^2\rangle - \langle\Delta\varphi\rangle^2}, \tag{A.19}$$

exploiting the fact that the variance of a variable can be expressed by the expectation value of the square minus the square of the mean value [21].

Two sinusoids: Referring to figure A.2 (a), the expectation values will give

$$\langle\Delta\varphi\rangle = \frac{1}{\pi}\int_{\frac{-\pi}{2}}^{\frac{\pi}{2}} \Delta\varphi d\Delta\varphi = 0, \tag{A.20}$$

and

$$\left\langle\Delta\varphi^2\right\rangle = \frac{1}{\pi}\int_{\frac{-\pi}{2}}^{\frac{\pi}{2}} \Delta\varphi^2 d\Delta\varphi = \frac{\pi^2}{12}. \tag{A.21}$$

for two sinusoids of the same amplitude and with the assumption that the circle does not cover both the first and fourth quadrants. In the last case the expectation values are higher because the modulo-2π produces stronger fluctuations.
For two cameras, the mean error and the standard deviation are

$$\langle\delta\rangle = 0 \tag{A.22}$$

$$\sigma_\delta = \frac{1}{\sqrt{48}} \approx 0.144. \tag{A.23}$$

3 or more cameras: The more cameras that are operating, the more the circle of resulting phasors appears as a circle around the origin of the coordinate system, as depicted in figure A.2 (b). For this extreme case, the expectation values are

$$\langle \Delta\varphi \rangle = \frac{1}{2\pi} \int_0^{2\pi} \Delta\varphi d\Delta\varphi = \pi, \tag{A.24}$$

$$\left\langle \Delta\varphi^2 \right\rangle = \frac{1}{2\pi} \int_0^{2\pi} \Delta\varphi^2 d\Delta\varphi = \frac{4}{3}\pi^2, \tag{A.25}$$

so that the corresponding parameters related to the relative distance, are

$$\langle \delta \rangle = 0.5 \tag{A.26}$$

$$\sigma_\delta = \frac{1}{\sqrt{12}} \approx 0.289. \tag{A.27}$$

A.4 Derivation of the Shot Noise Limited Accuracy for PN Modulation

The distance extraction in a PN modulated TOF measurement is based on the following formula

$$R = \frac{A_2 - A_3}{A_0 - A_1 + A_2 - A_3} \cdot R_{max}, \tag{A.28}$$

where the substitutions according to equation 3.42 have been performed. Based on the rule of Gaussian error propagation, the distance standard deviation is computed as

$$\sigma_R = \sqrt{\sum_{g=0}^{3} \left(\frac{\partial R}{\partial A_g} \right)^2 \mathrm{Var}(\hat{A}_g)}, \tag{A.29}$$

where $\hat{A}_g$ are the noisy sample values. The practical derivatives of R to A_0, A_1, A_2 and A_3, respectively, are described as

$$\frac{\partial R}{\partial A_0} = \frac{A_3 - A_2}{(A_0 - A_1 + A_2 - A_3)^2} \cdot R_{max} \tag{A.30}$$

$$\frac{\partial R}{\partial A_1} = \frac{A_2 - A_3}{(A_0 - A_1 + A_2 - A_3)^2} \cdot R_{max} \tag{A.31}$$

$$\frac{\partial R}{\partial A_2} = \frac{A_0 - A_1}{(A_0 - A_1 + A_2 - A_3)^2} \cdot R_{max} \tag{A.32}$$

$$\frac{\partial R}{\partial A_3} = \frac{A_1 - A_0}{(A_0 - A_1 + A_2 - A_3)^2} \cdot R_{max}. \tag{A.33}$$

By exploiting the fact that the variance of $\hat{A}_g$ is equal to the mean value when Poisson noise is assumed, the variance of the distance measurement can be expressed as

$$\sigma_R^2 = \frac{(A_2 - A_2)^2 \cdot A_0 + (A_2 - A_3)^2 \cdot A_1 + (A_0 - A_1)^2 \cdot A_2 + (A_1 - A_0)^2 \cdot A_3}{(A_0 - A_1 + A_2 - A_3)^4} \cdot R_{max}. \tag{A.34}$$

In the following, the nominator is abbreviated by NO and the denominator by DN. Furthermore, the following relationships are used:

$$A_0 + A_1 = 2A_{sig} + 2BG \tag{A.35}$$

$$A_2 + A_3 = 2A_{sig} + 2BG \tag{A.36}$$

$$A_0 - A_1 = 2c_d A_{sig}\left(1 - \frac{\tau}{T_c}\right) + \frac{2c_d BG}{n_p} \tag{A.37}$$

$$A_2 - A_3 = 2c_d A_{sig}\frac{\tau}{T_c} + \frac{2c_d BG}{n_p}. \tag{A.38}$$

The nominator is

$$\begin{aligned} NO &= (A_0 - A_1)^2 \cdot (A_1 + A_3) + (A_2 - A_3)^2 \cdot (A_0 + A_1) \\ &= 2 \cdot (A_{sig} + BG) \cdot \left[\left\{2c_d A_{sig}\left(1 - \tfrac{\tau}{T_c}\right) + \tfrac{2c_d BG}{n_p}\right\}^2 + \left\{2c_d A_{sig}\tfrac{\tau}{T_c} + \tfrac{2c_d BG}{n_p}\right\}^2\right] \end{aligned} \tag{A.39}$$

and the denominator is calculated as

$$\begin{aligned} DN &= \left[2c_d A_{sig}\left(1 - \tfrac{\tau}{T_c}\right) + \tfrac{2c_d BG}{n_p} + 2c_d BG\tfrac{\tau}{T_c} + \tfrac{2c_d BG}{n_p}\right]^4 \\ &= \left[2c_d BG + \tfrac{4c_d BG}{n_p}\right]^4. \end{aligned} \tag{A.40}$$

