Resolving Measurement Errors Inherent with Time-of-Flight Range Imaging Cameras

A thesis submitted in fulfilment of
the requirements for the degree of

Doctor of Philosophy

in Electronic Engineering at

The University of Waikato

by

Refael Zabdi Whyte

THE UNIVERSITY OF
WAIKATO
Te Whare Wānanga o Waikato

2015

Abstract

Range imaging cameras measure the distance to objects in the field-of-view (FoV) of the camera, these cameras enable new machine vision applications in robotics, manufacturing, and human computer interaction. Time-of-flight (ToF) range cameras operate by illuminating the scene with amplitude modulated continuous wave (AMCW) light and measuring the phase difference between the emitted and reflected modulation envelope. Currently ToF range cameras suffer from measurement errors that are highly scene dependent, and these errors limit the accuracy of the depth measurement. The major cause of measurement errors is multiple propagation paths from the light source to pixel, known as multi-path interference. Multi-path interference typically arises from: inter-reflections, lens flare, subsurface scattering, volumetric scattering, and translucent objects.

This thesis contributes three novel methods for resolving multi-path interference: coding in time, coding in frequency, and coding in space. Time coding is implemented by replacing the single frequency amplitude modulation with a binary sequence. Fundamental to ToF range cameras is the cross-correlation between the reflected light and a reference signal. The measured cross-correlation depends on the selection of the binary sequence. With selection of an appropriate binary sequence and using sparse deconvolution on the measured cross-correlation the multiple return path lengths and their amplitudes can be recovered. However, the minimal resolvable path length is dependent on the highest frequency in the binary sequence.

Frequency coding is implemented by taking multiple measurements at different modulation frequencies. A subset of frequency coding is operating the camera in a mode analogous to stepped frequency continuous wave (SFCW). Frequency coding uses techniques from radar to resolve multiple propagation paths. The minimal resolvable path length is dependent on the camera's modulation bandwidth and the spectrum estimation technique used to recover distance, and it is shown that SFCW can be used to measure depth of objects

behind a translucent sheet, while AMCW measurements can not. Path lengths below quarter a wavelength of the highest modulation frequency are difficult to resolve.

The use of spatial coding is used to resolve diffuse multi-path interference. The original technique comes from direct and global separation in computer graphics, and it is modified to operate on the complex data produced by a ToF range camera. By illuminating the scene with a pattern the illuminated areas contain the direct return and the scattering (global return). The non-illuminated regions contain the scattering return, assuming the global component is spatially smooth. The direct and global separation with sinusoidal patterns is combining with the sinusoidal modulation signal of ToF range cameras for a closed form solution to multi-path interference in nine frames. With nine raw frames it is possible to implement direct and global separation at video frame rates. The RMSE of a corner is reduced from 0.0952 m to 0.0112 m. Direct and global separation correctly measures the depth of a diffuse corner, and resolves subsurface scattering however fails to resolve specular reflections. Finally the direct and global separation is combined with replacing the illumination and reference signals with a binary sequence. The combination allows for resolving diffuse multi-path interference present in a corner, with the sparse multi-path interference caused mixed pixels between the foreground and background. The corner is correctly measured and the number of mixed pixels is reduced by 90%.

With the development of new methods to resolve multi-path interference ToF range cameras can measure scenes with more confidence. ToF range cameras can be built into small form factors as they require a small number of parts: a pixel array, a light source and a lens. The small form factor coupled with accurate range measurements allows ToF range cameras to be embedded in cellphones and consumer electronic devices, enabling wider adoption and advantages over competing range imaging technologies.

Acknowledgements

"And private in his chamber pens himself, shuts up his windows, locks fair daylight out, and makes himself and artificial night. Black and portentous must this humor prove unless good counsel may the cause remove." William Shakespeare, Romeo and Juliet.

This quote is originally about Romeo brooding, but could equally describe the long hours collecting data in an optics lab, and getting frustrated with poor quality data. Something that has been a regular experience during the course of this research. This thesis and research project has been an interesting, varied and long journey. At times the idea of finishing was a distant dream, never thought I would write a thesis, but here I am. I would like to thank my supervisors Adrian Dorrington, Michael Cree, and Lee Streeter for all there help and guidance in this undertaking. Special thanks to Lee Streeter for all the assistance with signal processing and mathematics, without him I would have been truly lost. My parents for their support of my higher education, and help with proofreading and editing. My fellow PhD students in the C.1.01 space, Richard, Robert and John it was great working along side all of you when I was based at Waikato. Microsoft® for the opportunity to complete a year long internship in Silicon Valley. The technical knowledge and experience gained from working on cutting edge technology was an invaluable experience. Last but not least my friends and flatmates over the years who have made the last few years incredibly enjoyable.

Contents

List of Figures

List of Tables

List of Symbols

Common Variables, Constants and Functions

j	The imaginary unit, given as $\sqrt{-1}$
a	Amplitude of reflected light
ω	Angular frequency of amplitude modulation
N	Number of samples on the correlation signal, $h(\tau)$
M	Number of modulation frequencies sampled
K	Number of propagation paths from the light source to pixel
ϕ	Phase of the amplitude modulation
c	Speed of light
τ	Phase offset of reference signal
d	Radial distance measurement
ζ	Complex time-of-flight range measurement, $ae^{j\phi}$
b	Background light
A	Amount of charge stored on tap 1 of a time-of-flight pixel
B	Amount of charge stored on tap 2 of a time-of-flight pixel
L	The radiance
R	The real component of $h(\tau)$
Q	The imaginary component of $h(\tau)$
T	Integration time for one raw frame of the time-of-flight range camera
Υ	Maximum frequency supported by the time-of-flight range camera
d_{amg}	The ambiguity distance, which is the distance at which the phase, ϕ, wraps from 2π back to 0.
Γ	The number of bits in a binary sequence
T_c	The transmission length of each bit in a binary sequence
p_γ	A binary sequence, with each bit indexed by γ
H	Circulant matrix of the measured cross-correlation
Λ	Circulant matrix of a binary sequence
BW	Bandwidth of the measured frequencies

Δf	Separation between measured frequencies
L_d	Direct component of radiance
L_g	Global component of radiance
L_s	Specular inter-reflection component of radiance
Ξ	The number of illumination sources
$s(t)$	The reflected light as a function of time
$g(t)$	The reference signal of the time-of-flight pixel
$h(\tau)$	The measured cross-correlation between $s(t)$ and $g(t)$

Chapter 1

Introduction

Over the last 15 years digital cameras have become prevalent in consumer electronics. However, digital cameras are limited to measuring 2D colour images of the three dimensional universe, providing machines with only a limited perception of their environment. This has led to the development of 3D vision systems which allow machines to perceive the universe in 3D, enabling robots to perceive the world in the same way as humans. In the last decade, 3D cameras have advanced from the research laboratory to industrial and consumer applications. This development is blazing the way for a revolution in robotics, manufacturing and human computer interaction. Before this dream can be fully realized, the measurement errors present in current 3D vision systems need to be resolved. In this thesis measurement errors inherent with time-of-flight (ToF) range cameras are explored and solutions developed. Optimistically this thesis will contribute to the wide spread adoption of 3D range imaging cameras.

1.1 Range Imaging

Range measurement has a long and varied history, and various technologies can be categorised into active and passive techniques. For active techniques the measurement device transmits a signal into the environment and measures the reflection to estimate range. Passive techniques rely on ambient light and use geometry to calculate distance.

Passive techniques were developed for surveying, and use geometry extensively in order to accurately measure distance and height in geographic areas for map production. Stereo imaging is the most common passive 3D imaging technique. It mirrors the human vision system by using triangulation between

cameras and relies on ambient light. For stereo imaging to achieve depth measurement, features in the scene are matched between cameras. In scenes with limited features, for example white walls, stereo matching fails.

Active range measurement began with the development of sonar during the First Wold War. Sonar was used to locate submarines and was further refined during the Second World War. Radar was developed during the Second War Two for detecting aircraft. These early systems were time of flight based, transmitting a pulse of energy, sound or radio, and measuring the time delay between the emitted and received pulse. In later years, modulation of the carrier allowed for continuous operation where the distance travelled was encoded by a change of phase or frequency in the reflection. Optical range measurement began in the 1960's shortly after the invention of the laser. Light detection and ranging (LIDAR) measures the distance from the device to a single point using a focused laser beam. In order to measure an entire scene, the laser beam has to be physically shifted, normally using motorized mirrors. LIDAR was originally used in optical meteorology to measure clouds and is used extensively in surveying and simultaneous localization and mapping (SLAM) applications.

Full-field range imaging cameras measure the distance to each pixel of the camera simultaneously. With amplitude modulated continuous wave (AMCW) time-of-flight (ToF) range cameras, a specialized pixel design is employed to measure distance. AMCW ToF range cameras work by illuminating the scene with amplitude modulated light and measuring the phase difference in the modulation envelope between the emitted and reflected light. One of the first AMCW ToF range cameras was developed by Lange (2000), as part of his Ph.D at the University of Siegen, Germany. Lange's ToF range camera had an array size of 64 by 25 pixels. The use of custom pixels allows for innovative imaging systems, however, it increases the cost and length of development. There are two methods of ToF range measurement: direct and indirect.

Direct ToF directly measures the time between transmitting a pulse and receiving the reflection. This requires precise timing, which is expensive and difficult. Direct ToF cameras have been realized with an array of 32 by 32 avalanche photodiodes by Niclass *et al.* (2005).

Indirect ToF range cameras infer the time of flight through the measurement of propagation time effects on some property of the emitted light. In one common case this is achieved by illuminating the scene with an amplitude modulated light source and measuring the phase shift of the modulation en-

velope in the reflected light. There are two implementations to recover the phase: a global shutter and lock-in pixels.

A global shutter is implemented so it can open and close at the same, or similar, frequency range as the light's amplitude modulation. The ratio of light striking the shutter to light transmitted through the shutter encodes the distance travelled. A global shutter was implemented by Dorrington *et al.* (2007) by using an image intensifier as the shutter. Tadmor *et al.* (2014) implemented a global shutter on a CCD array with a novel pixel design.

Lock-in pixels mix the returning light with a reference signal in order to measure the correlation between the reflected and reference signal. When a shutter is closed the returning light is discarded, while with lock-in pixels all the reflected light is collected. The phase of the correlation signal encodes the distance travelled. Lock-in pixels have been implemented by Lange (2000), Büttgen *et al.* (2004), and Van Nieuwenhove *et al.* (2008). Solid state implementations have the advantage of being realized in a standard CMOS process, leveraging existing semiconductor manufacturing capabilities.

Time-of-flight range imaging is a competing active 3D imaging technology. One of the other competing technologies is structured light. Structured light techniques overcome passive stereo imaging's limitation by combining a spatial light modulator (SLM) and an imaging sensor to measure distance. The triangulation is computed based on the projected patterns from the SLM. As structured light works on triangulation, physical separation between the SLM and camera is required. This limits the physical size of structured light range cameras. Structured light cameras leverage the existing technology of digital cameras and projection systems, both of which have made significant progress in the last two decades. Lanman and Taubin (2009) demonstrated how to build a structured light range camera with off the shelf components. The original Xbox® Kinect ® camera used structured light technology. The Kinect® camera brought available and affordable full field range imaging cameras to market.

1.2 Commercial Activity

The current companies producing ToF range cameras are: PMD technologies, MESA Imaging, Softkinetic, Microsoft®, Samsung®, Panasonic® and Texas Instruments®. PMD technologies was founded in 2002 as a spin out company of the University of Seigen in Germany. PMD's focus is on selling ToF

sensors and pixel technology. MESA imaging was founded in 2006 as a spin out of the Swiss Center for Electronics and Microtechnology in Switzerland. MESA's main focus has been on high quality cameras for industrial applications. A notable application of MESA's technology is a ToF range camera used for automated cup placement on a cow's udder for automated milking machines. Optrima was founded in 2009 as a spin out company of Vrije University in Belgium. Optrima focused on ToF sensor development. In 2011 they merged with SoftKinetic to provide a full solution stack of hardware and software. SoftKinetic has focused on consumer applications and licensed their technology for use in the first generation of Intel RealSense® cameras and gesture technology to Sony for use in the PlayStation 4®. 3DV was a ToF range camera design company founded in 2001 in Israel. They were acquired by Microsoft® in 2009. Canesta was a ToF range camera startup located in Silicon Valley and founded in 2001. They were acquired by Microsoft® in 2010. Canesta technology was used in the Xbox® One Kinect® camera. At its release in 2013, the Xbox® One Kinect® was the largest ToF sensor at 512 by 424 pixels. Samsung® has been active in developing ToF technology, and published research papers on RBGZ range cameras. However, as of 2015 they have not released a commercial product. Panasonic® released a ToF camera in 2010 and recently announced plans for larger sensor resolutions. In 2014 Texas Instruments® released a ToF sensor and controller chip, with the pixel technology licensed from SoftKinetic.

The commercial world surrounding non-ToF range imaging technology has also been active in recent years. PrimeSense was a structured light range camera company that was founded in Israel in 2005. Microsoft® licensed their technology for the original Xbox® Kinect® camera. PrimeSense was acquired by Apple® in 2013 for $300 Million USD. Intel® has invested $100 Million USD in perceptual computing focused on natural user interfaces (NUI) for computers and portable devices. In 2011 Intel® acquired the structured light range camera company InVision Biometrics. In 2015 Intel® started shipping laptops containing range cameras under the brand RealSense®. Leap Motion is a company founded in 2010 that develops novel hardware for measuring hand and finger motion to interact with computers. Their technology is based on the combination of stereo and photometric depth measurement techniques. Leap Motion has raised over $40 Million USD as of 2014. Google® released the Trango development phone in 2014 with an integrated structured light range camera. Later in 2014 Google® released a tablet with an integrated ToF range

camera. These devices are being used to develop innovative applications that use range imaging cameras on mobile platforms.

There has been significant capital investment and development on range imaging technology in the last decade and applications in the consumer market are emerging, such as computer interfaces, security, face recognition, automatic robotic navigation and assisted driving. As with current machine vision algorithms, processing and interpretation of range images is computationally expensive and the price for computation is always decreasing. It is anticipated range cameras will be integrated into consumer devices in coming years and this will lead to new applications of range imaging.

An AMCW ToF range camera requires: a custom pixel array, a light source capable of high frequency amplitude modulation, and a lens. While the research and development costs of new silicon are high, the final products bill of materials is small and can be manufactured at low cost. Current research on integrating optical elements into silicon wafers would allow for miniaturised system on chip (SoC) ToF range cameras. With significant investment capital, the availability and usage of 3D cameras is expected to grow significantly in the coming years.

1.3 Thesis Contribution

Time-of-flight range cameras suffer from measurement errors that limit their performance. A major cause of error is when multiple propagation paths are present from the light to the same pixel, and can be caused by inter-reflections, subsurface scattering and translucent objects. This has limited the guaranteed accuracy of ToF measurements because multi-path interference is highly scene dependent. In turn this has limited the adoption of ToF technology. In this thesis three novel methods of resolving multi-path interference in ToF range cameras are presented.

1.3.1 Objectives

The primary objective of this thesis is to find novel solutions to measurement errors in ToF range cameras, with the focus on resolving multi-path interference. ToF range cameras are designed to measure in excess of 30 fps of full field depth information, with low computational complexity, and graceful degradation with noise. Any multi-path interference restoration method developed should ideally not interfere with these properties.

1.3.2 Outline

This work is divided into seven chapters, and one appendix, with the conclusion and future outlook located in chapter 9.

The theory and implementation of time-of-flight range cameras is reviewed in **Chapter 2**. The causes of measurement errors and noise sources in ToF range cameras are discussed. They are: multiple propagation paths, harmonic aliasing, and phase wrapping. The physical causes of multiple propagation paths are explained and a mathematical forward model developed. Current state of the art correction techniques are reviewed for the discussed measurement errors.

In **Chapter 3** the system design process of a ToF camera is presented. Commercial ToF range cameras are inflexible. They have limited controllability and access to low level data. These reasons are behind the motivation to build a custom ToF range camera. The design includes the mechanical mounting, PCB design, light source, FPGA design, embedded software, and a Matlab interface and image processing. The ToF camera is designed around a PMD19k2 ToF sensor that contains 120 by 160 pixels. The performance of the camera design is overviewed and common design errors are explained.

In **Chapter 4** the use of binary sequences in time-of-flight range cameras is explored. The current literature around using binary sequences in ToF cameras is reviewed. Currently, binary sequences are used to allow multiple unsynchronized cameras to operate simultaneously. A new forward model using binary sequences is developed and methods to calculate distance from the new measurement are discussed. The use of sparse deconvolution to resolve multi-path interference is then demonstrated. Different sparse deconvolution methods are tried and the limits of restoration simulated.

In **Chapter 5** sampling over frequency and phase is explored. Currently ToF range cameras sample over phase. By sampling over frequency, the camera operates in a mode analogous to stepped frequency continuous wave (SFCW) radar. Spectrum estimation techniques are used to measure the depth. The advantages with sampling over frequency is that the multi-path interference and harmonic aliasing are measured as higher frequencies and therefore do not interfere. However, the accuracy depends on the spectral estimation technique used.

The transformation between a frequency sweep and a binary sequence is shown in **Chapter 6**. Measurements of a frequency sweep can be transformed into a binary sequence and vice versa. This is demonstrated by resolving

multi-path interference in both domains from the same measurement. The advantages of measuring in each domain are explored.

The use of spatial coding to resolve multi-path interference is shown in **Chapter 7**. By projecting a checkerboard image, multi-path interference can be removed. In the illuminated regions, both direct and multi-path interference are measured. In the non-illuminated regions, only multi-path interference is present. A novel method of mixing the phase shift in a ToF camera and phase shifting a projected pattern is developed. A closed form solution to multi-path interference that requires nine raw frames is developed.

In **Chapter 8** coded ToF from chapter 4 is combined with spatial coding from chapter 7 to solve both sparse and diffuse multi-path interference. In the combination both multi-path interference caused mixed pixels and inter-reflections is resolved.

In **Appendix A**, novel applications of ToF technology, that are not indirect range measurements, are presented. The velocity of objects is measured by operating the camera in a mode analogous to continuous wave (CW) Doppler radar. The camera is used for fluorescence lifetime imaging (FLI). Multiple light sources illuminate the scene simultaneously for both colour range imaging and dynamic relighting.

1.4 Publications Arising from this Thesis

This thesis resulted in two first name journal publications, three journal publications as a contributing author, two first author conference papers, two patents, and four conference papers as a contributing author.