By simplifying the quotient of nominator and denominator and taking the square root, the standard deviation of the distance measurement is obtained

$$\sigma_R = \frac{cT_c}{2\sqrt{2}}\frac{\sqrt{A_{sig} + BG}}{c_d(A_{sig} + \frac{2}{n_p}BG)^2} \cdot \sqrt{\left(A_{sig}(1 - \frac{\tau}{T_c}) + \frac{BG}{n_p}\right)^2 + \left(A_{sig}\frac{\tau}{T_c} + \frac{BG}{n_p}\right)^2}. \tag{A.41}$$

The left square root expression includes the noise components. The dark current shot noise can be added by a constant BD, when an equally distributed dark current is assumed without being influenced by the one-bit unbalancing property of the m-sequence. Furthermore, additional noise components are modelled by a general parameter N_{ps}, so that the complete model of distance resolution results in

$$\sigma_R = \frac{cT_c}{2\sqrt{2}}\frac{\sqrt{A_{sig} + BG + BD + N_{ps}}}{c_d(A_{sig} + \frac{2}{n_p}BG)^2} \cdot \sqrt{\left(A_{sig}(1 - \frac{\tau}{T_c}) + \frac{BG}{n_p}\right)^2 + \left(A_{sig}\frac{\tau}{T_c} + \frac{BG}{n_p}\right)^2}. \tag{A.42}$$

Appendix B

Physical Constants

Boltzmann's constant	$K_B = 1.3806505 \cdot 10^{-23}\mathrm{J/K}$
Charge mobility in the linear region of silicon	$\mu_0 = 1450\mathrm{cm}^2/\mathrm{Vs}$ at $\vartheta = 300\mathrm{K}$
Dielectric constant of silicon	$\epsilon_{si} = 11.9$
Dielectric constant of silicon-oxide	$\epsilon_{ox} = 3.9$
Effective mass of electrons in silicon	$m_{eff} = 0.2368444 \cdot 10^{-30}\mathrm{kg}$
Elementary charge	$q = 1.60217653 \cdot 10^{-19}\mathrm{C}$
Permittivity in vacuum	$\epsilon_0 = 8.85418 \cdot 10^{-14}\mathrm{F/cm}$
Planck's constant	$h = 6.6260693 \cdot 10^{-34}\mathrm{Js}$
Speed of light in vacuum	$c = 2.99792 \cdot 10^{8}\mathrm{m/s}$
Speed of ultrasound in air	$c_{us} \approx 343\mathrm{m/s}$ at $\vartheta = 293\mathrm{K}$

Bibliography

[1] American Society for Testing and Materials, http://www.astm.org.

[2] Database of complex refraction indices for numerous materials, http://www.sopra-sa.com/.

[3] Leica Geosystems: Distanzmessprinzip, http://www.leica-geosystems.com/media/new/product_solution/Messprinzip_de.pdf.

[4] Reference Solar Spectral Irradiance, Data Set, http://rredc.nrel.gov/solar/spectra/am1.5/.

[5] Riegl Laser Measurement Systems, http://www.riegl.co.at.

[6] Surphaser, http://www.surphaser.com.

[7] ZFOUNDRY Dresden, http://www.zfoundry.com.

[8] Zoller + Fröhlich GmbH, http://www.zf-laser.com/d_punktuelle.html.

[9] M. A. Albota, R. M. Heinrichs, D. G. Kocher, D. G. Fouche, B. E. Player, M. E. Obrien, G. F. Aull, J. J. Zayhowski, J. Mooney, B. C. Willard, and R. R. Carlson. Three-dimensional imaging laser radar with a photon-counting avalanche photodiode array and microchip laser. *Applied Optics*, 41(36):7671–7678, 2002.

[10] G.G. Amelio, M.F. Tompsett, and G.E. Smith. Experimental Verification of the Charge Coupled Device Concept. *The Bell system Technical Journal*, April 1970.

[11] M. Bass, E.W. Van Stryland, D.R. Williams, and W.L. Wolfe, editors. *Handbook of Optics*, volume 1, chapter Optical Properties of Films and Coatings. McGraw-Hill, Inc., 2. edition, 1995.

[12] S. Beer and P. Seitz. Real-Time Tomographic Imaging Without X-Rays: A Smart Pixel Array With Massively Parallel Signal Processing For Real-Time Optical Coherence Tomography Performing Close To The Physical Limits. In *IEEE Proc. of the Conference, 2005 PhD Research in Microelectronics and Electronics*, pages 335–338, 2005.

[13] S. Beer, S. Waldis, and P. Seitz. Video-rate optical coherence tomography imaging with smart pixels. In *Optical Coherence Tomography and Coherence Techniques. Edited by Drexler, Wolfgang. Proceedings of the SPIE, Volume 5140, pp. 69-76 (2003).*, pages 69–76, October 2003.

[14] S. Beer, P. Zeller, N. Blanc, F. Lustenberger, and P. Seitz. Smart pixels for real-time optical coherence tomography. In *Three-Dimensional Image Capture and Applications VI. Edited by Corner, Brian D.; Li, Peng; Pargas, Roy P. Proceedings of the SPIE, Volume 5302.*, pages 21–32, April 2004.

[15] J.D.E. Beynon and D.R. Lamb. *Charge-coupled devices and their applications*. McGraw-Hill Book Company, London, 1980.

[16] Alice Irene Biber. *Avalanche Photodiode Image Sensing in Standard Silicon BiCMOS Technology*. PhD thesis, ETH Zurich (Switzerland), Zurich, 2000.

[17] F. Blais. Review of 20 years of range sensor development. *Journal of Electronic Imaging*, 13(1):231–243, January 2004.

[18] Nicolas Blanc, Thierry Oggier, Gabriel Gruener, Jan Weingarten, Alain Codourey, and Peter Seitz. Miniaturized smart cameras for 3D-imaging in real-time. *Proceedings of IEEE*, 1:471–474, 2004.

[19] Stéphane Bourquin. *Low-Coherence Interferometry Based On Customized Detector Arrays*. PhD thesis, EPFL Lausanne, 2000.