1.4.1 Journal

Whyte, R., L. Streeter, M. J. Cree, and A. A. Dorrington. Application of lidar techniques to time-of-flight range imaging. *Appl. Opt.*, **54**(33), pp. 9654–9664 (2015).

Whyte, R., L. Streeter, M. Cree, and A. Dorrington. Resolving multiple propagation paths in time of flight range cameras using direct and global separation methods. *Optical Engineering*, **54**(11), pp. 113109 (2015).

1.4.2 Patents

R. Whyte, and A. Dorrington. Time of Flight Camera System which Resolves Direct and Multi-path Radiation Components. US Patent 20150253429 (2015).

A. Kadambi, R. Whyte, B. Ayush, S. Streeter, C. Barsi, A. Dorrington, and R. Raskar. Methods and Apparatus for Coded Time-of-Flight Camera US Patent 20150120241 (2015).

1.4.3 Conference

Whyte, R., L. Streeter, M. Cree, and A. Dorrington. Review of methods for resolving multi-path interference in time-of-flight range cameras. In: *2014 IEEE Sensors*, pp. 629–632 (2014).

Whyte, R., A. Bhandari, L. Streeter, M. J. Cree, and A. A. Dorrington. Time frequency duality of time-of-flight range cameras for resolving multi-path interference. In: *29th International Conference on Image and Vision Computing New Zealand (IVCNZ)*, pp. 247–252 (2014).

1.4.4 As a contributing author

A. Kadambi, R. Whyte, A. Bhandari, L. Streeter, C. Barsi, A. Dorrington, and R. Raskar, "Coded time of flight cameras: sparse deconvolution to address multipath interference and recover time profiles," in *ACM Transactions on Graphics (TOG)*, vol. 32, no. 6, p. 167, (2013).

M. Feigin, R. Whyte, A. Bhandari, A. Dorrington, and R. Raskar, "Modeling "wigglin" as a multi-path interference problem in AMCW ToF imaging," *Opt. Express*, vol. 23, no. 15, pp. 19213–19225 (2015).

A. Bhandari, A. Kadambi, R. Whyte, C. Barsi, M. Feigin, A. Dorrington, and R. Raskar, "Resolving multi-path interference in time-of-flight imaging via modulation frequency diversity and sparse regularization," *Optics Letters*, vol. 39, no. 6, pp. 1705–1708 (2014).

A. Kadambi, A. Bhandari, R. Whyte, A. Dorrington, and R. Raskar, "Demultiplexing illumination via low cost sensing and nanosecond coding," In *IEEE International Conference on Computational Photography (ICCP)*, pp.

1–10 (2014).

Barsi, C., Whyte, R., Bhandari, A., Das, A., Kadambi, A., Dorrington, A. A., & Raskar, R. (2014, April). Multi-frequency reference-free fluorescence lifetime imaging using a time-of-flight camera. In *Biomedical Optics* (pp. BM3A-53). Optical Society of America (2014).

Bhandari, A., Barsi, C., Whyte, R., Kadambi, A., Das, A. J., Dorrington, A., & Raskar, R. (2014, July). Coded Time-of-Flight Imaging for Calibration Free Fluorescence Lifetime Estimation. In *Imaging Systems and Applications* (pp. IW2C-5). Optical Society of America (2014).

Chapter 2

Time-of-Flight Range Imaging

The various methods of time-of-flight (ToF) range imaging have been introduced in chapter 1, and the remainder of this thesis is primarily focused on the homodyne amplitude modulated continuous wave (AMCW) method, and as such, only this method will be reviewed in depth. The term AMCW will be omitted for readability, and further references to ToF will normally imply the homodyne AMCW method, unless otherwise stated. In this chapter the theory of ToF range cameras is reviewed. The implementation of the theory in hardware and processing pipeline is explored and the measurement errors, and noise sources are explained. The current state of the art methods for resolving measurement errors are reviewed.

2.1 Theory of Operation

Time-of-flight (ToF) range cameras operate by illuminating the scene with amplitude modulated light (Lange and Seitz, 2001). The phase shift between the emitted modulation envelope and the reflected envelope encodes the distance the light has travelled. The reflected light is expressed as

$$s(t) = a\sin(\omega t - \phi) + b, \tag{2.1}$$

where a is the amplitude of the returning light, ω encodes the modulation frequency, ϕ is the phase offset due to the distance travelled, and b is the background light. To determine the phase, the reflected light is correlated with a reference signal, $g(t)$, which is amplitude modulated at the same frequency,

$$g(t) = \sin(\omega t). \tag{2.2}$$

Each pixel in ToF imaging sensor performs the correlation. The correlation of equations 2.1 and 2.2 is

$$h(\tau) = (s \star g)(t) = \lim_{T \to \infty} \frac{1}{T} \int_{-T/2}^{T/2} s(t) g(t + \tau) \, dt, \tag{2.3}$$

where T is the integration period. When the integral is evaluated the correlation function is

$$h(\tau) = \frac{a}{2} \cos(\omega\tau + \phi) + b, \tag{2.4}$$

where $\omega\tau$ encodes the phase of the reference signal $g(t)$, relative to the light source. Multiple measurements of the correlation signal are taken by phase shifting the reference signal N times, where each individual sample on $h(\tau)$ is called a raw frame. Since equation 2.4 has three unknowns, a, ϕ, and b, at least three raw frames are required to solve for the unknowns, typically four are used (Lange, 2000). The unknowns are calculated by taking the discrete Fourier transform (DFT) over τ. As the first Fourier bin contains the phase and amplitude information. The relationship between the phase and radial distance, d, is

$$d = \frac{c\phi}{2\omega}, \tag{2.5}$$

where c is the speed of light.

2.1.1 Heterodyne AMCW

The previous section covered the operation of ToF range cameras when the reference signal and light source were modulated at the same frequency. When the light source and reference signals are at slightly different frequencies this mode of operation is called heterodyne imaging. Dorrington *et al.* (2007) describes the implementation of a heterodyne range camera. If the illumination has an amplitude modulation frequency of f and the sensor modulation is $f + f_d$, where f_d is the difference in frequency, then by sampling over one cycle

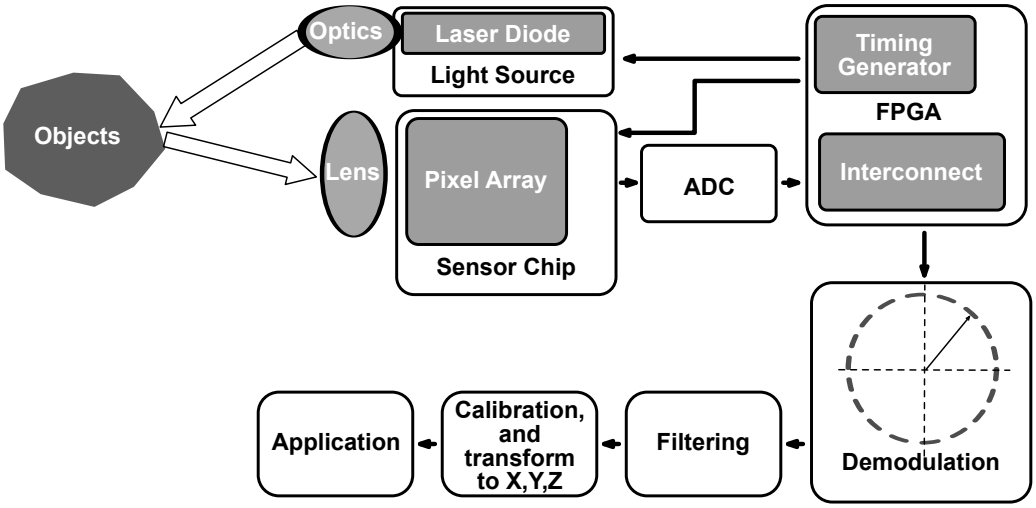


Figure 2.1: A typical time-of-flight range camera processing pipeline from hardware to application.

of the beat signal the phase and amplitude is calculated by

$$\phi = \tan^{-1}\left(\frac{\sum_{n=1}^{N} h_n \sin\left(2\pi n f_d/f_s\right)}{\sum_{n=1}^{N} h_n \cos\left(2\pi f_d/f_s\right)}\right), \tag{2.6}$$

$$a = \frac{2}{N}\sqrt{\left[\sum_{n=1}^{N}\left(h_n \cos\left(2\pi n f_d/f_s\right)\right)\right]^2 + \left[\sum_{n=1}^{N}\left(h_n \sin\left(2\pi n f_d/f_s\right)\right)\right]^2}, \tag{2.7}$$

where f_s is the sampling rate. Heterodyne has the advantage that it is resilient to higher order harmonics in the modulation signal (Dorrington *et al.*, 2008). The implementation of two closely spaced frequencies, which is critical to heterodyne imaging, is more complicated than homodyne range imaging, therefore homodyne is more common.

2.2 Implementation

This section introduces the implementation of a time-of-flight (ToF) range camera and the processing pipeline from raw frames to the depth frames which the end application uses. The diagram in figure 2.1 is an overview of the typical implementation and processing performed by a ToF range camera (Van Nieuwenhove *et al.*, 2008).

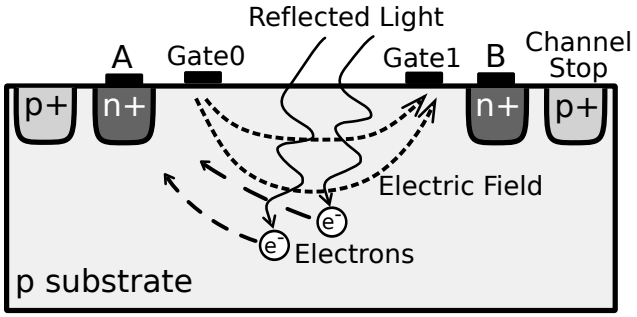


Figure 2.2: Cross section of a simplified time-of-flight range camera pixel design in a standard CMOS process.

2.2.1 Hardware

The cross-correlation from equation 2.4 is realized by a custom pixel architecture. An example pixel implementation in CMOS is shown in figure 2.2 (Xu *et al.*, 1998) (Texas Instruments, 2014). The photons captured in the silicon substrate are converted to electron hole pairs. The electric field applied across the pixel move the electrons to the selected side, where they are stored on a capacitor. The reference signal, $g(t)$ from equation 2.2, sets the direction of the electric field across the pixel by applying certain voltages to Gate0 and Gate1. The electrons migrate to the side of the pixel and are captured in the A and B capacitive regions knowns as taps. This type of pixel is known as a lock in pixel, and can be implemented in both standard CMOS and CCD fabrication processes. This allows current solid state imaging device fabrication processes to be used for ToF pixel arrays. ToF pixels are actively powered therefore power consumption is one of the variables that limits the maximum array size. It is expected the trend of smaller pixels and lower power consumption will continue to allow for more pixels on a sensor. Currently VGA resolution array sizes exist (Bamji *et al.*, 2015) with the road map to one megapixel resolution.

A number of research groups and companies have developed ToF lock in pixels. Lange and Seitz (2001) implemented an array of 64 by 25 ToF pixels, with a pixel size of 65 μm by 21 μm. The array achieved a fill factor of 20% and was fabricated on a slightly modified CMOS process. Büttgen *et al.* (2004) implemented a four tap pixel on a dual CMOS/CCD process. The pixel structure moved electrons to the taps due to lateral drift instead of diffusion, allowing for faster modulation frequencies. The four taps decreases the number of raw frames required for a depth measurement by a factor of two. Van Nieuwenhove *et al.* (2008) implemented a novel pixel design with a current assisted photo

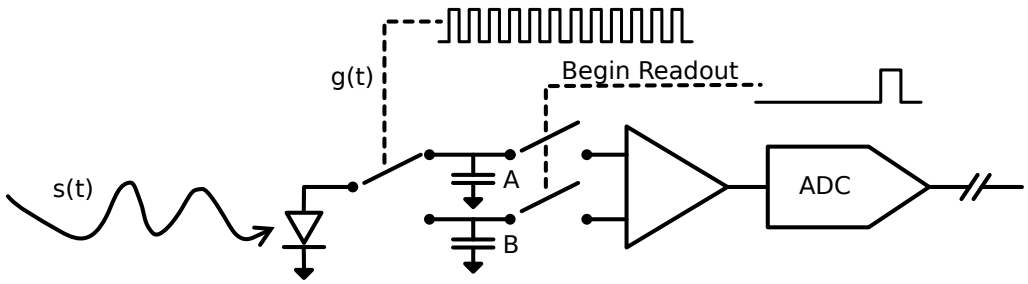


Figure 2.3: An example time-of-flight range camera readout process. Incoming photons are converted into electrons and stored on capacitors during the integration period. During readout the capacitors are read through an operational amplifier and then converted to a digital value for further processing.

demodulator (CAPD) structure for improved pixel performance. There are a number of other pixel implementations in the literature and they work on the same principle of separating the returning photons into different regions of the pixel.

The measured correlation, $h(\tau)$, is calculated by the subtraction of the two taps A and B.

$$h(\tau) = A(\tau) - B(\tau). \tag{2.8}$$

If the tap values, A and B, are added instead of subtracted the pixel measures the radiance of integrated light, including the reflected light and background light. Operating in a mode where the two taps are added together is known as common mode (Texas Instruments, 2014). Common mode is useful applications such as during lens calibration, and implementing optical flow for motion correction.

An example pixel readout circuit is in figure 2.3 (Payne *et al.*, 2009). The capacitors A and B store the electrons during the integration period, then during readout the charge stored is read through an operational amplifier, and the output voltage of the operational amplifier is converted to a digital value by an ADC, to be used to compute range (Gokturk *et al.*, 2004). As with traditional imaging arrays multiple ADCs can be used in parallel to reduce the readout time. The location of the subtraction depends on the pixel implementation, in some architectures the subtraction is performed by the readout amplifier, in others the subtraction is done in the digital logic after the ADC (Bamji *et al.*, 2015).

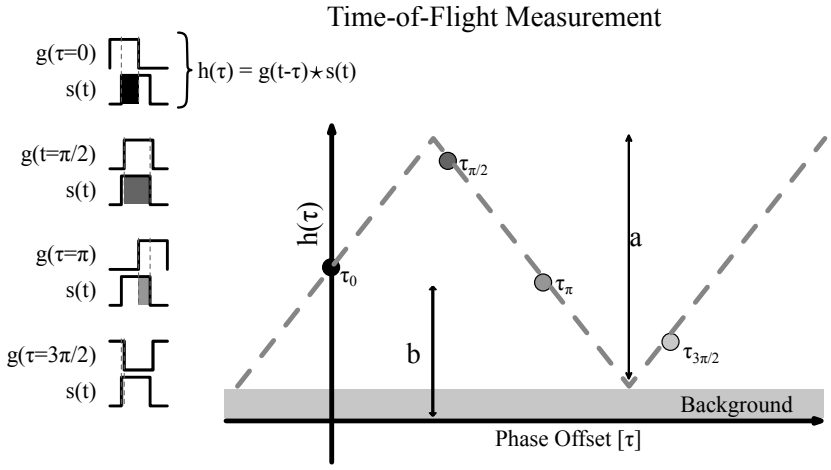


Figure 2.4: Visualization of the measurement process of a time-of-flight range camera.

Due to implementation in digital logic square wave modulation is typically used. Therefore equations, 2.1 and 2.2, for the reflected light and reference signal become

$$s(t) = a \sum_{\upsilon=1,3,\ldots}^{\Upsilon} \left[\frac{\sin(\upsilon(\omega t - \phi))}{\upsilon} \right] + b, \tag{2.9}$$

$$g(t) = \sum_{\upsilon=1,3,\ldots}^{\Upsilon} \frac{\sin(\upsilon \omega t)}{\upsilon}, \tag{2.10}$$

where Υ encodes the bandwidth of the camera. In theory there are an infinite number of harmonics, but in reality the higher order harmonics are attenuated, eventually below the noise floor, due to bandwidth limitations. As with all hardware ToF range cameras are bandlimited devices. When equation 2.4 for the cross-correlation is evaluated the resulting function is

$$h(\tau) = \frac{a}{2} \sum_{\upsilon=1,3,\ldots}^{\Upsilon} \left[\frac{\cos(\upsilon(\omega \tau + \phi))}{\upsilon^2} \right] + b. \tag{2.11}$$

Equation 2.11 is approximately the Fourier series of a triangle wave. Figure 2.4 visualizes the typical ToF camera measurement process of taking four samples on the triangular correlation signal.

The timing generation module generates the modulation signals. A variety of implementations exists for a configurable frequency and phase offset

generator. Bamji *et al.* (2015) used a current controlled oscillator (ICO) to derive the modulation clocks from a common high speed clock. For fine programmable control over the frequency and phase configurable linear feedback shift registers (LFSRs) that produced a pulse when their programmed count was reached were used on the output of the ICO. Payne (2008) used a direct digital synthesizer (DDS) to generate the modulation signals from a reference clock. The DDS allowed for clock generation a few hertz apart making it ideal for heterodyne operation. Jongenelen (2011) used a phase locked loop (PLL) inside a field programable gate array (FPGA) to generate the modulation signals. The PLL uses a voltage controlled oscillator (VCO) and programmable output counters to generate the desired modulation frequency. The phase step of each clock of the PLL can be dynamically reconfigured.

2.2.2 Demodulation

Demodulation is the process of calculating the phase and amplitude, which can be done by taking discrete Fourier transform (DFT) over the raw frames. The phase can be converted to radial distance using equation 2.5. The phase and amplitude information is contained in the first frequency bin so the calculation can be optimized. The measured values can be separated into the real and imaginary components, R and Q respectively by

$$R = \sum_{n=1}^{N} h(\tau_n) \sin(\tau_n), \tag{2.12}$$

$$Q = \sum_{n=1}^{N} h(\tau_n) \cos(\tau_n), \tag{2.13}$$

where there are N samples on the cross-correlation signal. The phase and amplitude from the real and imaginary components is

$$\phi = \tan^{-1}\left(\frac{R}{Q}\right), \tag{2.14}$$

$$a = \frac{\sqrt{Q^2 + R^2}}{2}. \tag{2.15}$$

Commonly four raw frames are used with $\omega\tau = [0, \pi/2, \pi, 3\pi/2]$, then the phase and amplitude calculation is simplified to

$$\phi = \tan^{-1}\left(\frac{h(\tau_1) - h(\tau_3)}{h(\tau_2) - h(\tau_4)}\right), \tag{2.16}$$

$$a = \frac{\sqrt{(h(\tau_1) - h(\tau_3))^2 + (h(\tau_2) - h(\tau_4))^2}}{2}. \tag{2.17}$$

The measurement can be represented as a complex variable, ζ,

$$\zeta = R + jQ = ae^{j\phi}. \tag{2.18}$$

2.2.3 Filtering

ToF range cameras suffer from noisy and erroneous measurements. Filtering is designed to reduce noise, and correct or remove invalid measurements.