[20] W.S. Boyle and G.E. Smith. Charge Coupled Semiconductor Devices. *The Bell System Technical Journal*, pages 587–593, April 1970.

[21] I.N. Bronstein, K.A. Semendjajew, G. Musiol, and H. Mülig. *Taschenbuch der Mathematik*. Verlag Harri Deutsch, Thun and Frankfurt am Main, 3. edition, 1997.

[22] B. Büttgen, T. Oggier, R. Kaufmann, P. Seitz, and N. Blanc. Demonstration of a novel drift field pixel structure for the demodulation of modulated light waves with application in three-dimensional image capture. In *Three-Dimensional Image Capture and Applications VI. Edited by Corner, Brian D.; Li, Peng; Pargas, Roy P. Proceedings of the SPIE*, volume 5302, pages 9–20, April 2004.

[23] B. Büttgen, T. Oggier, M. Lehmann, R. Kaufmann, and F. Lustenberger. CCD/CMOS Lock-In Pixel for Range Imaging: Challenges, Limitations and State-of-the-Art. In *1st Range Imaging Research Day*, pages 21–32. Swiss Centre for Electronics and Microtechnology (CSEM), Eidgenössische Technische Hochschule Zurich, September 2005.

[24] B. Büttgen, T. Oggier, M. Lehmann, R. Kaufmann, S. Neukom, M. Richter, M. Schweizer, D. Beyeler, R. Cook, C. Gimkiewicz, C. Urban, P. Metzler, P. Seitz, and F. Lustenberger. High-speed and high-sensitive demodulation pixel for 3D imaging. In *Three-Dimensional Image Capture and Applications VII. Proceedings of the SPIE*, volume 6056, pages 22–33, February 2006.

[25] Bernhard Büttgen. Untersuchung und Aufbau eines neuartigen Laserradars auf MSM-PMD Basis. Diploma thesis, University of Siegen, 2002.

[26] B. Buxbaum, R. Schwarte, T. Ringbeck, M. Grothof, and X. Luan. MSM-PMD as correlation receiver in a new 3D-ranging system. In U. Schreiber, C. Werner, G. W. Kamerman, and U. N. Singh, editors, *Proc. SPIE Vol. 4546, p. 145-153, Laser Radar: Ranging and Atmospheric Lidar Techniques III, Ulrich Schreiber; Christian Werner; Gary W. Kamerman; Upendra N. Singh; Eds.*, pages 145–153, January 2002.

[27] Bernd Buxbaum. *Optische Laufzeitentfernungsmessung und CDMA auf Basis der PMD-Technologie mittels phasenvariabler PN-Modulation*. PhD thesis, University of Siegen (Germany), Siegen, 2001.

[28] J.E. Carnes, W.F. Kosonocky, and E.G. Ramberg. Free Charge Transfer in Charge-Coupled Devices. *IEEE Transactions on Electron Devices*, 19(6):798–808, June 1972.

[29] E. Charbon. Techniques for CMOS Single Photon Imaging and Processing. Shanghai, 2005. ASICON.

[30] Q. Chen, D. Zhao, C. Yang, and Y. Huo. Self-triggering pulsed time-of-flight laser range-finding method. *Optical Engineering*, 42(12):3608–3611, December 2003.

[31] S. Cova, M. Ghioni, A. Lacaita, C. Samori, and F. Zappa. Avalanche photodiodes and quenching circuits for single-photon detection. *Applied Optics*, 35(12):1956–1975, April 1996.

[32] K. David and T. Benkner. *Digitale Mobilfunksysteme*. Teubner, Stuttgart, 1996.

[33] R.C. Dixon. *Spread Spectrum Techniques*, chapter Introduction, pages 1–14. IEEE Press, New York, 1976.

[34] P.G. Drennan and C.C. McAndrew. Understanding MOSFET Mismatch for Analog Design. *IEEE Journal of Solid-State Circuits*, 38(3):450–456, March 2003.

[35] S.E. El-Khamy, S.E. Shaaban, and E.A. Thabet. Efficient Multiple-Access Communications using Multi-User Chirp Modulation Signals. In *Spread Spectrum Techniques and Applications Proceedings*, volume 3, pages 1209–1213, Mainz, 1996. IEEE.

[36] M'Hamed-Ali El Mechat. Examination of modern modulation schemes for future 3D-TOF cameras. Master's thesis, EPFL Lausanne, 2006.

[37] O. Elkhalili, O.M. Schrey, P. Mengel, M. Petermann, W. Brockherde, and B.J. Hosticka. A 4x64 Pixel CMOS Image Sensor for 3-D Measurement Applications. *IEEE Journal of Solid-State Circuits*, 2004.

[38] L.J.M. Esser and A.J.P. Theuwissen. *Handbook on Semiconductor*, volume 4, chapter Charge Coupled Devices: Physics, Technology and Imaging. North-Holland, Amsterdam, 1993.

[39] T. Ferris. Design of two tone ultrasonic distance measurement system. *Applications of Electromagnetic Phenomena in Electrical and Mechanical Systems*, 2000.

[40] E.R. Fossum. Active pixel sensors: are CCDs dinosaurs? In *Proc. of SPIE, Charge-Coupled Devices and Solid State Optical Sensors III*, pages 2–14, July 1993.

[41] E.C. Fox, Nixon O, M.S. Agwani, D.R. Dykaar, T.J. Mantell, and R.W. Sabila. A high speed linear CCD sensor with pinned photo-diode photosite for low lag and low noise imaging. In *Proc. of SPIE*, volume 3301, page 17, 1998.

[42] H.G. Geib, I. Koren, U. Ramacher, B. Schilling, and H. Schwendner. Generationenwechsel - CMOS statt CCD. *Elektronik*, 2000.

[43] P. Geissler, T. Dierig, and H.A. Mallot. *Handbook of Computer Vision and Applications*, chapter Three-Dimensional Imaging Algorithms. Academic Press, April 1999.