A number of noise reduction techniques from 2D imaging have been ported to work on ToF data. The filtering can occur on the raw images, $h(\tau)$, the complex images, R and Q, the amplitude, a, and phase, ϕ, images, or the point cloud, $[X, Y, Z]$. Edge preserving filters are now applied to preserve the edges in the depth image. The bilateral filter was applied to depth images by Le *et al.* (2014). A real time implementation of the bilateral and guided filter with a GPU for ToF range images was done by Wasza *et al.* (2011). Lenzen *et al.* (2011) implemented adaptive total variation on the phase and amplitude images for noise reduction. Median filtering on the depth image for de-noising the SR4000 range camera is supported by Mesa Imaging (2010).

Some of the measured pixels are too corrupted by noise or measurement errors to be useful. The measured amplitude is an estimate of the reflected signal to noise ratio and is used as a confidence measure (Kolb *et al.*, 2009). Reynolds *et al.* (2011) used machine learning techniques to generate a confidence mask for each pixel to remove invalid pixels.

To circumvent the limited array size of ToF cameras they have been mounted with RBG cameras. The combination of data types allows for super resolution and filtering of the depth data to remove invalid pixels. These techniques have been developed by Huhle *et al.* (2008), and Lindner *et al.* (2008). Current commercial cameras such as the Microsoft Xbox Kinect and Softkinetic DepthSense 325 are packaged with a ToF and RBG sensor. Super resolution and up-sampling techniques have been developed to work on depth images. Kim and Yoon (2012) used a method based on Markov random fields to up-sample

a depth image. Schuon *et al.* (2008) used joint bilateral up-sampling for depth super resolution on a ToF camera. Other image enhancement techniques have been applied to ToF cameras. Applying shape from shadow constraints to the range image by using the amplitude image to improve the quality of the depth image by Böhme *et al.* (2010). Kim *et al.* (2014) used four light sources to apply photometric stereo based techniques to enhance the depth image.

2.2.4 Calibration

There are two calibrations required for a ToF range camera. The first one for accurate distance measurement, and the second is a traditional camera calibration. The calibration methods for ToF range cameras has been explored in detail by Lindner *et al.* (2010), and Kahlmann and Ingensand (2008).

Each pixel has a unique phase offset caused by the path length between the modulation driver and pixel. The phase offset of each pixel of a PMD19K3 ToF sensor is shown in figure 2.5. The propagation of the clock tree from the top of the sensor to the bottom is visible in the distance offset. The PMD19k3 sensor contains 120 by 160 pixels and has five regions in the sensor. The phase offset can be a function of temperature which can add complexity to the calibration process, and has lead to the inclusion of temperature sensors on the pixel array.

Camera calibration is a well researched area with a number of camera calibration toolboxes available. Two popular calibration methods are Zhang (2000) and Tsai (1987). Both calibration methods use a pinhole camera model and then correct the lens distortion. The small array size of ToF range cameras can cause problems for the calibration method. Once the intrinsic and extrinsic camera parameters are known the camera point $[w, v, d]$, where w, and v are the pixel coordinates and d is the radial distance can be transformed to the world coordinate $[X, Y, Z]$. The generated point clouds are ready for use by the application.

2.3 Measurement Errors

A number of factors contribute to the accuracy and precision of ToF range cameras. Accuracy is the closeness of the measurement with the ground truth, while precision is the distribution of random errors. In this section the measurement errors and noise sources in ToF range cameras are overviewed and

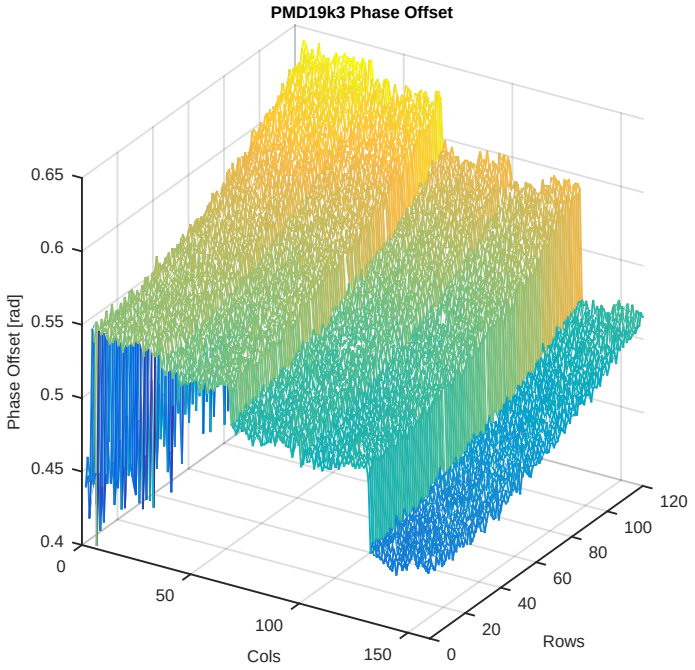


Figure 2.5: Phase offset of each pixel in a PMD19k-3 based time-of-flight range camera operating at 50 MHz. The five different regions of the sensor are clearly visible, along with the clock propagation through the pixel array.

state of the art correction techniques reviewed. The measurement errors affect the measurement accuracy while the noise sources affect the precision.

2.3.1 Multiple Propagation Paths

A major cause of measurement errors in ToF range cameras is multiple propagation paths between the light source and pixel, and this error is called multi-path interference. Multi-path interference is highly scene dependent and can be caused by: inter-reflections, subsurface scattering, volumetric scattering, translucent objects and mixed pixels. Inter-reflections are caused by light reflected from one object striking other objects in the surrounding area. Inter-reflections are commonly caused by corners, as demonstrated in figure 2.6, and lens flare. Subsurface scattering occurs when the light enters the object and is scattered inside the object before emerging. Sub-surface scattering occurs on objects such as wax and biological tissue. Volumetric scattering is scattering of light by particles in the atmosphere, and is present in environments as dust, smoke, and fog. Translucent objects, such as glass, transmit a portion of the light and reflect a portion. The diagram in figure 2.7 demonstrates how

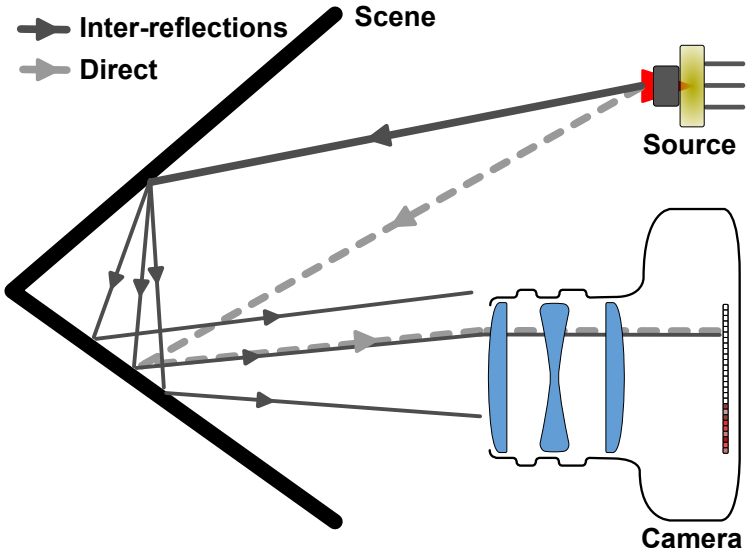


Figure 2.6: The measured signal in a corner is the combination of the direct return and the multiple propagation paths caused by inter-reflections in the corner.

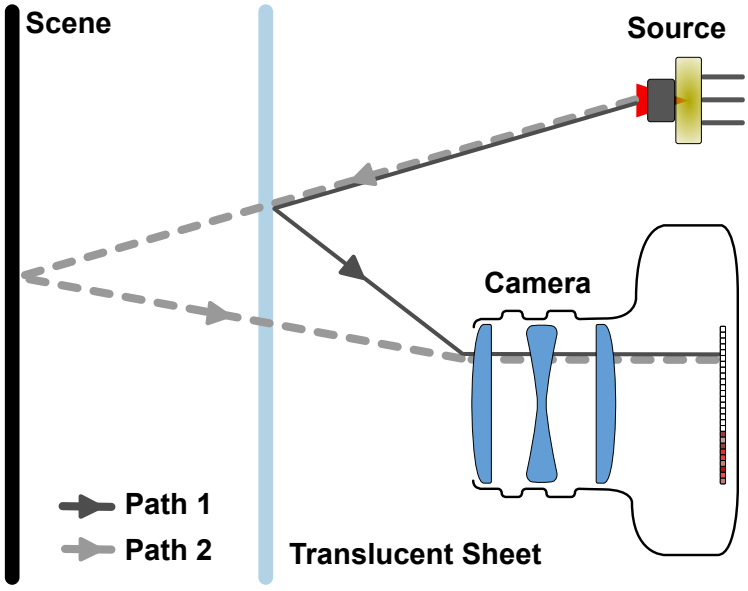


Figure 2.7: The measured signal is the combination of the return off the translucent sheet and the return behind the translucent sheet from the scene.

multi-path interference is caused by a translucent sheet. Mixed pixels occur when one pixel is a combination of two objects at different distances, and is exacerbated by defocused objects.

Currently two forward models are used to describe multi-path interference in ToF range cameras. One model assumes diffuse scattering (Jiménez *et al.*, 2012), (Fuchs *et al.*, 2013), (Naik *et al.*, 2015), the other a sparse number of

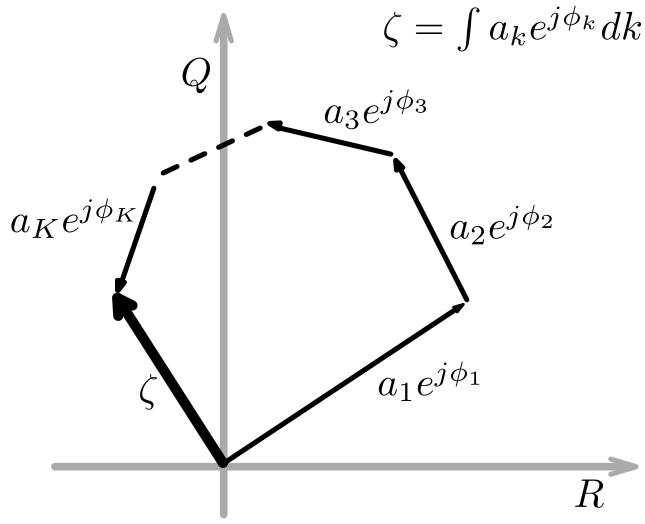


Figure 2.8: The measured vector ζ is the result of the contour integral on the complex plane.

discrete returns (Godbaz *et al.*, 2012), (Kirmani *et al.*, 2013), (Freedman *et al.*, 2014). In the sparse forward model the multi-path interference is expressed as the sum of K complex exponentials such that

$$\zeta = \sum_{k=1}^{K} a_k e^{j\phi_k}. \tag{2.19}$$

The measured complex value is the sum on the complex plane as demonstrated in figure 2.8. An infinite combination of vectors is possible that sum to the measured vector. The infinite number of combinations makes resolving multi-path interference a difficult problem to solve as it is a trapdoor function. A trapdoor function is a function that is easy to compute in one direction but is difficult to compute the inverse without special information.

The diffuse forward model expresses the return as one direct return and the sum of everything else at a further distance. This is expressed as

$$\zeta = a_1 e^{j\phi_1} + \int_{d_1}^{\infty} a_k e^{j\phi_k} dk. \tag{2.20}$$

In reality a combination of both forward models better describes the combination of diffuse and specular multi-path interference than the individual models in equations 2.19 and 2.20.

As plotted in figure 2.8 the measured value is the integral on the complex plane. When only one return is present the amplitude is constant over

all modulation frequencies and the phase increases linearly with frequency (baring phase wrapping). When multi-path interference is present the relationship between modulation frequency and measured phase and amplitude is non-linear and techniques that take measurements at multiple modulation frequencies focus on the inversion of this non-linear relationship. Dorrington *et al.* (2011) demonstrated resolved multi-path interference on two commercial ToF range cameras, the Mesa Imaging SR4000 and the Canesta XY-422 Demonstrator, by taking measurements at two modulation frequencies. Dorrington *et al.* (2011) assumed two propagation paths, therefore there are four unknown variables, each paths' amplitude and phase, the unknown variables were solved by iteratively minimizing the objective function

$$\underset{a_1,a_2,\phi_1,\phi_2}{\arg\min} = |\zeta_1 - \hat{\zeta_1}|^2 + |\zeta_2 - \hat{\zeta_2}|^2, \tag{2.21}$$

where ζ_1 and ζ_2 are the complex measurements at modulation frequency f_1 and f_2 respectively and $\hat{\zeta}$ is the forward estimate of the measurement. While Dorrington *et al.* (2011) was able to demonstrate removal of multi-path interference, each pixel had to iteratively converge to the solution, which is computationally expensive and not possible to solve for the entire imaging array at 30 fps of depth information. Godbaz (2012) also sampled two modulation frequencies and assumed two propagation paths and solved the speed issue from Dorrington *et al.* (2011) by using a two dimensional lookup table. Along one dimension is the ratio of the measured amplitudes a_1/a_2 and the other dimension is the difference between the phases while accounting for the frequency difference. The other constraint is the second modulation frequency has to be double the first. Freedman *et al.* (2014) sampled three modulation frequencies and created a compressed lookup table to resolve multi-path interference on the Xbox One Kinect camera. The reason for the compressed lookup table is a three dimensional lookup table occupies an large amount of memory, Freedman *et al.* (2014) noted that there was redundancy in one of the frequency measurements so two measurements could be combined resulting a two dimensional lookup table, but had improved results over Godbaz (2012). With the assumption of two propagation paths Godbaz (2012) found a closed form solution for the amplitude and phase of each return when four evenly spaced modulation frequencies are sampled. Kirmani *et al.* (2013) used measurements at five evenly spaced frequencies and used robust spectral estimation with low computational complexity to detect and resolve multi-path

interference and perform phase unwrapping. Kirmani *et al.* (2013) showed results based on simulations but did not test the presented method on real data. Bhandari *et al.* (2014) generalized the solution for $M > 4$ modulation frequencies that are evenly sampled and for K returns, this expanded on the previous assumption of two propagation paths. Using $K = 3$ for the number of propagation paths Bhandari *et al.* (2014) was able to image through two translucent sheets. The previous reviewed multi-frequency techniques for solving multi-path interference have measured the phase at each frequency, an alternative approach is to keep the phase offset constant and only change the frequency. This approach is similar to stepped frequency continuous wave (SFCW) radar, and early in the development of ToF range cameras Gulden *et al.* (2002) applied SFCW so the ToF could operate as a level gauge, but was limited by the frequency bandwidth of the technology at the time. Whyte *et al.* (2015b) applied known techniques from radar to ToF range cameras, in particular stepped frequency continuous wave (SFCW) to make measurements that were resilient to multi-path interference and phase wrapping. Whyte *et al.* (2015b) took 15 measurements from 10 MHz in increments of 10 MHz up to 150 MHz and applied the MUSIC spectrum estimation technique to measure the depth and showed higher harmonics and multi-path interference showed up at a higher frequency than the direct (shortest) return. Whyte *et al.* (2015b) experimentally showed that 6 measurements was the minimum for accurate depth measurement,but the systematic accuracy was still worse than AMCW measurements.

Fundamental to ToF range cameras is correlation between the returning optical signal and a reference signal as given in equation 2.4. Use of sparse deconvolution on the correlation signal was shown by Godbaz (2012) to resolve multi-path interference. Godbaz (2012) showed the results were limited by the lack of spectral content in the correlation signal as the correlation signal compressed of the first and third harmonic, and the third harmonic was one ninth of the amplitude of the first. Kadambi *et al.* (2013) increased the spectral content of the correlation signal by replacing the square wave modulation with a binary sequence which improved the results of the sparse deconvolution to recover the multiple propagation paths, assuming a discrete number of returns.

In computer graphics light transport is an important field of study for photo-realistic rendering. The light transport in a scene can be represented by a direct component, which is the shortest path between the light source and camera, and the global component which encompasses everything else. Separ-

ating the light transport components into the direct and global returns using projected patterns was shown by Nayar *et al.* (2006). Nayar *et al.* showed that when illuminating the scene with a checkerboard pattern in the non-illuminated regions only the global component was present and in the illuminated regions the direct and global components were present. When the pattern is inverted there are two measurements per pixel and two unknowns. The primary constraint is the global component has to have a lower spatial frequency than the projected checkerboard pattern. In theory it is possible to use two pattern in reality due to mis-alignment between the camera and projector optics means the pixels on the edge of the pattern are not resolved. Nayar *et al.* (2006) proposed using 25 patterns for the best results. Naik *et al.* (2015) applied direct and global separation to ToF range cameras using checkerboard patterns and compared the results to ground truth data collected with a lidar point scanner. The method presented by Naik *et al.* (2015) required 25 depth frames to operate. Further constraints to direct and global separation on the complex data produced by ToF cameras when using checkerboard illumination patterns were added by Whyte *et al.* (2015a) as the data is complex the addition is not always constructive unlike radiance which is always constructive. Whyte *et al.* (2015a) mixed the sinusoidal projected pattern with the correlation signal of ToF cameras for a closed form solution to diffuse multi-path interference that required nine raw frames.