[44] J.G. Gottling. Node and Mesh Analysis by Inspection. *IEEE Transactions on Education*, 38(4), 1995.

[45] J. Gühring. Dense 3-D surface acquisition by structured light using off-the-shelf components. In *Proc. of SPIE, Videometrics VII*, volume 4309, 2001.

[46] P. Hammond. On resolution, accuracy and calibration of digital ultrasonic thickness gauges. *The online Jounal of Nondestructive Testing and Ultrasonics*, 2(10), 1997.

[47] S. Hengstler, Kasilingam D.P., and A.H. Costa. A novel chirp modulation spread spectrum technique for multiple access. In *2002 IEEE Seventh International Symposium on Spread Spectrum Techniques and Applications*, volume 1, pages 73–77, Santa Clara, 2002. IEEE.

[48] K. Hoffmann. Surface Charge Transport with an MOS-Transmission-Line. *Solid-State Electronics*, 20:177–181, 1977.

[49] C.R. Hoople and J.P. Krusius. Characteristics of Submicrometer Gaps in Buried-Channel CCD Structures. *IEEE Transactions on Electron Devices*, 38(5), May 1991.

[50] New Wave Instruments, M-Sequences, LFSR, Feedback Taps, www.newwaveinstruments.com/resources/articles/ m_sequence_linear_feedback_shift_register_lfsr.htm.

[51] C. Jacoboni, C. Canali, G. Ottaviani, and A. Alberigi Quaranta. A review of some charge transport properties of silicon. *Solid State Electronics*, 20:77–89, February 1977.

[52] J.R. Janesick. *Charge-Coupled Devices*. SPIE - International Society for Optical Engineering, 1987.

[53] Y. Ju and B. Barkat. A new efficient modulation technique for multi-user access communications systems. In *IEEE International Conference on Acoustics, Speech, and Signal Processing, 2004. Proceedings. (ICASSP'04)*, volume 4, pages 937–40. IEEE, 2004.

[54] I. Kaisto, J. Kostamovaara, M. Manninen, and R. Myllyla. Optical range finder for 1.5-10-m distances. *Applied Optics*, 22(20):3258–3264, October 1983.

[55] R. Kaufmann, M. Lehmann, M. Schweizer, M. Richter, P. Metzler, G. Lang, T. Oggier, N. Blanc, P. Seitz, G. Gruener, and U. Zbinden. A time-of-flight line sensor: development and application. In B. Culshaw, A. G. Mignani, and R. Riesenberg, editors, *Optical Micro- and Nanometrology in Manufacturing Technology. Edited by Gorecki, Christophe; Asundi, Anand K. Proceedings of the SPIE*, volume 5459, pages 192–199, September 2004.

[56] Rüdiger Klein. *Ein laseroptisches Entfernungsmessverfahren mit frequenzvariabler Pseudo-Noise-Modulation*. PhD thesis, University of Siegen, Siegen, October 1993.

[57] S. Kleinfelder, S.H. Lim, X. Liu, and A. El Gamal. A 10000 Frames/s CMOS Digital Pixel Sensor. *IEEE Journal of Solid-State Circuits*, 36(12):2049–2059, December 2001.

[58] P. Knoll. Driver assistance systems - driver support or driver tutelage? In *VDE Micro.tec 2003, Conf. Proc.*, volume 1, pages 341–346, 2004.

[59] R. Lange and P. Seitz. Solid-state time-of-flight range camera. *IEEE Journal of Quantum Electronics*, 37(3):390–397, March 2001.

[60] Robert Lange. *3D Time-of-flight distance measurement with custom solid-state image sensors in CMOS/CCD-technology*. PhD thesis, University of Siegen (Germany), Siegen, 2000.

[61] M. Lehmann, L. Cavalier, S. Lauxtermann, M. Richter, F. Lustenberger, N. Blanc, and P. Seitz. A 2x2k photo line sensor for high-speed applications. *Optics and Lasers in Engineering*, 43:349–364, 2005.

[62] M. Lehmann, T. Oggier, B. Büttgen, C. Gimkiewicz, M. Schweizer, R. Kaufmann, F. Lustenberger, and N. Blanc. Smart pixels for future 3D-TOF sensors. In *IEEE Workshop on CCD and Advanced Image Sensors*, Nagano, June 2005.

[63] D.R. Lide, editor. *Handbook of Chemistry and Physics*. CRC Press, 82. edition, 2001-2002.

[64] D. Litwiller. CMOS vs. CCD: Maturing Technologies, Maturing Markets. *Photonics Spectra*, pages 54–60, August 2005.

[65] T. Lulé, S. Benthien, H. Keller, F. Mütze, P. Rieve, K. Seibel, M. Sommer, and M. Böhm. Sensitivity of CMOS Based Imagers and Scaling Perspectives. *IEEE Transactions on Electron Devices*, 47(11):2110–2122, November 2000.

[66] Rioux M. Laser range finder based on synchronized scanners. *Applied Optics*, 23(21):3837–3844, November 1984.

[67] J. May. *The Greenpeace Book of Dolphins*. Sterling Pub. Co. Inc., 1990.

[68] G. Meynants, B. Dierickx, D. Uwaerts, and J. Boggaerts. Fixed pattern noise suppression by a differential readout chain for a radiation-tolerant image sensor. In *2001 IEEE Workshop on CCDs and AISs*, pages 56–59, 2001.

[69] M. Mohsen, M.F. Tompsett, and C.H. Séquin. Noise Measurements in Charge-Coupled Devices. *IEEE Transactions on Electron Devices*, ED-22(5):209–218, May 1975.

[70] R.N. Mutagi. Pseudo noise sequences for engineers. *Electronics & Communication Engineering Journal*, April 1996.