Iterative computational approaches and filtering have been used to reduce multi-path interference. These approaches do not take additional measurements like the above methods, instead they make assumptions (a prior) about the multi-path interference or use scene preparation. Lens flare is a cause of multi-path interference identified by Mure-Dubois and Hügli (2007), and to correct this error the complex point spread function (PSF) of the lens was estimated. The complex PSF is the expansion of the PSF to include the phase in the coupling, and the multi-path interference was removed by deconvolving the measurement with the complex PSF. Instead of pre-measuring the PSF Falie and Buzuloiu (2008) estimated the lens flare by placing two tags of known reflectivity in the scene at the same distance and using the difference in amplitude and phase from the ground truth to estimate the PSF and then apply the correction. However the issue with using a constant PSF over the entire camera was demonstrated by Godbaz (2012) who measured the PSF of a ToF range camera and showed the PSF is non-localised and dependent on the location of the light source. Fuchs *et al.* (2013), and Jiménez *et al.* (2012) both posed

iterative optimization problems constrained by the measured data to perform inverse ray tracing. Both these methods assume the only mode of multi-path interference present is lambertian inter-reflections in the scene and estimate the surface normal and albedo of the material. While these approaches work in controlled environments with only lambertian inter-reflections they are computationally expensive. Larkins *et al.* (2009) applied the normal angle filter to detect mixed pixels and correct them to the foreground or background object. The normal angle filter operates on calculated the surface normal of each pixel, if the normal approaches perpendicular to the camera then the pixel is marked as mixed.

Table 2.1 summarizes the previous work for resolving multi-path interference. This thesis is the major contributor to the methods (Kadambi *et al.*, 2013), (Whyte *et al.*, 2015b), (Whyte *et al.*, 2015a), and a contributor to (Bhandari *et al.*, 2014).

2.3.2 Computational Imaging

Measuring the propagation of light in a scene is of interest in computer graphics and optical engineering to verify ray tracing for optical engineering (lens designs etc.) and for photo realistic rendering. Transient imaging is the process of measuring the light propagation through the scene at small time intervals, which is a form of high speed photography. While CMOS and CCD sensors have allowed for frame rates up to 10^7 frames per second they are limited by their readout time and memory storage capacities. Frames rates of 10^7 are extremely high compared to the human vision system but are still orders of magnitude below capturing light in flight. The capture of light in flight, known as transient imaging, measures the x,y,t of light, where x and y are the spatial location and t is the time interval. The smaller the time interval the higher the temporal measurement of the measurement.

Velten *et al.* (2013) coined the term femto-photography to visualize the propagation of light by using a streak camera and femto-second laser to illuminate the scene. A streak camera varies the electromagnetic field to transform a temporal signal into a spatial signal across the imaging sensor, by synchronizing the 50 fs laser pulse with the streak camera the light propagation for a single point can be visualized. To obtain a reasonable signal to noise ratio Velten *et al.* (2013) averaged a number of light pulses, so only static scene measurements were possible. Micro-mirrors were used to raster scan across the scene to measure the x and y spatial information. This equipment setup

Table 2.1: Comparison of multi-path interference correction techniques.

Paper	**Model**	**Computation**	**Constraints**
Dorrington *et al.* (2011)	Two-path	High	Two modulation frequencies
Godbaz (2012)	Two-path	Low	Two or three modulation frequencies
Freedman *et al.* (2014)	Sparse	Low	Three modulation frequencies
Kirmani *et al.* (2013)	Two-paths	real-time	Five modulation frequencies
Kadambi *et al.* (2013)	K-paths	High	Large number of phase steps
Bhandari *et al.* (2014)	K-paths	Medium	Five or more evenly spaced frequencies
Fuchs *et al.* (2013)	Diffuse	High	
Jiménez *et al.* (2012)	Diffuse	"Minutes"	
Falie and Buzuloiu (2008)	Diffuse	Low	Scene preparation required
Mure-Dubois and Hügli (2007)	Diffuse	"Real Time"	
Whyte *et al.* (2015a)	Diffuse	Low	Projected Patterns Required
O'Toole *et al.* (2014a)	Combined	High	5D measurements, phase, frequency and pattern.
Whyte *et al.* (2015b)	Sparse	Medium	Six or more samples.

allowed Velten *et al.* (2013) to record the propagation of light through a plastic bottle and a tomato.

Avalanche photo-diodes can be manufactured into an imaging sensor, commonly called SPAD arrays. Gariepy *et al.* (2015) used a 32 by 32 pixel array to image Rayleigh scattering of light in flight through air. Each pixel in the array contains a pico-second timer and the time of each arriving photon is recorded and using this information and sensor calibration the propagation of light in flight is recorded. The use of an imaging array and a high repetition laser (in the MHz) would allow for a capture in the order of one second. While SPAD arrays achieve high temporal resolution for transient imaging they remain expensive.

Gao *et al.* (2014) developed compressed ultra-fast photography (CUP) that observed events with a temporal resolution in the tens of picoseconds. This was achieved by coupling the streak camera with a micro-mirror array (DMD). The streak camera observed the event, and the DMD spatially encoded the incoming light. The use of computation, similar to that of compressed sensing based image restoration, was used to reconstruct the 2D image from the streak camera output. These adaptations allowed for imaging of light interactions at boundaries between different gases.

Wu *et al.* (2014) used the streak camera setup from Velten *et al.* (2013) to measure the time profile of light in a scene using direct measurements. Wu *et al.* (2014) showed the time profile could be represented by Gaussian function for the direct return and for inter-reflections, and as a decaying exponential for subsurface scattering measurements. This understanding of the propagation of light and potential basses for representation allow for better indirect measurements that exploit computational models.

Computational imaging is the combination of imaging sensors and computational models to recover information from the scene. With the increase in computational power, following Moore's law, computational imaging is a growing area of research and producing consumer imaging products, such as Lytro's light field camera, Pelican imaging image sensor, and Light Co's new high performance camera. The use of computational models has allowed image recovery from a limited number of samples using a single pixel camera (Duarte *et al.*, 2008) and the recovery of images placed behind scattering medium (Tajahuerce *et al.*, 2014). Computational techniques have been combined with time of flight measurements to gain additional information about the scene. Velten *et al.* (2012) used a streak camera and pulsed laser (same setup as Vel-

ten *et al.* (2013)) to image around a diffuse corner. By combing the streak camera measurements with the measurements at different spatial locations and a model of diffuse of the wall, the objects behind the wall were able to be reconstructed.

The above mentioned techniques for transient imaging require expensive equipment, both streak cameras and femto-second lasers are expensive and difficult to operate. This motivated the usage of ToF imaging sensors for transient imaging applications. Heide *et al.* (2013) used a modified PMD sensor based camera to perform transient imaging, primarily of a disco ball in a corner. By taking measurements from 10 MHz up to 120 MHz in increments of 0.5 MHz and that capture took 90 s. By solving a convex optimization problem the time profile of each pixel is recovered. The temporal resolution of the transient image is proportional to the highest modulation frequency of the ToF range camera.

The use of computational imaging has been applied to ToF range cameras. Heide *et al.* (2014b) measured the distance to objects in a 80 litre vat of dilute milk by taking measurements at three modulation frequencies. In a scattering medium the multiple propagation paths are no longer diffuse, to overcome this Heide *et al.* (2014b) demonstrated the returns were sparse when represented with a Gaussian basis. Using this information an optimization problem can be solved to recover the depth of objects in a scattering medium. The problem of looking around a corner can be solved with a ToF range camera as demonstrated by Heide *et al.* (2014a) and Kadambi *et al.* (2015). Unlike traditional ToF range imaging where the scene is flood light Heide *et al.* (2014a) focused the light source onto a point on the diffuse wall and measured the resulting inter-reflections back onto the wall. The difference between Velten *et al.* (2012) and Heide *et al.* (2014a) is the size of the objects behind the corner that could be recovered, (Velten *et al.*, 2012) could recover objects on the order of centimeters while (Heide *et al.*, 2014a) could recover objects in the order of meters.

O'Toole *et al.* (2014a) illuminated the scene with temporal patterns to resolve the propagation paths into the direct, indirect and specular retro-reflective using the scene probed techniques previously developed by O'Toole *et al.* (2014b). The scene probing allowed for high resolution transient imaging and distinguished between a direct view and mirror reflections.

Lin *et al.* (2014) applied Fourier analysis to recover transient images from the dataset produced by Heide *et al.* (2013). By applying the inverse discrete

Fourier transform to the complex data the time profile can be recovered. However the missing frequency content from the experimental data caused blurring in the time resolution, to alleviate this problem the missing frequencies are interpolated. Further post processing is performed to enhance the results to be closer to the ground truth measurements.

An alternative approach was taken by Kadambi *et al.* (2013). The illumination signal of the ToF was changed from a square wave to a binary sequence, which changed the measured correlation signal from a sinusoidal signal to a triangular peak and constant everywhere else. Then by applying Tikhonov regularization the correlation was transformed into signal that emulated the time profile of each pixel.

Recently Peters *et al.* (2015) measured three modulation frequencies for fast transient imaging and removal of multi-path interference. This is achieved by spectral estimation using a modification of the Burg method (Burg, 1967). Peters *et al.* (2015) removed the problem of harmonic aliasing from the measurement by careful sampling of the correlation signal. The fast transient imaging presented by Peters *et al.* (2015) has strong cross over to resolving multi-path interference.

2.3.3 Motion Artifacts

ToF range cameras like conventional cameras suffer from measurement artifacts caused by motion. ToF cameras are especially susceptible to motion because they are a multi-shot depth measurement system. Motion can be categorized into transverse and longitudinal motion, with transverse motion describing motion in the $[x, y]$ dimension and longitudinal motion in the z dimension.

Optical flow is a motion correction technique that attempts to match pixels between temporal frames and was developed by Horn and Schunck (1981). Optical flow was adapted to work with ToF range cameras by Lindner and Kolb (2009). The optical flow was performed on the common mode images as they are the closest image type to conventional cameras. However not all ToF cameras have access to the common mode image. Lee *et al.* (2012) detected motion between raw frames and replaced the pixels affected by motion with the nearest neighbor. Coded exposure is a motion correction technique that works by opening and closing the camera's shutter during the integration period to properly condition the deconvolution that removes motion (Raskar *et al.*, 2006). Coded exposure was adapted to ToF range cameras by Streeter and Dorrington (2014).

The previous methods of solving motion artifacts have focused on their detection and correction, another approach is to measure the motion directly. Heide *et al.* (2015) operated a ToF camera in heterodyne mode to measure the range and velocity simultaneously. The velocity of the object causes a Doppler shift in the reflection and because of the heterodyne modality of the camera the change in frequency can be measured. Whyte *et al.* (2015b) measured only the velocity of the object with a method analogous to continuous wave Doppler radars by sampling at a constant frequency and phase offset. If the object moved by one ambiguity interval, see equation 2.22, then the output is one period of a sine wave and the time it takes for one interval encodes the velocity of the object.

2.3.4 Distance Ambiguity

Distance ambiguity occurs when the phase of the reflected light's modulation envelop wraps from 2π back to 0. Objects further than the ambiguous distance are measured closer than they actually are. The maximum distance that can be measured before phase wrapping occurs is called the ambiguity distance, d_{amg}, and is defined by

$$d_{amg} = \frac{\pi c}{\omega}. \tag{2.22}$$

As the modulation frequency increases the ambiguity distance decreases. At 50 MHz the ambiguity distance is 3 m. Phase unwrapping techniques fall into two broad areas: using multiple modulation frequencies, and using the spatial information of the sensor to estimate the ambiguity interval. These techniques can be combined for robust phase unwrapping in the presence of noise.

To extend the ambiguity distance Jongenelen (2011) combined measurements at two different modulation frequencies using a variation of the new chinese remainder theorem (Wang, 2000). A visualisation of using two frequencies for phase unwrapping is in figure 2.9. Using three modulation frequencies to perform phase unwrapping was detailed in a patent filed by Microsoft Corp (Benedetti *et al.*, 2014). It was disclosed by Bamji *et al.* (2015) that three frequencies were used in the Xbox Kinect for phase unwrapping. By matching the possible ambiguity brackets with the three frequency measurements phase unwrapping is possible. Using generalized approximate message passing (GAMP) and two modulation frequencies Mei *et al.* (2013) performed phase unwrapping and noise reduction on ToF images. The use of multiple frequen-

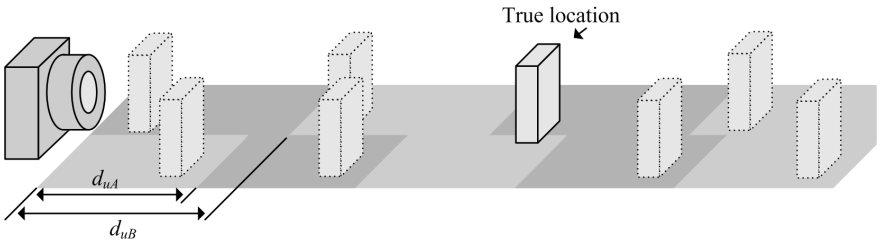


Figure 2.9: Phase unwrapping with two measurement frequencies. At the true location the phase values for both measurements align. Figure reproduced with permission from (Jongenelen, 2011).

cies with an iterative optimization framework for consistent phase unwrapping was presented by Choi *et al.* (2013).

Spatial phase unwrapping uses the spatial information from the sensor to perform phase unwrapping, which also includes statistical information and path walking for image phase unwrapping. McClure *et al.* (2010) segmented the objects in the scene and then estimated the ambiguity bracket of each object. The use of a graphical model with loopy belief propagation was applied to phase unwrapping of ToF images by Droeschel *et al.* (2010). Crabb and Manduchi (2014) used probabilistic models and spatial information for phase unwrapping on a single frequency ToF measurement.

2.3.5 Harmonic Aliasing

Harmonic aliasing is a measurement error caused by the implementation using square waves instead of sinusoidal signals, or by using distorted sinusoidal signals. As shown in equation 2.9 there are odd harmonics present in the correlation signal. If the number of samples on the correlation signal is less than twice the highest frequency present in the correlation signal then aliasing will occur. In the case of $N = 4$ the higher order harmonics alias onto the first harmonic and cause measurement errors. The measured error caused by harmonic aliasing is plotted in figure 2.10. The error is measured with the ToF camera described in chapter 3, operating at 50 MHz. The error bars are plus and minus one standard deviation, and the standard deviation of the distance measurement increases with distance because as the object moves further away less light returns to the camera, as observed in figure 2.10.

Payne *et al.* (2008) modified the camera's operation to phase shift the reference signal during the integration period. Payne's modification cancels the third harmonic however it decreases the precision of the measurement.

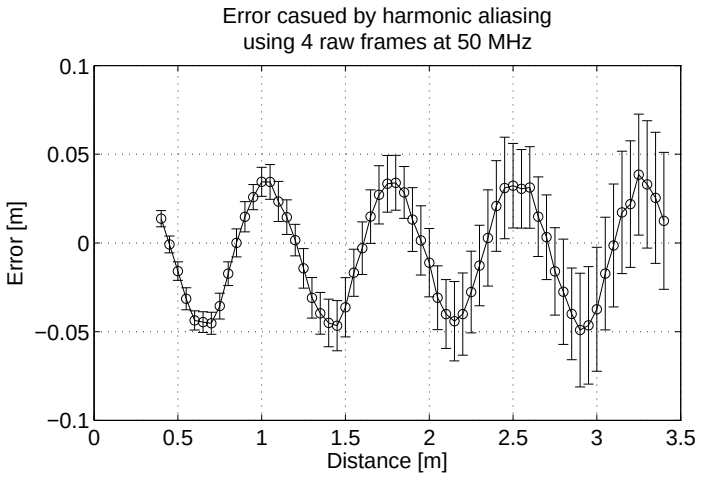


Figure 2.10: Error caused by harmonic aliasing as measured with four raw frames at a modulation frequency of 50 MHz using the camera described in chapter 3. The error bars are one standard deviation, and they increase with distance because the measured object is further away so less light is returning to the camera resulting in a decreased signal to noise ratio.

Jongenelen (2011) showed that using three samples, instead of four, reduces the error caused by aliasing because the third harmonic did not alias directly onto the first harmonic.

A number of calibration methods have been used to correct for harmonic aliasing. Kahlmann *et al.* (2006) used a lookup table to calibrate for harmonic aliasing. Lindner and Kolb (2006) fitted cubic B-splines to the error for calibration. When other errors are present the harmonic calibration will introduce further errors.

2.4 Noise

A variety of noise sources exist in ToF range cameras. These include photon shot noise, amplifier noise, ADC quantization, crosstalk, clock jitter, reset noise, and dark current. These noise sources affect the signal to noise ratio of the range measurement.

Photon shot noise is present in every CMOS imaging system and is well quantified. It is governed by the probability of the generation of a photon and is a function of the number of photons generated. Amplifier noise and ADC quantization are noise sources introduced during the sensor readout. The standard deviation of the phase measurement, σ_ϕ, is inversely proportional to

the signal to noise ratio (SNR) of the pixel measurement

$$\sigma_\phi = \frac{1}{2m\sqrt{SNR}}, \tag{2.23}$$

where m is the modulation index which is the difference between the maximum and minimum levels of the modulation signal. For solid state image sensors the demodulation contrast, c_d, is a similar measure and calculated as

$$c_d = \frac{a}{b}, \tag{2.24}$$

where a is the measured demodulated pixel amplitude and b is the DC offset. The demodulation contrast depends on a number of factors, namely the modulation frequency, illumination wavelength, and the imaging sensor design (Lange and Seitz, 2001). The phase standard deviation, ϕ_{sigma}, can be estimated by (Büttgen *et al.*, 2005)

$$\sigma_\phi = \frac{\sqrt{b}}{\sqrt{2}c_d a}. \tag{2.25}$$

The distance standard deviation σ_d is

$$\sigma_d = \frac{c\sigma_\phi}{2\omega}. \tag{2.26}$$

Increasing the modulation frequency decreases the distance variance. The other options are improving the demodulation contrast by increasing the light power or decreasing the DC component which increases the number of useful photons collected by the imaging sensor.

Jitter is the undesired random deviation from the true frequency in a signal. The random deviation between the reference signal and optical signal is a cause of uncertainty in ToF range cameras. The impact of cycle to cycle jitter in ToF cameras was investigated by Seiter *et al.* (2014). Streeter *et al.* (2013) demonstrated jitter is a significant contributor to uncertainty in phase and presented a variety of image processing filters to reduce the uncertainty caused by jitter. Phase noise analysis is derived from the analysis of jitter in digital clocks. The phase noise, $\theta(t)$, can be expressed in the sinusoid reference signal as

$$g(t) = \sin(\omega t + \theta(t)). \tag{2.27}$$

The phase noise is a random variable that is zero mean. When the reference signal is correlated with the optical signal the measured correlation is

$$h(\tau) = (g \star s)(t) = a\cos(\omega\tau + (1 + \Delta f)\phi), \tag{2.28}$$

where Δf is the error in modulation frequency due to signal jitter. The phase noise due to jitter increases linearly with amplitude, therefore increasing the illumination power does not affect the relative jitter. Streeter *et al.* (2013) showed the minimum phase variance achievable is limited by the jitter. The phase noise can be converted to a jitter period by

$$J_{PER} = \frac{\theta(t)}{\omega}. \tag{2.29}$$

The jitter period is often constant in the camera while the phase noise is a function of modulation frequency.

2.4.1 Other Noise Sources

The pixel uses capacitors to store charge during the integration period and capacitors have an inherent reset noise (Texas Instruments, 2014). Correlated double sampling can be used to remove the pixel's offset voltage.

With increases in temperature a larger number of thermally generated electrons are present in CMOS imaging sensors. Along with the increase in noise the phase offset calibration is temperature dependent. The effect of temperature on ToF range cameras has been characterized by Kahlmann (2007).

Signal integrity is important in ToF range cameras as they operate in the megahertz range. Cross talk between signals can negatively affect performance by limiting the maximum modulation frequency and can add non-linearities in the modulation signal which can cause measurement errors.

Chapter 3

Time-of-Flight Range Camera Design

"He who controls the spice controls the universe." Frank Herbert, Dune.

This quote could also be said of data, where access and control of data is playing an ever important role in society. For this research a flexible configurable time-of-flight (ToF) range camera was needed to develop novel methods of resolving measurement errors. Current commercial ToF range cameras limit the access to low level data, and are inflexible, making them unsuitable for in-depth systems research. Having access to the entire processing pipeline from raw data to depth, enables the development of novel algorithms that tightly couple the hardware and image processing. The following chapter details the design, implementation and testing of the custom ToF camera developed as part of this research. Recently Texas Instruments® have developed ToF range imaging solutions, providing a sensor and processing integrated circuit (IC) (Texas Instruments, 2014), lowering the barrier to entry for custom ToF range cameras.

This chapter is divided into five sections: hardware design, field programable gate array (FPGA) design, the Matlab interface, system bring up, and camera characterization. The main design requirements of the camera are:

- Access to the lowest level of data possible.
- Dynamically configurable phase offset and modulation frequency.
- Configurable binary sequence for modulation.
- Configurable integration period.
- Robust and reliable operation.

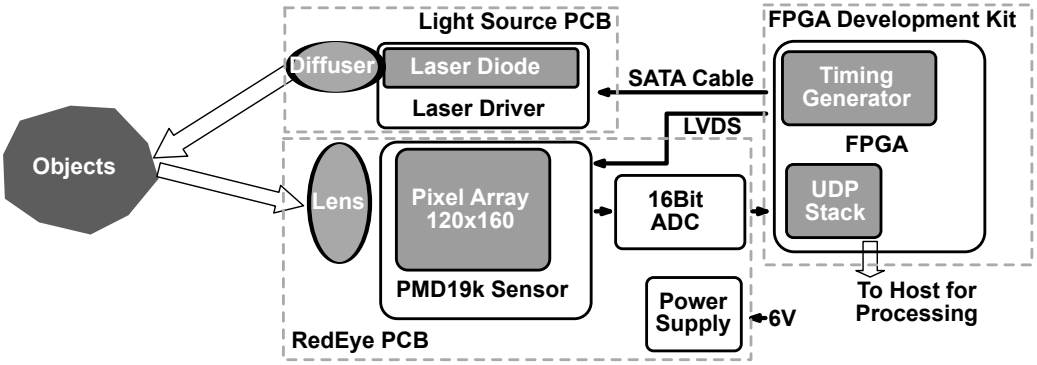


Figure 3.1: Overview of a time-of-flight range camera design.

A block diagram giving an overview of the ToF camera design is in figure 3.1. The hardware consists of the light source illumination, and the sensor printed circuit board (PCB) called RedEye. The FPGA design provides connectivity to the outside world and precise control over modulation of the pixel array. The host computer provides a Matlab interface for ease of configuration and processing of the collected data.

3.1 Hardware Development

The hardware consists of two custom PCBs, and their mechanical mounting. The PCBs encompass the ToF sensor, analog to digital converter (ADC), signal conditioning, and light source. The mechanics consists of the light source diffuser, lens mounting, and connecting to standard optical components. The hardware of the camera design can be broken down into two revisions the PMD19k2 and the PMD19k3 version. The purpose of the revision was to update the design in order to use an updated ToF sensor.

The camera has three PCBs: the light source, RedEye which contains the ToF sensor and ADC, and the FPGA development kit. The light source and RedEye PCBs are custom made and the FPGA development kit is off-the-shelf.

3.1.1 PMD19k2 Version

The original PCB design provided by Jongenelen (2011) had issues with the high speed mezzanine card (HSMC) connector, and the ADC power supply. Due to the HSMC connector being improperly soldered the ground plane floated on the RedEye PCB causing damage to the modulation drivers, resulting in a PCB redesign for this research.

The RedEye PCB contains the ToF pixel array, the ADC, modulation drives, power supplies, and signal end to LVDS converters. The RedEye PCB has four copper layers to maintain signal integrity. Due to the high speed modulation, cross-talk and ringing on the PCB traces can be an issue, therefore short traces are used between the modulation drivers and the ToF sensor. Sufficient decoupling capacitors with minimum ground loops are part of the ADC layout design. The decoupling capacitors maintain a constant voltage to the ADC which reduces readout noise.

ToF pixels are actively driven in the megahertz range. As the PMD19k2 modulation drivers are off chip, eight external Intersil EL7158 (Intersil, 2007) drivers are used to switch the A and B taps of the sensor. The sensor is divided into four regions, and a separate driver is used for each region to decrease the capacitive load on each driver. The -3 dB point of the Intersil EL7158 is 40 MHz which limits the maximum modulation frequency of the ToF camera.

The output of each pixel from the imaging array is a voltage which encodes the number of photons collected. The analogue voltage is converted to a 16 bit digital value by the Analog Devices AD9826 ADC (Analog Devices, 2012). The RedEye PCB physically connects to the FPGA development board with three 40-pin IDC connectors.

The Stratix III development board (Altera, 2008a) was used as the FPGA development kit. The development board contains the required RAM, I/O, and Ethernet conductivity for the camera development. At the start of the research project the Stratix III FPGA (Altera, 2011) was the top of the line FPGA from Altera, having sufficient logic elements for the camera design and imaging processing. The FPGA has dynamically configurable phase lock loops (PLLs), which in 2009 Xilinx FPGAs did not have. Dynamically reconfigurable PLLs are required for controlling the modulation frequency and phase.

Low-voltage differential signalling (LVDS) was used as the physical layer communication standard between the FPGA and external PCBs. At the used signal speeds and PCB trace lengths, the transmission line properties are important. serial ATA (SATA) cables were used to communicate between the FPGA and light source, as they are low cost, sufficient length, and impedance matched. These features allow for flexible light source placement.

3.1.2 PMD19k3 Version

During the course of the Ph.D a newer camera version was developed, which upgrades the FPGA development kit and the ToF sensor. The new sensor is

the PMD19k3 ToF pixel array from PMD Technologies. The sensor contains 120 by 160 pixels, and has a maximum operating frequency of 80 MHz. Due to confidentiality additional information about the sensor is not provided herein. The new RedEye PCB connects to the DE2-115 development board (Terasic Technologies, 2013) with two 40 pin IDC connectors. The DE2-115 development board contains a Cyclone IV E FPGA (Altera, 2013a), 16 MB of SDRAM, sufficient I/O and a 1 Gbps Etherent PHY, and all the additional parts of the camera. Using knowledge gained from the previous revision a smaller and cheaper FPGA was sufficient for the design.

3.1.3 Light Source Design

The light source is a critical part of the ToF camera design. If the light sources waveform is malformed then the phase and amplitude can become non-linear with distance, as demonstrated by Payne *et al.* (2011). Electrical cross talk and optical coupling between the light source and sensor also introduces measurement errors.

The light source PCB is designed to drive two laser diodes to be driven independently. Photographs of the light source design are in figures 3.3 and 3.2. The iC-HK driver (iC Haus, 2009) laser driver was used as it can modulate up to 155 MHz, and switch up to 700 mA per channel. The PCB was designed to fit a variety of 5.6 mm diameter laser diode cases, intended for three laser diodes in the design: the HL6545MG (Opnext, 2008) 660 nm 120 mW diode, D405-120 (US Lasers, 2013) 405nm 120 mW diode, and L850P010 (Thorlabs, 2010) 850 nm 10 mW laser diode. The different wavelengths and powers are useful for fluorescence and subsurface scattering measurements and other applications. The mechanical design allows for one inch mounted ground glass diffusers to be placed in the optical path of the light source and the PCB be mounted to standard optical parts.

3.1.4 Mechanical Design

The FPGA development kit, sensor, lens, and light source are mounted onto an optical table. One error with the design is that no lens mounting holes were included, which limits the lens calibration accuracy as the lens is unstable. A photo of the mounted setup is in figure 3.4.

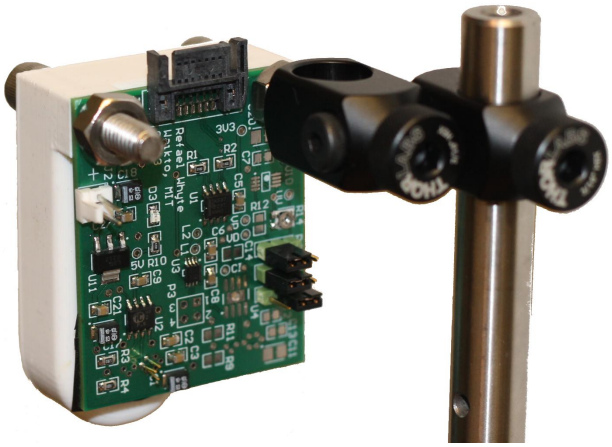


Figure 3.2: Photograph of back side of light source.

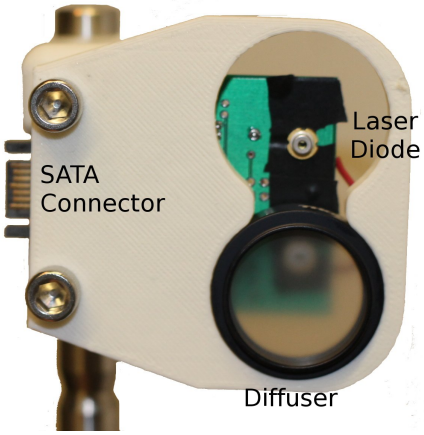


Figure 3.3: Photograph of front side of light source with one diffuser mounted.

3.1.5 Projector Integration

A spatial light modulator (SLM) was integrated into the light source to allow configurable projected patterns. The SLM used is a digital micro-mirror device (DMD) from Texas Instruments. DMDs contain an array of highly reflective aluminium micro-mirrors, each mirror in the array is electrostactaliy positioned to one of two selectable possible positions at different angles.

The projector integrated into the camera was a DLP LightCrafter Evaluation Module from Texas Instruments (Texas Instruments, 2013b). The projector used three LEDs and a 608 by 686 DMD array. The red LED was removed and replaced by a 660 nm laser diode driven by the described light source. The C++ API for the DLP LightCrafter[1] was interfaced to Matlab to control the projected image. A photograph of the integrated projector is

[1] `http://www.ti.com/tool/dlplightcrafter`

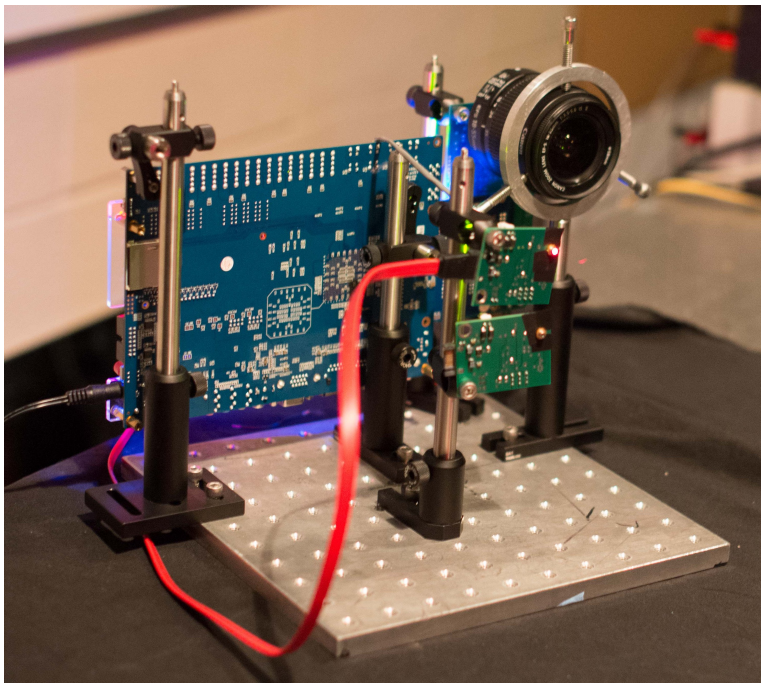


Figure 3.4: Photo of assembled PMD19k3 time-of-flight range camera with two operating light sources.

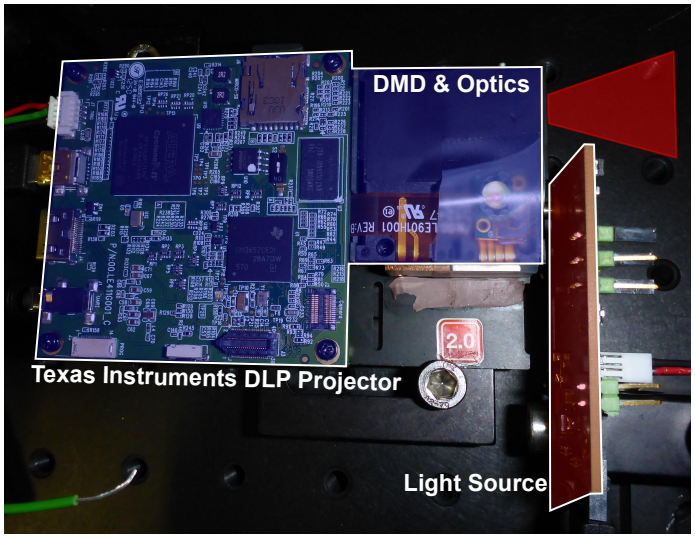


Figure 3.5: Photograph of integrated DLP LightCrafter projector. The original projector is labelled along with the mounted time-of-flight light source.

in figure 3.5. A photograph of a projected checkerboard pattern is in figure 3.6. The projector's field of view does not match the camera's, therefore some regions are not directly illuminated.

The luminance of each pixel of the DMD is controlled by the DLPC300 IC (Texas Instruments, 2013a). The DLPC300 takes an input 8 bit value for controlling the luminance of the pixel by controlling the duty cycle of the

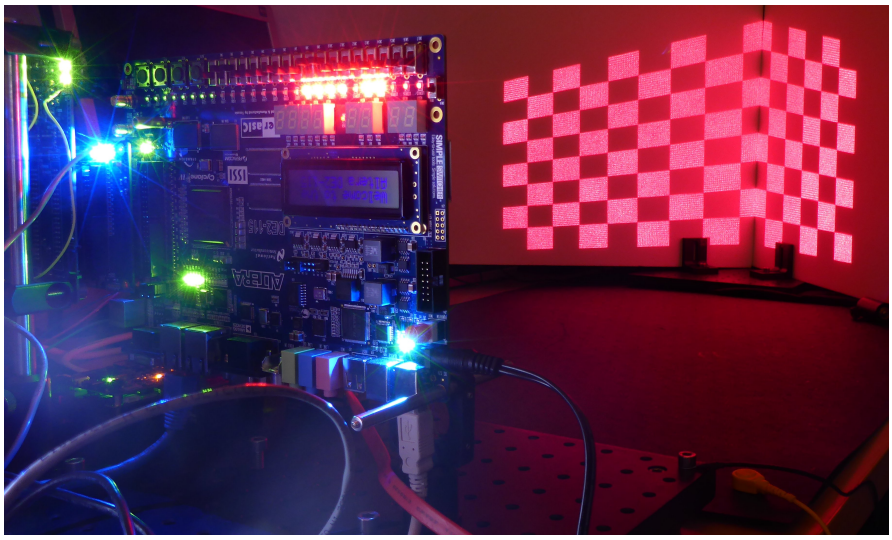


Figure 3.6: Photograph of projector illuminating a corner with a checkerboard pattern.

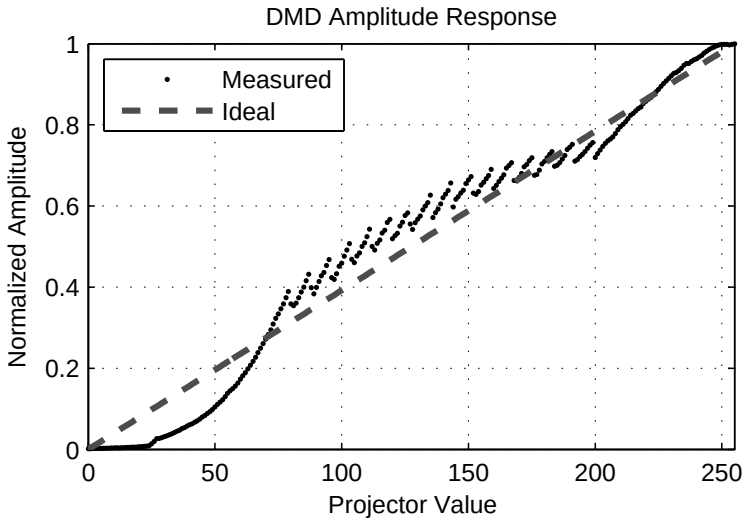


Figure 3.7: Amplitude response of the DLP projector plotted against ideal response.

pixel between the on and off states. However the 8 bit value is not linear with the amplitude measured with the ToF range camera. The measured amplitude compared to the ideal is plotted in figure 3.7. A lookup table is used to calibrate images in order to have a linear amplitude response, as a linear response is critical for some patterns, as demonstrated in chapter 7.

Component	Description
Nios II	A softcore microprocessor provided by Altera, used to implement a UDP and ARP stack for networking.
PMD19k Interface	The interface to the ADC and PMD19k ToF sensor. Controls readout and initial setup.
UDP Packetizer	Generates valid UDP packets from blocks of pixel values and streams them to the Ethernet core.
Modulation Controller	Generates the modulation signals for the sensor and light source, controlling the phase, binary sequence, and frequency of each signal.
TSE Controller	The Triple Speed Ethernet (TSE) controller provided by Altera (Altera, 2013b).