[71] C. Niclass and E. Charbon. A Single Photon Detector Array with 64x64 Resolution and Millimetric Depth Accuracy for 3D Imaging. In *2005 IEEE International Solid-State Circuits Conference, ISSCC, Digest of Technical Papers*, February 2005.

[72] C. Niclass, A. Rochas, P.-A. Besse, and E. Charbon. Design and Characterization of a CMOS 3-D Image Sensor Based on Single Photon Avalanche Diodes. *IEEE Journal of Solid-State Circuits*, 40(9), September 2005.

[73] C.L. Niclass, A. Rochas, P.-A. Besse, and E. Charbon. A CMOS Single Photon Avalanche Diode Array for 3D Imaging. In *2004 IEEE International Solid-State Circuits Conference, ISSCC, Digest of Technical Papers*, February 2004.

[74] K. Oberhauser, A. Nemecek, and H. Zimmermann. Monolithische Korrelationsempfänger für optische Abstandsmessung. *Technisches Messen*, 72:230–235, 2005.

[75] T. Oggier, B. Büttgen, F. Lustenberger, R. Becker, B. Rüegg, and A. Hodac. SwissRanger SR3000 and First Experiences based on Miniaturized 3D-TOF Cameras. In *1st Range Imaging Research Day*, pages 97–108. Swiss Centre for Electronics and Microtechnology (CSEM), Eidgenössische Technische Hochschule Zurich, September 2005.

[76] T. Oggier, R. Kaufmann, M. Lehmann, B. Büttgen, S. Neukom, M. Richter, M. Schweizer, P. Metzler, F. Lustenberger, and N. Blanc. Novel pixel architecture with inherent background suppression for 3D time-of-flight imaging. In *Proceedings of the SPIE*, volume 5665, pages 1–8, December 2004.

[77] T. Oggier, R. Kaufmann, M. Lehmann, P. Metzler, G. Lang, M. Schweizer, M. Richter, B. Büttgen, N. Blanc, K. Griesbach, B. Uhlmann, K.-H. Stegemann, and C. Ellmers. 3D-Imaging in Real-Time with Miniaturized Optical Range Camera. In *Proc. OPTO conference*, Nürnberg, 2004.

[78] T. Oggier, M. Lehmann, R. Kaufmann, M. Schweizer, M. Richter, P. Metzler, G. Lang, F. Lustenberger, and N. Blanc. An all-solid-state optical range camera for 3D real-time imaging with sub-centimeter depth resolution (SwissRanger). In *Optical Design and Engineering. Edited by Mazuray, Laurent; Rogers, Philip J.; Wartmann, Rolf. Proceedings of the SPIE, Volume 5249, pp. 534-545 (2004).*, pages 534–545, February 2004.

[79] J.-R. Ohm and H.D. Lüke. *Signalübertragung: Grundlagen der digitalen und analogen Nachrichtenübertragungssysteme*. Springer, Berlin, 8. edition, 2002.

[80] J. Ohta, A. Uehara, T. Tokuda, and M. Nunoshita. Pulse-Modulated Vision Chips with Versatile-Interconnected Pixels. In *IPDPS Workshops, Parallel and Distributed Processing, 15 IPDPS 2000 Workshops*, volume 1800 of *Lecture Notes in Computer Science*. Springer, 2000.

[81] J. Ohta, K Yamamoto, T. Hirai, K. Kagawa, M. Nunoshita, M. Yamada, Y. Yamasaki, S. Sugishita, and K. Watanabe. An Image Sensor With an In-Pixel Demodulation Function for Detecting the Intensity of a Modulated Light Signal. *IEEE Transactions on Electron Devices*, 2003.

[82] B. Pain, L. Matthies, B. Hancock, and C. Sun. A Compact Snap-shot Range-Imaging Receiver. In *IEEE Workshop on CCD and Advanced Image Sensors*, pages 234–237, Nagano, June 2005.

[83] M.J.M. Pelgrom, C.J. Duinmaijer, and A.P.G. Welbers. Matching Properties of MOS Transistors. *IEEE Journal of Solid-State Circuits*, 24(5):1433–1440, October 1989.

[84] P.-F. Rüedi, P. Heim, F. Kaess, E. Grenet, F. Heitger, P.-Y. Burig, S. Gyger, and P. Nussbaum. A 128x128 Pixel 120dB Dynamic Range Vision Sensor Chip for Image Contrast and Orientation Extraction. In *Digest of Technical Papers. ISSCC. 2003 IEEE International Solid-State Circuits Conference*, volume 1, 2003.

[85] C. Ruhland. Nachrichtenübermittlung. Vorlesungsskript, Institut für digitale Kommunikationssysteme, University of Siegen.

[86] B.E.A. Saleh and M.C. Teich. *Fundamentals of Photonics*. Wiley, New York, 1991.

[87] S.V. Sarwate and M.B. Pursley. Crosscorrelation Properties of Pseudorandom and Related Sequences. *Proceedings of the IEEE*, 68(5):593–619, 1980.

[88] J. Schiller. *Mobilkommunikation: Techniken für das allgegenwärtige Internet*. Addison-Wesley, München, 2000.

[89] D.K. Schroder. The Concept of Generation and Recombination Lifetimes in Semiconductors. *IEEE Transactions on Electron Devices*, 29(8):1336–1338, August 1982.

[90] D.K. Schroder. Carrier Lifetimes in Silicon. *IEEE Transactions on Electron Devices*, 44(1):160–170, January 1997.

[91] D.K. Schroder. Private communication, May 2006.

[92] R. Schwarte. Ein Jahrzehnt 3D-Bilderfassung - Zielsetzungen, Fortschritte, Zukunftsperspektiven. In *Messen in der Fertigung: 3D-Messtechnik in Produktion und Entwicklung (Fachtagung Optische Formerfassung)*. DGZfP, VDI/VDE-GMA, Stuttgart, 1999.

[93] R. Schwarte, G. Häusler, and R.W. Malz. *Handbook of Computer Vision and Applications*, chapter Three-Dimensional Imaging Techniques. Academic Press, April 1999.