Table 3.1: Short descriptions of each component in the camera's FPGA design.

3.2 FPGA Design

FPGAs are semiconductors designed so that the digital logic can be reprogrammed in the field, which makes them ideal for ideal for implementing custom digital logic. The FPGA is used to control the modulation signals, readout process, and transmission of the camera data over Ethernet to a host computer. The major IP cores of the FPGA are shown in figure 3.8, and a short description of each component is in table 3.1. The Nios II and triple speed ethernet (TSE) are two IP cores provided by Altera are integrated into the FPGA design. The other IP cores are custom designed and written in VHDL.

Each component in the FPGA design is connected with the Avalon Interconnect standard, allowing for portability and reusability with Altera FPGAs. Use of standard interfaces allowed for a smooth transition from the Stratix III FPGA (Altera, 2011) to the Cyclone IV FPGA (Altera, 2013a).

The maximum sampling rate of the ADC is 15 MSPS and each sample is 16-bits therefore generating 240 Mbps of data. The bottleneck in data throughput is the sampling rate of the ADC, as the Ethernet is capable of transmitting 1 Gbps of data.

3.2.1 PMD19k Interface

The PMD19k Interface interfaces to the ToF sensor and ADC. Both the PMD and ADCs datasheets (Analog Devices, 2012) specify the correct readout interface. The datasheet for the PMD19k2 and PMD19k3 sensors are confidential, therefore details are not provided in this thesis.

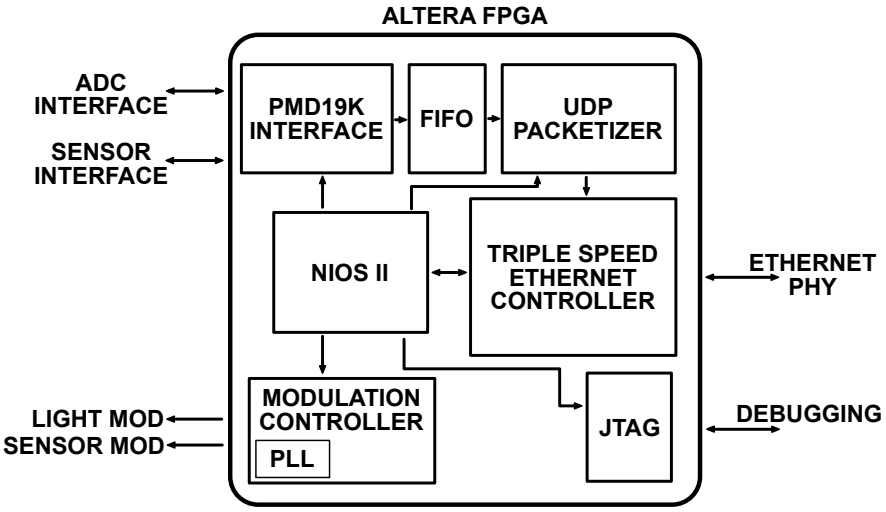


Figure 3.8: Internal block diagram of the FPGA design for the PMD19k time-of-flight range camera.

3.2.2 Modulation Controller

The modulation controller generates and controls the modulation signals for the light source and pixel array. The modulation controller component controls the phase, frequency, and binary sequence of the modulation signal. Fundamental to the modulation controller is the phase lock loop (PLL) in the Altera FPGA (Altera, 2008b). The output frequency and phase of the PLL is dynamically reconfigurable, and the modulation controller provides a wrapper around the PLL to update the phase and frequency between raw frames of the camera. This section describes the generation of the binary sequences and configuration of the PLL for the desired frequency and phase.

A key aspect of the research presented here is the modulation of the sensor and light source with custom binary sequences, as described in chapter 4. Figure 3.9 is the implementation for generating a binary sequence in the camera's FPGA. The length of binary sequences is configurable from 2 bits up to 2048 bits. The binary sequences are phase shifted by phase shifting the output clocks of the PLL.

The PLLs in the FPGA take an input clock that drives a voltage controlled oscillator (VCO). The output clock of the PLL is down sampled from the VCO by two 8 bit registers. The frequency, f_{VCO}, of the VCO is given by,

$$f_{VCO} = \frac{m f_{in}}{n}, \tag{3.1}$$

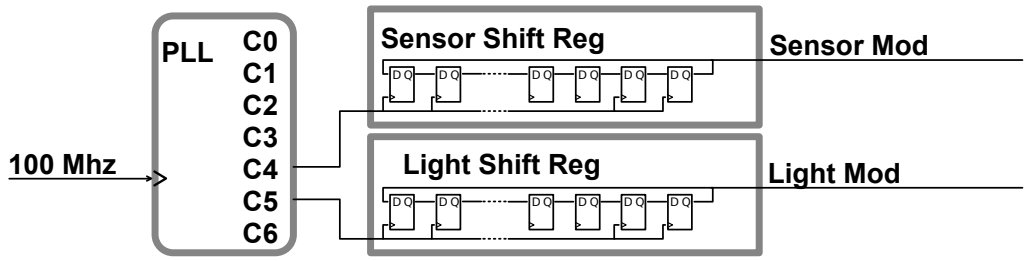


Figure 3.9: Method of generating binary sequences in the PMD19k camera. The binary sequence is preloaded into the D flip–flops and shifted out on the rising edge of the generated clock. The sequences are phase shifted by changing the relative phase of the driving clocks.

where m and n registers are 9 bits to give the range from 0 to 511. The VCO frequency in the Altera Cyclone IV FPGA operates between 600 MHz and 1600 MHz. The output frequency is set by two 8 bit counters c_{hi} and c_{low} such that

$$f_{out} = \frac{f_{VCO}}{c_{hi} + c_{low}}. \tag{3.2}$$

As the modulation frequency is set by a number of integer registers, not all frequencies are possible. The number of phase steps for a 2π phase shift is

$$S = 8(c_{hi} + c_{low}). \tag{3.3}$$

For a binary sequence of J bits long, the number of phase steps required to phase wrap the shift register chain in figure 3.9 is

$$S = 16J(c_{hi} + c_{low}). \tag{3.4}$$

The maximum VCO frequency sets the minimum phase shift in time. With a maximum frequency of 1600 MHz the minimum phase shift is 1/(8×1600 MHz) = 78.1 ps. During the readout process each PLL clock is phase shifted to the updated phase offset. As the phase offset is an integer it can be impossible to evenly sample over the $[0, 2\pi]$ interval.

3.2.3 UDP Packetizer

The UDP Packetizer generates the UDP, IP, and MAC packets that encapsulate the image data, and stamps the frame and row number of the im-

age to each packet so the image can be reconstructed on the host computer. The valid packets are streamed into the triple speed ethernet (TSE) component, which is a 10/100/1000-Mbps Ethernet Mac provided by Altera (Altera, 2013b). The UDP stack is implemented in hardware to minimize latency and increase throughput. The overheads of TCP are prohibitive to implement in hardware therefore UDP was used as the transport layer.

3.2.4 Nios-II Design

A Nios-II soft core microprocessor in the FPGA design is used to implement a software UDP stack to communicate to the host computer. The UDP Packetizer UDP stack is only capable of transmitting packets. The Software UDP stack transmits and receives information from the host computer. The communication enables configuration of the camera by enabling registers in the camera to be read and written to. An ARP stack is implemented allowing the Ethernet switch to know which MAC address the camera's IP address corresponds to. In the future an embedded version of linux would allow for a full network stack and other options such as DHCP.

It was intended that the Nios-II configure the camera operation during and between each raw frame. An interrupt between each raw frame would allow for dynamic configuration of the camera between each raw frame, giving a high degree of reconfigurability with a simple interface. However the need for such a system was never realized.

3.3 Matlab Interface

Matlab (Mathworks, Natick, MA, USA) is a powerful numerical computing programming language, making it ideal for signal and image processing. To have a seamless development environment platform for testing novel algorithms that tightly couple the hardware and processing a Matlab interface was developed to configure the camera and capture data directly into Matlab. The camera API provides functionality to read and write all registers in the FPGA and capture image streams from the camera. Additional higher level functionality simplified the configuration of the modulation controller and ADC. The Matlab API for configuring the camera and capturing data is documented in Appendix B.

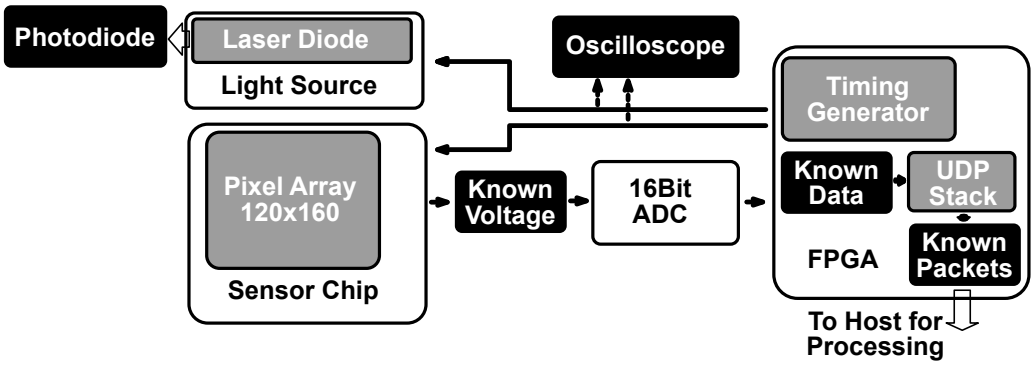


Figure 3.10: Camera design with test features in black. Known data is injected and confirmed received to verify the data processing pipeline.

3.4 Bring-Up

The process of bringing up a new system can be difficult and time consuming as one bug in any part can cause issues further down the chain. Each part of the design was built so that each part of the data output chain could be tested individually and verified before testing the next part. Figure 3.10 shows the test points of the data pipe on the PDM19k camera design. At each test point known data is injected and confirmed as received at the end of the pipe.

Known packets were generated and injected into the TSE Ethernet PHY. The packets were inspected with Wireshark on the host computer. Once correct packets were confirmed as being received the data acquisition tools were tested, verifying data was correctly imported into the processing framework. Then known data was injected into the FIFO data buffer, thus simulates data arriving at the ADC. With the design it is possible to de-couple the ADC input from the sensor. Known voltages are applied to the input of the ADC to verify its operation and readout process. The ADC noise, linearity, and gain can be measured with this design.

3.5 Characterization

Characterization of the camera enables and assessment of its performance. Jongenelen (2011) described and performed a full characterization on a ToF range camera. The characterization reported herein is not as comprehensive as Jongenelen, but is sufficient for performance of the camera is understood for better design of experiments.

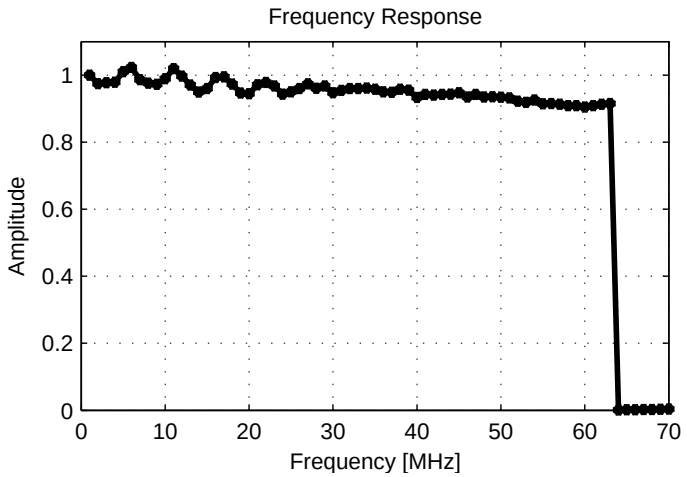


Figure 3.11: Measured amplitude with the modulation frequency of a time-of-flight range camera designed with a PMD19k3 sensor.

3.5.1 Frequency Response

The camera electronics are a low pass filter, which limits the maximum modulation frequency of the camera. To measure the frequency response the modulation frequency is swept from 1 MHz in increments of 1 MHz to 70 MHz. The measured amplitude is normalised against the first amplitude measurement. The frequency response is plotted in figure 3.11. After 63 MHz the camera stops working. The limited modulation frequency is believed to be due to impedance mismatch on the PCB, which causes ringing on the PCB trace. The ringing distorts the modulation signal such that the input high voltage threshold for the PMD19k3 sensor is never reached and this stops the pixel array from receiving a modulation signal causing the abrupt drop in figure 3.11.

3.5.2 Binary Sequence Performance

Comparison between the expected and the measured cross correlation of the binary sequence [0 0 0 1 0 1 1 1 0 1] is plotted in figure 3.12. The expected power spectrum is calculated from the binary sequence as ideally the measured value is the circular convolution of the binary sequence with itself. The spectral power of the cross correlation for the expected and measured power spectrum is plotted in figure 3.13. The difference is caused by the frequency response of the camera, as the higher frequencies are attenuated. Within the discrepancies caused by the camera hardware the expected and measured power spectrum are in agreement for the cross-correlation shape.

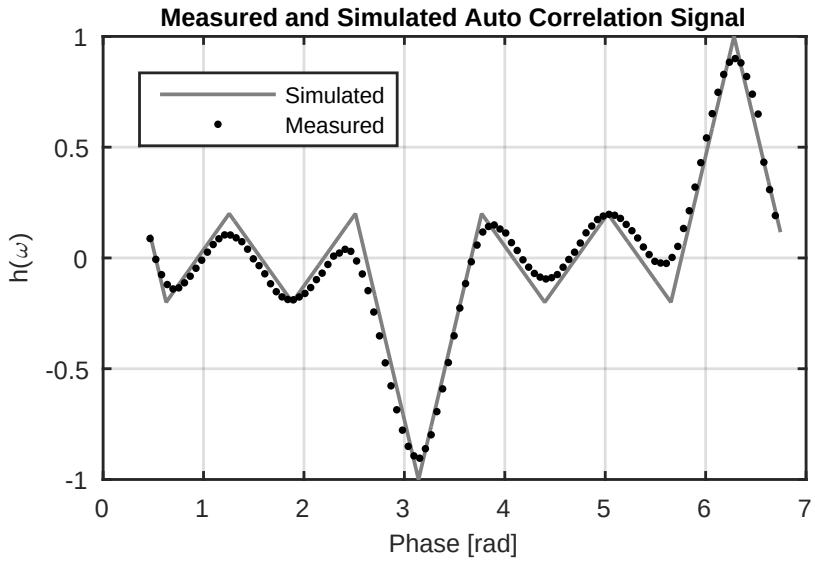


Figure 3.12: Measured and simulated cross correlation signal for the binary sequence [0 0 0 1 0 1 1 1 0 1].

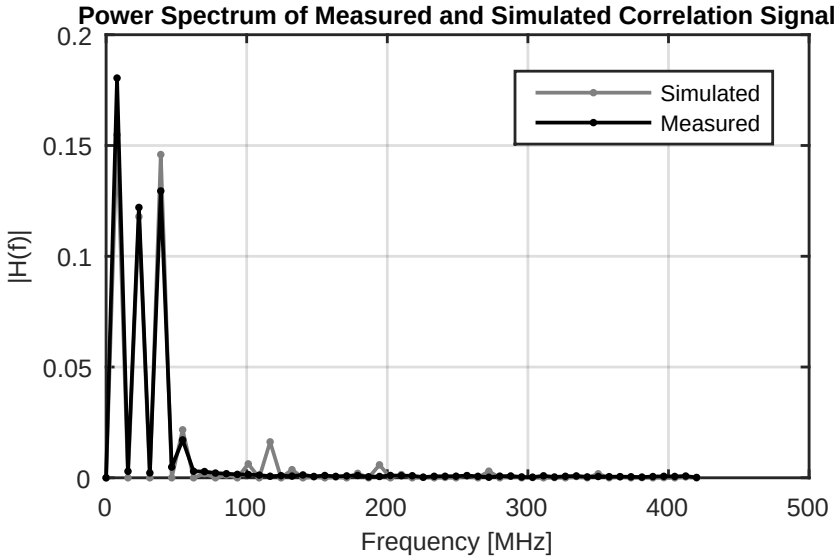


Figure 3.13: Measured and simulated power spectrum for the signals plotted in figure 3.12.

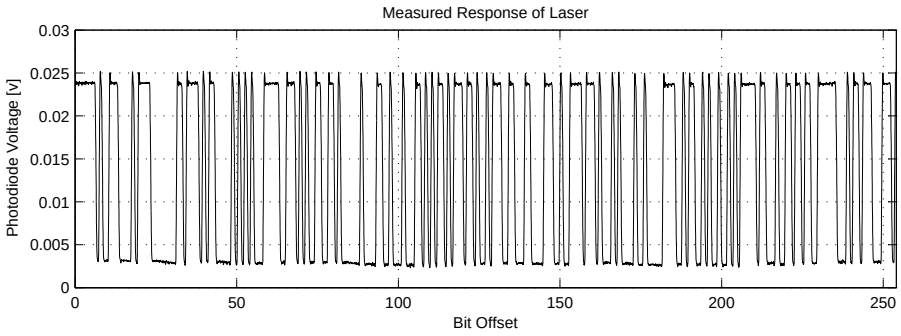


Figure 3.14: Measured response of light source with a 254 bit long m-sequence with a bit time of 1/80 MHz.

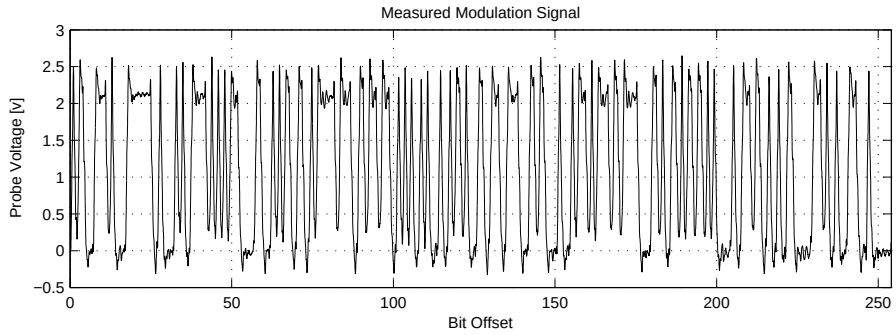


Figure 3.15: Measured input modulation to sensor array with a 254 bit long m-sequence with a bit time of 1/80 MHz.