[94] R. Schwarte, Z. Xu, H.-G. Heinol, J. Olk, R. Klein, B. Buxbaum, H. Fischer, and J. Schulte. New electro-optical mixing and correlating sensor: facilities and applications of the photonic mixer device (PMD). In *Proc. of SPIE*, volume 3100, pages 245–253, 1997.

[95] P. Seitz. *Handbook of Computer Vision and Applications*, chapter Solid-State Image Sensing. Academic Press, April 1999.

[96] P. Seitz. Unified analysis of the performance and physical limitations of optical range-imaging techniques. In *1st Range Imaging Research Day*, pages 9–19. Eidgenössische Technische Hochschule Zurich, September 2005.

[97] P. Seitz and N. Blanc. CCD and APS/CMOS technology for smart pixels and image sensors. In *Detectors and Associated Signal Processing. Edited by Chatard, Jean-Pierre; Dennis, Peter N. J. Proceedings of the SPIE, Volume 5251, pp. 142-153 (2004).*, pages 142–153, February 2004.

[98] C.H. Séquin and M.F. Tompsett. *Charge Transfer Devices*. Academic Press, Inc., New York, 1975.

[99] I. Shcherback, A. Belenky, and O. Yadid-Pecht. Empirical Dark Current Modeling for Complementary Metal Oxide Semiconductor Acitve Pixel Sensor. In *SPIE Optical Engineering*, volume 41, San Jose, 2002.

[100] W. Shockley and W.T. Read. Statistics of the Recombinations of Holes and Electrons. *Physical Review*, 87(5):835–842, September 1952.

[101] M.I. Skolnik. *Introduction to Radar Systems*. McGraw-Hill, New York, 2001.

[102] T. Spirig, M. Marley, and P. Seitz. The Multitap Lock-In CCD with Offset Subtraction. *IEEE Transactions on Electron Devices*, 1997.

[103] Thomas Spirig. *Smart CCD/CMOS Based Image Sensors with Programmable, Real-time, Temporal and Spatial Convolution Capabilities for Applications in Machine Vision and Optical Metrology*. PhD thesis, ETH Zurich (Switzerland), Zurich, 1997.

[104] Thomas Spirig and Peter Seitz. Vorrichtung und Verfahren zur Detektion und Demodulation eines intensitätsmodulierten Strahlungsfeldes. Patentschrift, July 1996. DE 44 40 613 C1.

[105] S.M. Sze. *Semiconductor Devices - Physics and Technology*. John Wiley & Sons, Inc., New York, 2002.

[106] A.J.P. Theuwissen. *Solid-State Imaging with Charge-Coupled Devices*. Kluwer Academic Publishers, Dordrecht, 1995.

[107] H. Tian, X.Q. Liu, S.H. Lim, S. Kleinfelder, and A. El Gamal. Active Pixel Sensors Fabricated in a Standard 0.18um CMOS Technology. In *Proc. of SPIE Electronic Imaging 2001*, volume 4306, San Jose, January.

[108] T. Ushinaga, I. A. Halin, T. Sawada, S. Kawahito, M. Homma, and Y. Maeda. A QVGA-size CMOS time-of-flight range image sensor with background light charge draining structure. In B. D. Corner, P. Li, and M. Tocheri, editors, *Three-Dimensional Image Capture and Applications VII. Proceedings of the SPIE*, volume 6056, pages 34–41, February 2006.

[109] H. van Brug and R.G. Klaver. On the effective wavelength in two-wavelength interferometry. *Journal of Optics. A: Pure and Applied Optics*, 7:1465–1471, 1998.

[110] M. J. Van der Tol and S. G. Chamberlain. Potential and Electron Distribution Model for the Buried-Channel MOSFET. *IEEE Transactions on Electron Devices*, 36(4):670–689, April 1989.

[111] L. Viarani, D. Stoppa, L. Gonzo, M Gottardi, and A. Simoni. A CMOS Smart Pixel for Active 3-D Vision Applications. *IEEE Sensors Journal*, 4(1):145–152, 2004.

[112] Oliver Vietze. *Active pixel image sensors with application specific performance based on standard silicon CMOS processes*. PhD thesis, ETH Zurich (Switzerland), Zurich, 1997.

[113] H. Vogel. *Gerthsen Physik*. Springer Verlag, Berlin, 20. edition, 1999.

[114] R.H. Walden, R.H. Krambeck, R.J. Strain, J. McKenna, N.L. Schryer, and G.E. Smith. The Buried Channel Charge Coupled Device. *The Bell System Technical Journal*, pages 1635–1640, September 1972.

[115] L. Warmerdam and H. Wallinga. A bruied-channel charge-coupled device with non-overlapping gate structure for a CMOS/BCCD process. *Semiconductor Science Technology*, 7:658–663, 1992.

[116] L. Wiegrebe, Den Baum trotz der Blätter sehen - Wie Fledermäuse komplexe 3D-Strukturen erkennen, http://www.innovations-report.de/html/berichte/biowissenschaften_chemie/bericht-27656.html, 2004.

[117] M. Willemin, N. Blanc, G.K. Lang, S. Lauxtermann, P. Schwider, P. Seitz, and M. Wäny. Optical characterization methods for solid-state image sensors. *Optics and Lasers in Engineering*, 36(2):185–194, 2001.

[118] Georg Wiora. *Optische 3D-Messtechnik: Präzise Gestaltvermessung mit einem erweiterten Streifenprojektionsverfahren*. PhD thesis, Ruprechts-Karls-Universität Heidelberg, 2001.

[119] H.-S. Wong. Technology and Device Scaling Considerations for CMOS Imagers. *IEEE Transactions on Electron Devices*, 43(12):2131–2142, December 1996.

[120] D.X.D. Yang, B. Fowler, and A. El Gamal. A Nyquist-Rate Pixel-Level ADC for CMOS Image Sensors. *IEEE Journal of Solid-state Circuits*, 34(3):348–356, March 1999.