The illumination system is critical to the camera's performance. To test the light source performance the light source is modulated with a 254 bit maximum length sequence, with a bit time of 1/80 MHz. The measured optical waveform with a photodiode is plotted in figure 3.14. The fast rise time and minimal ringing on the light source output indicates good performance. The measured input modulation signal for the PMD19k2 Tap A is in figure 3.15. The predicted response from the measured input modulation and the actual measurement are shown in figure 3.16.

The difference between the predicted and the measured response in figure 3.16 is most likely due to the measured input modulation signal not representing the internal workings of the pixel. The difference between predicted and actual response motivated Payne *et al.* (2009) to characterize a ToF pixel array using a femtosecond laser.

3.5.3 ADC Performance

The ADC performance of the PMD19k2 design was negatively affected by improper decoupling capacitors between the power supply and the ADC. The

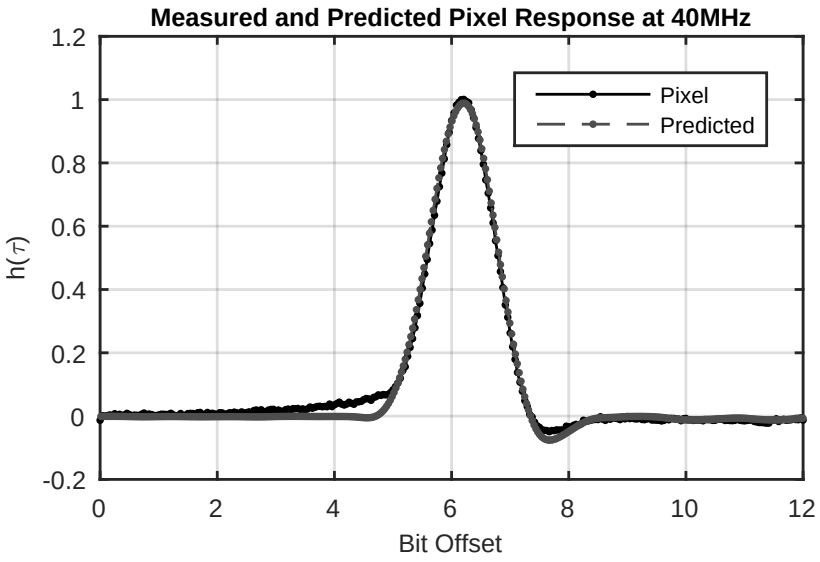


Figure 3.16: Measured versus predicted cross-correction of a maximum length sequence with a bit time of 1/80 MHz. The predicted is based on the laser response and input modulation.

distribution of ADC measurements for a constant voltage for version 1 and version 2 PMD19k2 designs are plotted in figure 3.17, note the y-axis is logarithmic. The long tails in the V0 design have been reduced due to changes in the V1 design, resulting in a tighter distribution which ultimately means less noise in the range measurement. The readout of a constant voltage on the ADC in the V0 design is imaged in figure 3.18. There are repetitive increase and drops in the ADC value over the entire image. It is hypothesized these are caused by the reference voltage dropping and recovering during the readout due to improper power supply decoupling.

3.5.4 Signal Integrity

ToF range cameras operate in the tens to hundreds of megahertz frequency range therefore signal integrity is an important design issue. Incorrectly designed PCBs will cause ringing and crosstalk on the modulation and analogue signals which introduces errors. An example of systematic phase noise introduced by cross talk on the PCB is shown in figure 3.19, where a systematic diagonal ripple is present in the phase image. Across the entire readout there is a diagonal ripple pattern present, ideally only white noise should be present in the phase image, any systematic offsets of this nature should not exist. Meeting regulatory standards for spectral emissions can be challenging because of the high speed switching of the high current light source and pixel drivers.

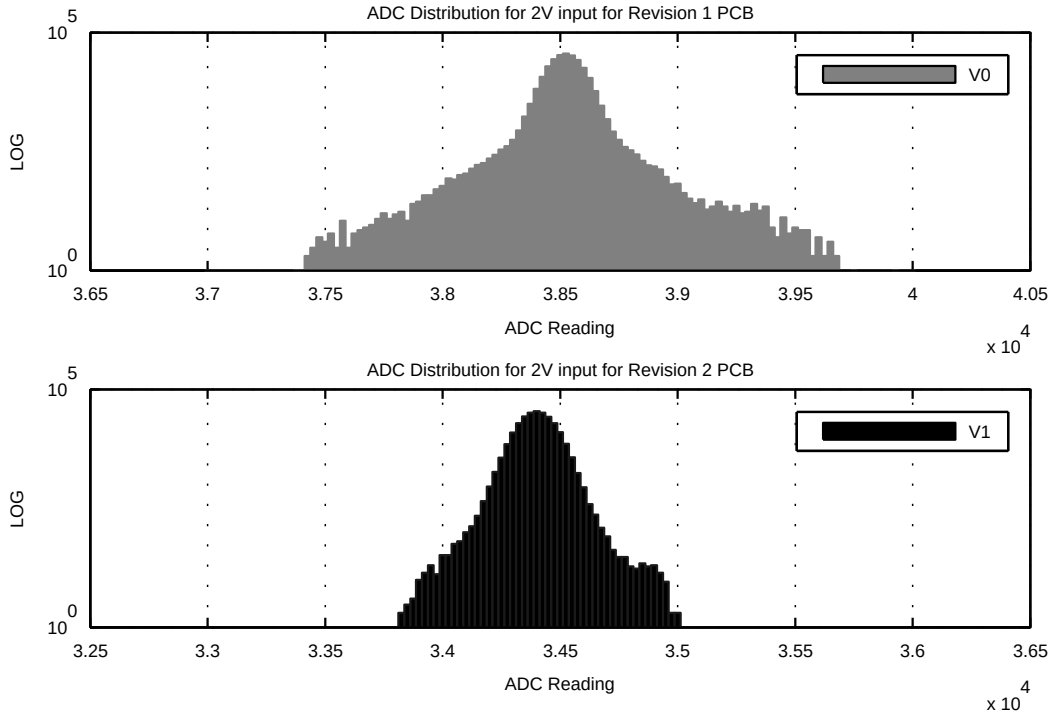


Figure 3.17: Distribution of ADC measurements for a constant voltage for two different PCB designs.

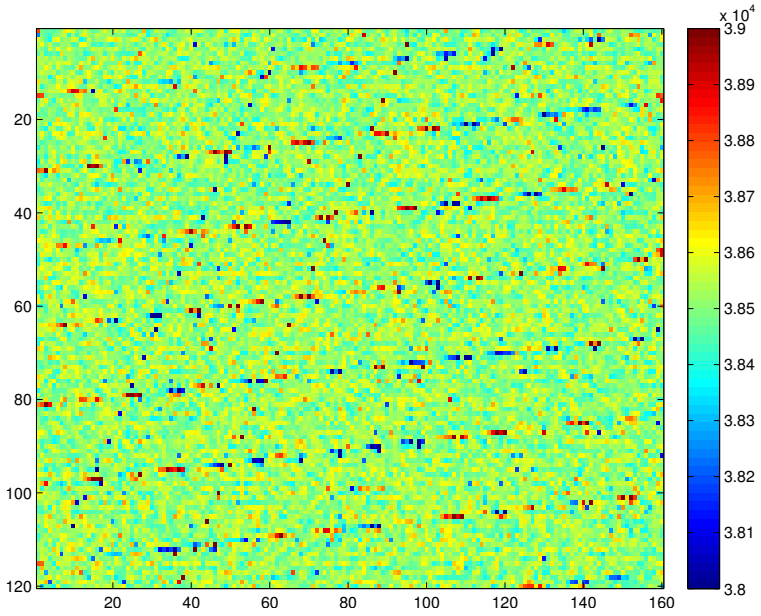


Figure 3.18: ADC values for a constant input voltage showing systematic error during the readout process, caused by improper decoupling of the ADC power supply.

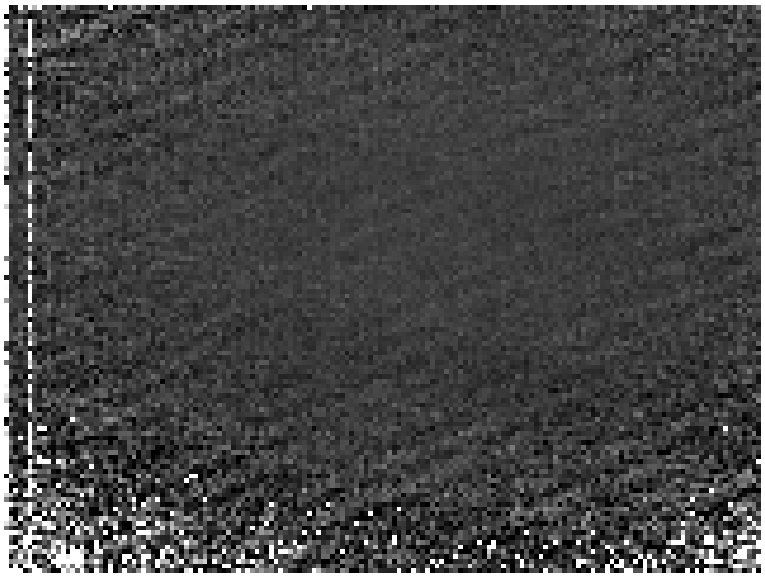


Figure 3.19: A diagonal ripple pattern is present in the phase image. This is an example of artifacts introduced by cross talk on an improperly designed PCB.

Impedance matching and good design practice is essential to minimize signal integrity issues.

3.6 Conclusion

This chapter has given a brief overview of a time-of-flight range camera system design. A camera encompasses mechanical, optical, PCB, signal integrity, FPGA, and software design. It is a multiple discipline process to design an effective ToF range camera. Each component has the potential to negatively affect the camera's performance if designed incorrectly. Minimizing cross talk and correct impedance matching is critical for camera performance. The camera is designed for flexible control of a ToF range camera for development of novel processing algorithms that tightly couples the hardware with signal processing.

As the demand for smaller devices increases it is expected that ToF range cameras will become a system on a chip (SoC). The SoC will combine the pixel array, ADCs, and digital processing logic. The components required to build a ToF range camera will become a SoC, a light source, diffuser, and lens. A small form factor will allow for integration into multiple devices creating new applications for range imaging.

Chapter 4

Temporal Coding

The current paradigm of time-of-flight (ToF) range imaging cameras is to illuminate the scene with amplitude modulated light, with the phase of the reflected modulation envelope encoding the distance travelled. The distance is measured by correlating the reflected light with a signal that has an identical frequency of amplitude modulation as the emitted signal. This technique is known as amplitude modulated continuous wave (AMCW) homodyne detection. In this chapter the usage of binary sequences as the illumination and reference signals is explored and referred to as coded ToF, a term coined by Kadambi *et al.* (2013). AMCW light modulation is a simple on off repeating pattern. Coded ToF uses more complex patterns of light modulation. The focus of this chapter is the application of coded ToF for resolving multi-path interference.

In this chapter the operation of ToF cameras with binary sequences is introduced. The selection of binary sequences, with the use of binary sequences to resolve multi-path interference, is investigated by using sparse deconvolution.

4.1 Previous Work

The application of binary sequences started with cryptography, radar, sonar, and communication systems (Golomb and Gong, 2005). Spread spectrum modulation is a communication technique that utilizes binary sequences to provide transmissions that are resistant to interference and multi-path interference (Haykin, 2008). ToF range cameras can be described as a communication system: the light source is the emitter, the scene is the communication channel, and the pixel is the receiver. The communication channel (scene) can suffer

from errors caused by multiple propagation. Techniques from communication systems can be applied to ToF range cameras for performance improvement.

The two most common types of spread spectrum communication are direct sequence spread spectrum (DSSS) and frequency hopping spread spectrum (FHSS). When applied to ToF range cameras DSSS modulation is applied by replacing the modulation scheme of the light source and reference signal with the spreading code. FHSS is applied by changing the modulation frequency during the camera's integration period. Previous work on the application of binary sequences in ToF range cameras has focused on providing resistance to interference from other cameras. Büttgen and Seitz (2008) used DSSS modulation with maximum length sequences (m-sequences); their low cross-correlation results in minimal interference between multiple cameras. The measured cross-correlation is a function of the binary sequence used, therefore Büttgen and Seitz developed a new computation of depth from the cross-correlation measurement. Whyte *et al.* (2010) modulated an m-sequence onto the carrier frequency, allowing for the current calculation of distance from the phase of the carrier and multiple cameras to operate with minimal interference. Bamji (2008) applied FHSS to the modulation frequency to operate multiple cameras simultaneously.

Wireless communication channels often suffer from signal degradation caused by multiple propagation paths from the transmitter to receiver. To correct for multi-path interference wireless systems can estimate the channel parameters by probing the channel with a known signal and measuring the channel response. Under the assumption the channel response is sparse Bajwa *et al.* (2010) used compressed sensing techniques to estimate the channel response of a wireless connection, by probing the channel with known binary sequences.

The ideas for channel estimation from wireless communication can be applied to ToF range cameras, where the channel is the scene. By probing the scene with known signals the multi-path interference can be estimated. The ToF pixel output is the cross-correlation between two signals. The cross-correlation results in an indirect measurement of the returning signal, complicating the process for resolving multi-path interference. Godbaz *et al.* (2008) used the assumption of sparsity and performed deconvolution on the measured cross-correlation of two square waves. However the cross-correlation of a square wave only contains odd harmonics which decay at $1/n^2$, so the sparse deconvolution results were limited by the spectral content. Hofbauer *et al.* (2014) used triangle basis functions to estimate the phase and amplitude of

the measured triangle cross-correlation. Using this approach the error due to harmonic aliasing was reduced, and multi-path interference was detectable.

There are many families of binary sequences including, but not limited to, maximum length sequences (m-sequences) (Golomb and Gong, 2005), MURA codes (Gottesman and Fenimore, 1989), gold codes (Gold, 1967), or Walsh-Hadamard codes (Hathi *et al.*, 2002). The application in question naturally influences the choice of code family from which to select a binary sequence.

The properties of maximum length sequences (m-sequences) have been throughly explored by Golomb and Gong (2005), and are described by a primitive polynomial over the Galois field of two elements (GF(2)), and implementable with a linear feedback shift register. To resolve multi-path interference deconvolution is applied to the measured cross-correlation, where the convolution matrix is a circulant matrix generated from the auto-correlation of the binary sequence.

Signals can often be well-approximated as a linear combination of just a few elements from a known basis or dictionary, in the case of ToF range cameras the multi-path interference can be approximated by a discrete number of path lengths and amplitudes. Mathematically a signal, x, is k-sparse when it has k nonzero elements, $\|x\|_{\ell_0} \leq k$. If there are a limited number of propagation paths from the light source to pixel, then the measured cross-correlation signal can be represented by a sparse number of rows from the dictionary matrix H. With y being the measured signal (in the ToF case the correlation signal) then the formation model is

$$Hx = y \tag{4.1}$$

Sparsity has been exploited in signal and image processing techniques such as compression, machine learning and noise reduction. Under the assumption of sparsity there have been a number of signal recovery algorithms developed, and this has been a growing area of research due to the explosion of interest in compressed sensing in recent years.

To recover x in equation 4.1 an optimization problem can be formulated of the form

$$x = \arg\min \ \|y\|_0 \text{ subject to } y = Hx \tag{4.2}$$

The problem with the formulation in equation 4.2 is that resolving it is difficult, in fact even finding a solution that approximates the true minimum is NP-hard

(Muthukrishnan, 2005). In order to make equation 4.2 solvable in a reasonable amount of time the ℓ_0 minimization is replaced with ℓ_1 minimization, such that

$$x = \arg\min \ \|y\|_1 \text{ subject to } y = Hx \tag{4.3}$$

Provided that $Hx = y$ is convex then the resulting problem can be resolved with linear programming. An alternative approach is to use a penalizing term which results in

$$\arg\min_x \|Hx - y\|_2 + s\|H\|_1, \tag{4.4}$$

for $s \geq 0$. This formulation is known as LASSO and was presented by Tibshirani (1996). LASSO has the property that in higher dimensions it will bias towards sparse solutions when compared to ℓ_2 minimization.

While convex optimization techniques are powerful tools for computing sparse representations there is another family of greedy/iterative methods for solving such problems. Greedy methods operate by iteratively identifying the support signal until the convergence criteria are meet. One of the simplest greedy methods is orthogonal matching purist (OMP) which begins by finding the column of H that is most correlated with the measurement vector (Pati *et al.*, 1993). This column is then subtracted from the measurement and this step is repeated with the residual of the measurement with the remaining columns of H. The stop criteria is the error of the residuals or the number of nonzero elements of x.

In the case of the ToF range camera measurements the dictionary matrix H can be created as a Circulant matrix of the measurement when no multi-path interference is present. Using Circulant matrices as as the basis for sparse deconvolution for compressed sensing has been a growing area of research (Bajwa *et al.*, 2007) (Rauhut, 2009). This research is related to the selection of the most ideal binary sequence for resolving multi-path interference with sparse deconvolution methods.

4.2 Time-of-Flight Range Cameras With Binary Sequences

The illumination and sensor modulation signal is replaced by a binary sequence, p_γ, that is Γ bits long and expressed as

$$p_\gamma \in \{0, 1\}, \quad \gamma = 0, 1, .., \Gamma - 1. \tag{4.5}$$

Each bit has a transmission length of T_c and the index γ of the binary sequence is a function of time. A mapping from the continuous variable time to the discrete bit index is therefore

$$\gamma = \left\lfloor \frac{t}{T_c} \right\rfloor \bmod (\Gamma - 1). \tag{4.6}$$

The binary sequence is circular and is a function of time. By substituting equation 4.5 into equations 2.1 and 2.2 the reference and reflected signals become

$$s(t) = ap(t - t_0) + b, \tag{4.7}$$

$$g(t) = p(t), \tag{4.8}$$

where t_0 encodes the phase offset of the binary sequence due to the distance travelled, and b is the background light. The measurement at each pixel is the circular convolution of the reflected light's binary sequence and the reference binary sequence

$$h(\tau) = (s \star g)(t). \tag{4.9}$$

When $g(t)$ is a square wave then the modulation frequency of the source can be written in terms of the bit length T_c

$$\omega = \frac{2\pi}{2T_c}. \tag{4.10}$$

The correlation function $h(\tau)$ is the cross correlation of the two binary sequences, where one binary sequence has an unknown phase shift and amplitude due to its interaction with the scene. It is exactly this phase shift that we seek to recover from which we infer the ToF, and hence the range.