[121] G. Yang, O. Yadid-Pecht, C. Wrigley, and B. Pain. A Snap-Shot CMOS Active Pixel Imager for Low-Noise, High-Speed Imaging. In *IEDM'98 Technical Digest., International Electron Devices Meeting*, pages 45–48, December 1998.

[122] F. Zappa, M. Ghioni, S. Cova, and L. Varisco. Integrated array of avalanche photodiodes for single-photon counting. In *Proceeding of the 27th European Solid-State Device Research Conference (ESSDERC)*, September 1997.

[123] F. Zappa, A.L. Lacaita, S.D. Cova, and P. Lovati. Solid-state single-photon detectors. *Optical Engineering*, 35(4):938–945, April 1996.

[124] H.J. Zepernick and A. Finger. *Pseudo Random Signal Processing: Theory and Application*. Wiley, 2005.

[125] E. Zimmermann, Y. Salvadé, and R. Dändliker. Stabilized three-wavelength source calibrated by electronic means for high-accuracy absolute distance measurement. *Optics Letters*, 21:531–533, 1996.

[126] O. Zinke and H. Brunswig. *Hochfrequenztechnik*, volume 1. Springer, 6. edition, 2000.

[127] M. Žˆivković. A Table of Primitive Binary Polynomials. *Mathematics of Computation*, 62(205):385–386, 1994.

[128] M. Žˆivković. Table of Primitive Binary Polynomials.II. *Mathematics of Computation*, 63(207):301–306, 1994.

Acknowledgments

First of all I would like to thank my supervisor Prof. Dr. Peter Seitz for his guidance during the last four years. He inspired me to develop fancy pixel structures that are based on common semiconductor technologies. His optimistic attitude is very motivating and makes one very creative.
The co-examination was accepted by Prof. Dr. Edoardo Charbon from the EPFL in Lausanne, Prof. Dr. Pierre-André Farine from the University of Neuchâtel and Prof. Dr. Shoji Kawahito from the Shizuoka University in Japan. I would like to thank them for their availability.

I am thankful to Dr. Nicolas Blanc, first head of the image sensing group and later head of the photonics division, and Dr. Felix Lustenberger, recently head of the image sensing group. They enabled the accomplishement of this work and made possible the impressive team spirit of our young group, resulting in numerous innovations and winning awards during the last few years.

In addition to that, I acknowledge the support of Thierry Oggier. He introduced me to the world of 3D-image sensing based on smart CCD/CMOS lock-in pixels and taught me his pragmatic way of working. His wide knowledge was helpful in all kinds of measurements, setting up the test equipment and finding new pixel concepts and technologies. Thanks to him I am able to understand a little bit the dialect of Valais people, which is really challenging, and we know the secrets of the real Valais Raclette.

I am very grateful to Dr. Rolf Kaufmann for his support in the semiconductor simulations, the pixel design and the quantum-efficiency analysis. His understanding of the physical fundamentals is appreciated. The long discussions with him, either at CSEM or at the "Turbinenhalle", were stimulating and entertaining. And Rolf, what happens with the holes when they recombine with electrons?
Further, I like to remember the invaluable cooperation of Michael Lehmann. Only his help in the various aspects of the pixel design and his creativity enabled the realisation of the new pixel concepts.

Many thanks go to all members of the image sensing group, with whom I was able to work. Matthias Schweizer helped in the implementation of the real-time 3D-camera that uses pseudo-noise optical modulation. Michael Richter's support in all kinds of questions concerning the sensor test equipment is very appreciated. Michael Stamm participated in the measurements with the high-dynamic range pixels. Peter Metzler contributed to the design of the analogue components of the test boards and I could profit from his long-term experience.
The motivation of the image sensing group was impressive. The following persons played an important role in holding this extraordinary group spirit: Stephan Beer, Roger Cook, Jörn Sparre Pedersen, Thierry Zamofing, Christiane Gimkiewicz, Claus Urban, Simon Neukom, Simon Schneiter and David Beyeler. I thank all of them for making this group so interesting to work with.

I like to remember the agreeable cooperation with the students Björn Rüegg from the Fachhochschule Rapperswil, Adrian Steiner from the Fachhochschule Sion and M'Hamed Ali El Mechat from the EPFL in Lausanne.

At this point I would also like to thank Prof. Dr. Rudolf Schwarte from the University of Siegen, Dr. Marcel Rohner and Dr. Knut Siercks, both from Leica Geosystems AG in Heerbrugg, for the initial education in 3D-imaging during my studies at the University of Siegen and for the motiviation to pursue a Ph.D. thesis at CSEM in the field of real-time 3D-imaging. Marcel, I was really able to test both some bad and some good ideas of pixel structures that I could not have done anywhere else. So I made the right decision!

Special thanks go to Graham Lang who took care of my rather bad English constructions in this work. He corrected everything with great patience.

I am happy having found a good friend, with whom I worked in the same office for more than two years: Philipp Zeller gave me helpful advice in all cultural and social ambiguities. He taught me the complexity and understanding of the Swiss dialects. I already miss the Philip Maloney sessions with him. They really had become a cult.

Last but not least I would like to thank my parents Hiltrud and Karl-Wilhelm Büttgen, my sister Dagmar and her husband Tobias Pöhland and in particular my girl-friend Nadine Scholl for the patience and mental motivation during the last four years. I love all of you!

Bernhard Büttgen
Zürich, 2006

Curriculum Vitae

Bernhard Büttgen
Born December 9, 1976 in Kirchen, Germany, single

Address Breitenstrasse 10, 8134 Adliswil (Switzerland)

Education

1984 - 1987	*Primary School* in Herdorf (Germany)
1987 - 1996	*Secondary School* in Neunkirchen (Germany) Termination: A-level
1996 - 1997	*Military Service*
1997 - 2002	*Studies in Electrical Engineering* at the University of Siegen (Germany) Termination: Diploma
1999 - 2000	*Technical Student at CERN* in Geneva (Switzerland) Development of Communication Software in an Industrial Environment
2002	*Diploma work at Leica Geosystems AG* in Heerbrugg (Switzerland) in cooperation with the Institute of Signal Processing (University of Siegen) Examination and Set-Up of a new Laserradar based on MSM-PMD-Technology
2003 - 2006	*Ph.D. student at CSEM* in Zurich (Switzerland) Engaged by the image sensing group of Dr. Felix Lustenberger as Ph.D. student of Prof. Dr. Peter Seitz.

Other activities

1998	*Student Assistant* at the Institute of High Frequency Technology (Siegen)
Since 2004	*Student Mentor*, Alumni Mentoring Programme (University of Siegen)

Publication List

First author:

[1] B. Büttgen and P. Seitz. Robust and Accurate Optical Time-of-Flight Range Imaging for Outdoor Applications. In preparation.

[2] B. Büttgen, M.-A. El Mechat, F. Lustenberger, and P. Seitz. Pseudo-Noise Optical Modulation for Real-Time 3D-Imaging with Minimum Interference. *IEEE Transactions on Circuits and Systems*, 2006. Submitted for publication: June 1, 2006.

[3] B. Büttgen, F. Lustenberger, and P. Seitz. Demodulation Pixel Based on Static Drift Fields. *IEEE Transactions on Electron Devices*, 53(11):2741–2747, November 2006.

[4] B. Büttgen, T. Oggier, M. Lehmann, R. Kaufmann, S. Neukom, M. Richter, M. Schweizer, D. Beyeler, R. Cook, C. Gimkiewicz, C. Urban, P. Metzler, P. Seitz, and F. Lustenberger. High-speed and high-sensitive demodulation pixel for 3D imaging. In *Three-Dimensional Image Capture and Applications VII. Proceedings of the SPIE*, volume 6056, pages 22–33, February 2006.

[5] B. Büttgen, T. Oggier, M. Lehmann, R. Kaufmann, and F. Lustenberger. CCD/CMOS Lock-In Pixel for Range Imaging: Challenges, Limitations and State-of-the-Art. In *1st Range Imaging Research Day*, pages 21–32. Swiss Centre for Electronics and Microtechnology (CSEM), Eidgenössische Technische Hochschule Zurich, September 2005.

[6] B. Büttgen, T. Oggier, R. Kaufmann, P. Seitz, and N. Blanc. Demonstration of a novel drift field pixel structure for the demodulation of modulated light waves with application in three-dimensional image capture. In *Three-Dimensional Image Capture and Applications VI. Edited by Corner, Brian D.; Li, Peng; Pargas, Roy P. Proceedings of the SPIE*, volume 5302, pages 9–20, April 2004.

[7] B. Büttgen. Noise Measurement with a Digital Storage Oscilloscope. In *Proc. MICRO.tec 2003, 2nd VDE World Microtechnologies Congress*, pages 589–594, Munich, October 2003. VDE Verlag.

Co-author:

[1] T. Oggier, B. Büttgen, F. Lustenberger, R. Becker, B. Rüegg, and A. Hodac. SwissRanger SR3000 and First Experiences based on Miniaturized 3D-TOF Cameras. In *1st Range Imaging Research Day*, pages 97–108. Swiss Centre for Electronics and Microtechnology (CSEM), Eidgenössische Technische Hochschule Zurich, September 2005.

[2] Y. Liebold, D. Haensse, T. Oggier, B. Büttgen, P. Metzler, P. Niederer, F. Lustenberger, and M. Wolf. Quantitative examination of biological tissue with frequency-domain near-infrared spectroscopy: Development of an experimental setup on the basis of a 3D-TOF camera. 2005.

[3] M. Lehmann, T. Oggier, B. Büttgen, C. Gimkiewicz, M. Schweizer, R. Kaufmann, F. Lustenberger, and N. Blanc. Smart pixels for future 3D-TOF sensors. In *IEEE Workshop on CCD and Advanced Image Sensors*, Nagano, June 2005.

[4] T. Oggier, R. Kaufmann, M. Lehmann, B. Büttgen, S. Neukom, M. Richter, M. Schweizer, P. Metzler, F. Lustenberger, and N. Blanc. Novel pixel architecture with inherent background suppression for 3D time-of-flight imaging. In *Proceedings of the SPIE*, volume 5665, pages 1–8, December 2004.

[5] T. Oggier, R. Kaufmann, M. Lehmann, P. Metzler, G. Lang, M. Schweizer, M. Richter, B. Büttgen, N. Blanc, K. Griesbach, B. Uhlmann, K.-H. Stegemann, and C. Ellmers. 3D-Imaging in Real-Time with Miniaturized Optical Range Camera. In *Proc. OPTO conference*, Nürnberg, 2004.

[6] R. Kaufmann, B. Büttgen, T. Oggier, N. Blanc, and P. Seitz. Optical Time-of-Flight 3D Imaging with a Novel Pixel Architecture Based on Lateral Electric Fields. In *Topical Meeting on Advanced Imaging Techniques*, pages 19–20, Delft, 2003. ISBN 3-00-012842-5.

Patents:

[1] Bernhard Büttgen, Felix Lustenberger, and Peter Seitz. Large-area pixel for use in an image sensor. European Patent Application, Publication date: February 8, 2006. EP 1 624 490 A1.

[2] Bernhard Büttgen. Device and method for the demodulation of modulated electromagnetic fields, Filing date: October 19, 2005.

[3] Bernhard Büttgen, Michael Lehmann, Thierry Oggier, Simon Neukom, and Felix Lustenberger. Device and method for the demodulation of modulated electric signals, Filing date: October 19, 2005.

[4] Thierry Oggier, Bernhard Büttgen, and Michael Lehmann. Highly sensitive, fast pixel for use in an image sensor, Filing date: August 11, 2005.