When multiple propagation paths are present from the light source to pixel the returned light is the integration of all the paths (Whyte *et al.*, 2014b). When using binary sequences the reflected signal is

$$s(t) = \int_0^\infty a(k)p(t - t(k))dk. \tag{4.11}$$

Using the assumption that a limited number of discrete propagation paths exist then equation 4.11 can be simplified to

$$s(t) = \sum_{k=1}^{K} a_k p(t - t_k), \tag{4.12}$$

where there are K returns. Equation 4.12 can be rewritten as a binary sequence convolved with a sparse spike train, where the location of the spike encodes the time delay, and the magnitude encodes the amplitude.

$$s(t) = p(t) \star a_k \delta(t - t_k) \tag{4.13}$$

The returning light is the circular convolution with the reference signal

$$h(t) = g(t) \star (a_k p(t) \star \delta(t - t_k)). \tag{4.14}$$

Using the associative property of convolution the equation can be rearranged to

$$h(t) = (g(t) \star p(t)) \star \delta(t - t_k). \tag{4.15}$$

The measured signal is the summation of individual signals at different offsets and amplitudes.

Figure 4.1 is a visualisation of the sparse forward model, for which the measured cross-correlation is the sum of circular shifted signals. Sinusoidal signals, when added together, produce another sinusoidal signal of the same frequency, therefore sparse deconvolution can not recover the individual components. This is the motivation for using a binary sequence as the individual components are recoverable.

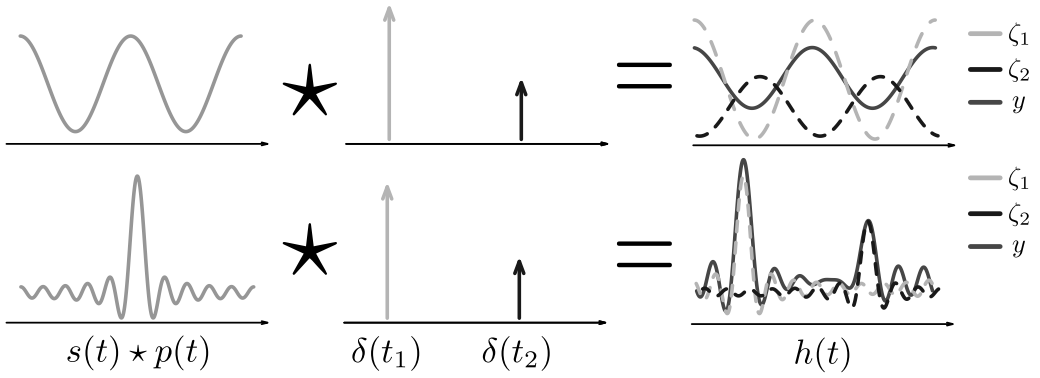


Figure 4.1: The cross-correlation of the binary sequence is convolved with the scene response generating the measured correlation signal. With a sinusoidal signal the convolution result is another sinusoidal so recovery is impossible. With another correlation signal the scene response can be recovered using sparse deconvolution.

4.3 Resolving Multi-path Interference

In the preceding section we derived a forward model that describes multi-path interference when ToF range cameras use binary sequences as the illumination and reference signals. In this section the forward model is used to develop a solution to multi-path interference. Using the measured cross-correlation as a basis, sparse deconvolution can be used to recover the multiple returns. The kernel vector is defined as the measured correlation signal with only one return, $k(t) = h(t)$, and the convolution matrix H is the circulant matrix (Davis, 1979) of the kernel vector. The measured correlation signal $\mathbf{y}$ is

$$\mathbf{y} = H\mathbf{x}, \tag{4.16}$$

where $\mathbf{x}$ is a sparse vector. The location of the non-zero elements of $\mathbf{x}$ encode the phase offset and their value encodes the amplitude of each return. By applying sparse deconvolution techniques to equation 4.16 the vector $\mathbf{x}$ can be recovered.

The formulated sparse deconvolution results in the following problem

$$\arg\min \|H\mathbf{x} - \mathbf{y}\|_2^2 \text{ subject to } \|\mathbf{x}\|_0 \leq K, \tag{4.17}$$

where $\|\mathbf{x}\|_0$ is the number of non-zero elements in $\mathbf{x}$, and there are K or fewer returns. Equation 4.17 is the classical compressed sensing formulation as developed by Donoho (2006), where H is the sensing matrix. Using the compressed sensing framework then if H meets the null space and restricted

isometry property (RIP) then $\mathbf{x}$ can be correctly reconstructed from L1 minimization methods (Donoho, 2006). Along with $\mathbf{x}$ being sparse all the elements are positive, as negative propagation paths do not exist. The circular convolution matrix, H, is determined by the cross correlation of the binary sequence.

$$H = \Lambda\Lambda^T, \tag{4.18}$$

where Λ is the circulant matrix of the binary sequence.

The rank of H determines the performance of the sparse deconvolution, the higher the rank the better. The rank of Λ is the same as H, the binary sequence and the number of samples determines the rank. Λ is an N by N matrix, the more samples taken the larger the matrix will be, and the discrete distance steps resulting from the quantization will be smaller. As ToF range cameras act as low pass filters the measured cross-correlation is a bandlimited signal. Once the Nyquist sampling criteria have been met there will be zeros in the Fourier domain and additional sampling will not yield more information, and decrease the rank of H, but ideally Λ is full rank. The sampling adds redundancy to Λ, but more samples decrease the distances quantization step size. As the number of samples increases the rank of H does not increase at the same rate and this limits the recovery of non-sparse elements in $\mathbf{x}$.

4.4 Sequence Selection

There are many binary sequences, and a number of variables which determine the ideal sequence for a given application. For the application of resolving multi-path interference using sparse deconvolution the circulant matrix of the binary sequence should ideally be full rank because of their desirable properties. One representation of a binary sequence is its Fourier transform which represents the frequency content of the binary sequence. Chapter 6 explores the duality between a binary sequence and a frequency sweep. In this application a binary sequence which maximises the frequency content over certain frequencies is desired. In applications for minimizing crosstalk between cameras, using families of binary sequences with low cross-correlation is desirable. Where the speed of depth acquisition is essential the length of the sequence is important. The sequence selection depends on the application, whether it is be resolving multi-path interference, measuring multiple frequencies or real-time depth acquisition.

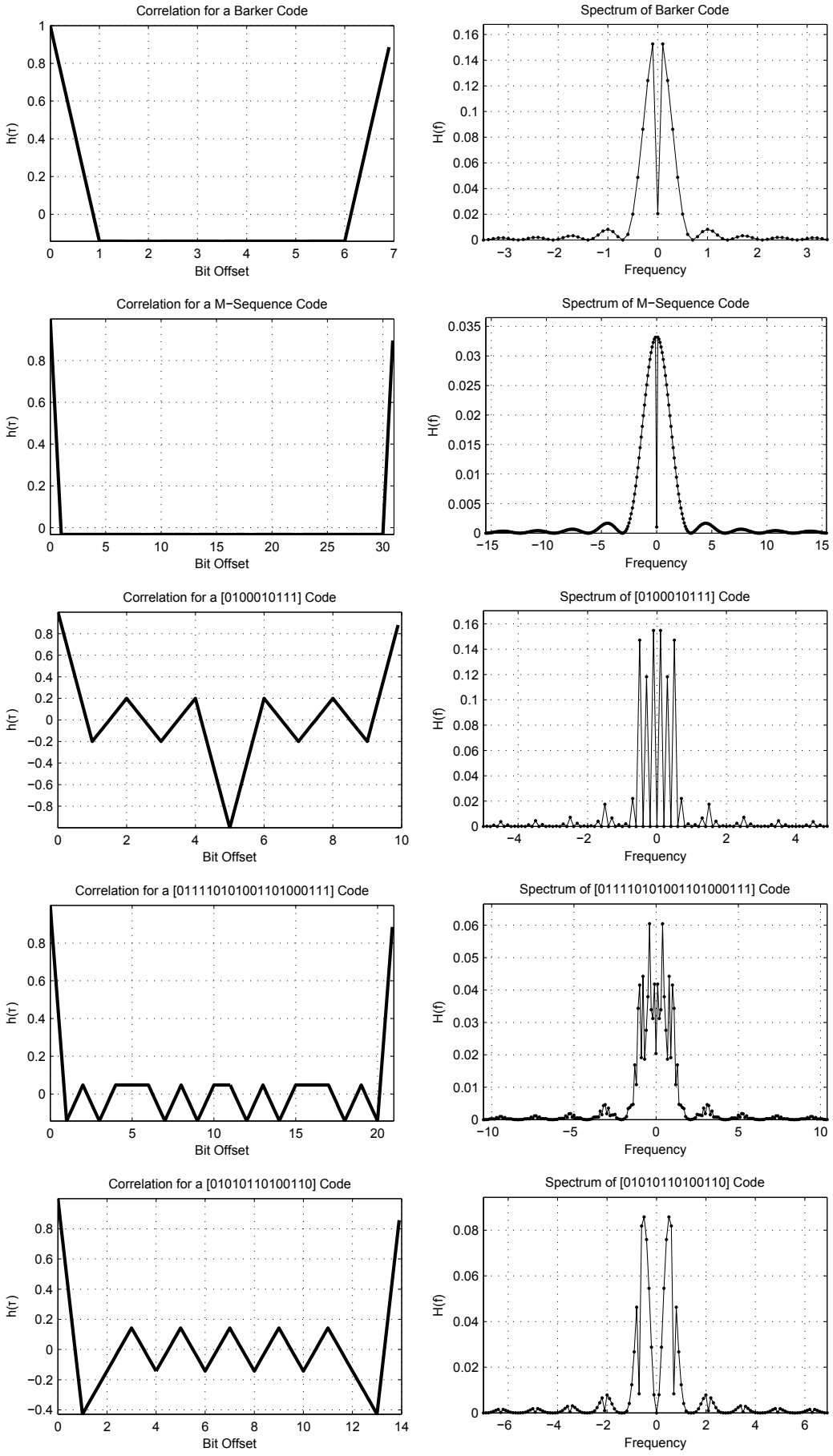


Figure 4.2: Comparison of binary sequences' normalised auto-correlation and spectra.

Various binary sequences' auto-correlation and their spectra are plotted in figure 4.2. The Barker and m-sequence produce nearly identical auto-correlation signals, however the m-sequence has a broader spectrum because it is a longer code. Barker codes are described by Golomb and Scholtz (1965), and they only exist up to length thirteen, and only nine codes are known to exist. The amplitude and distance information is encoded in the peak location and size, which is the same width for the m-sequence and Barker codes. The elements that are non changing in $h(\tau)$ contain no useful information about the distance and amplitude. Therefore using the Fourier bases and evenly sampling the correlation signal is not efficient as most of the samples are redundant. This is related to the property of the Fourier transform that a dirac delta function transform expands uniformly over the sampling interval.

M-sequences are well known binary sequences, and they are generated by a linear feedback registers (LFR). The LFR represents a primitive polynomial, and for a given register length of M D flip-flops the output sequence length is $2^M - 1 = N$ bits long, with an auto-correlation of

$$h(t) = \begin{cases} 1 - \|t - N\gamma\|, & \text{for } (N\gamma + 1) < t < (N\gamma - 1) \\ -1/(N-1), & \text{elsewhere.} \end{cases} \tag{4.19}$$

An example of this is plotted in figure 4.2, where there is a maximum at the point where the codes are phase aligned, and constant elsewhere. M-sequences generated with a different primitive polynomial have a small cross-correlation. This property makes then ideal for communication channels that have no control. The circulant matrix generated from an m-sequence is full rank, therefore has desirable properties for solving equation 4.17.

In the case where each frequency component of the correlation signal is used for further processing, as explored in chapter 6, the frequency response is ideally flat across all frequencies of interest. Given a binary sequence where the magnitude of each Fourier coefficient is α_k and the average magnitude is $\bar{\alpha}$, then minimizing the energy function

$$E(\alpha) = \sum_{k=1}^{K} ||\alpha_k - \bar{\alpha}||, \tag{4.20}$$

locates the most ideal binary sequence for this application.

4.5 Computing Depth

Equation 2.14 is the estimation of phase when using the standard AMCW technique, and assumes a sinusoidal cross-correlation signal. The binary sequence determines the cross correlation, therefore the computation of depth from the sampled values must change when deviating from the standard AMCW technique. With a sinusoidal cross-correlation the Fourier basis is ideal as the amplitude and phase are part of the basis. With different cross-correlation functions alternative bases may be preferential.

M-sequences cross-correlation is described by equation 4.19 and an example is plotted in figure 4.2. Büttgen and Seitz (2008) calculated distance, by sampling the rising and falling edge C_1 and C_2 respectively of the m-sequence cross-correlation,

$$d = \frac{C_2}{(C_1 + C_2)} \frac{cT_c}{2}. \tag{4.21}$$

This computation assumes an ideal correlation signal. The measured correlation for two light sources is plotted in figure 4.3. An ideal cross-correlation for an m-sequence is plotted in figure 4.2. Comparing figure 4.3 and 4.2 shows an ideal correlation is an invalid assumption. The actual delay versus that measured delay using equation 4.21 is plotted in figure 4.4. The difference between measured and ideal in figure 4.4 is caused by the non-ideal cross-correlation signal.

Some binary sequences' cross-correlation contain a well defined peak, as shown in figure 4.2. The size and location of the peak encodes the amplitude and distance travelled respectively. However a peak search quantizes distance and is susceptible to error if multiple peaks of similar amplitude are present. As shown by Hofbauer *et al.* (2014) an alternative basis can be used to demodulate the distance and amplitude. The Fourier basis can still be used as the individual frequency component's phase encodes the distance travelled (Whyte *et al.*, 2014a).

Sparse deconvolution can be used to measure distance and amplitude and recover multiple propagation paths. However, the limitations of sparse deconvolution methods are the distance is quantized to the number of samples and it is computationally expensive to perform.

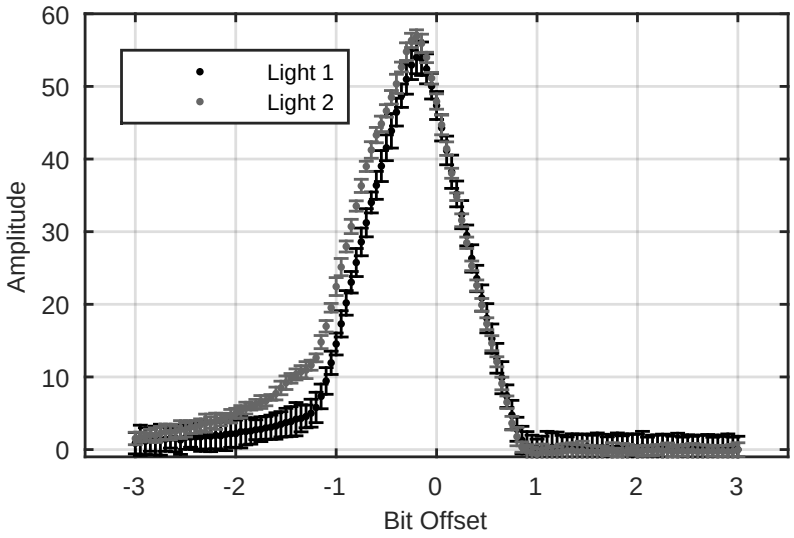


Figure 4.3: Measured correlation of a maximum length sequence using a ToF camera. The different plots are separate light sources with error bars at three standard deviations.

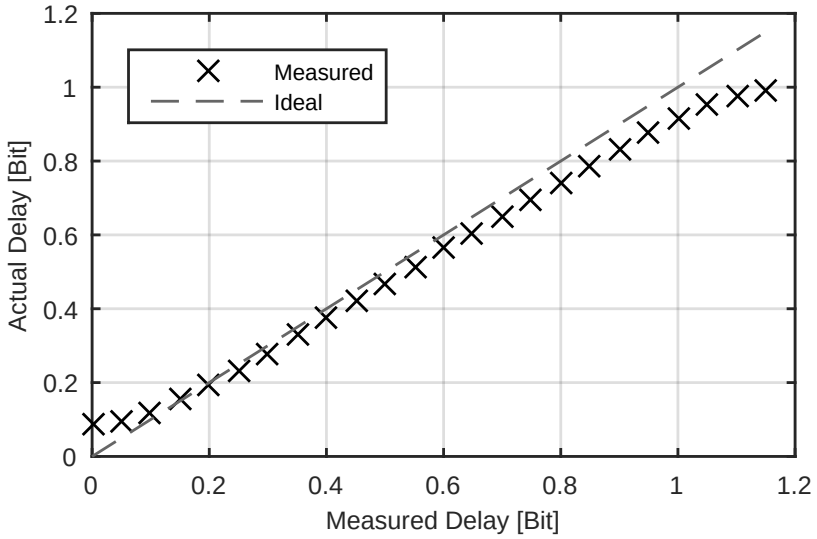


Figure 4.4: Calculated delay of m-sequence correlation and actual delay. The non-ideal correlation signal causes measurement errors.

4.6 Experimental Setup, Results and Discussion

The ToF camera described in chapter 3 and a 31 bit m-sequence with $T_c = 1/100$ MHz, is used for all experimentation. Six experiments are conducted: one simulation, four resolving two propagation paths and one resolving three propagation paths.

To compare results between experiments the x-axis on relevant figures is the offset distance in bits, where a square wave is two bits long, so a traditional π phase offset is one bit. This measure provides a convenient measure to compare different binary sequences of different lengths. With the used bit length of $T_c = 1/100$ MHz the bit distance is 1.5 m.

4.6.1 Sparse Deconvolution Testing

Pseudo inverse, LASSO and OMP methods are tested for resolving equation 4.17 to recover the propagation paths for when two returns are present, with the returns separated by 3.88 m and 2976 samples taken on the cross-correlation signal.

The measured cross-correlation when two returns are present is plotted in figure 4.5. The results of recovery methods to resolve the multiple propagation paths are in figure 4.6. Figure 4.5 shows the measured signal is the sum of varying amplitude and phase shifted cross-correlation measurements. The pseudo inverse fails to locate any returns, while LASSO is able to locate the two returns, however it blurs the location of the larger return, and OMP correctly reconstructs both returns.

4.6.2 Simulation

The simulation is a test of the performance of resolving two propagation paths using OMP as the sparse deconvolution method. The simulation changes the distance of the returns and relative amplitudes without the presence of noise. The results are the error in distance and percentage error in amplitude of the closest return.

Simulated results of the sparse deconvolution separation when no noise is present is shown in figure 4.7. Separation becomes possible when the returns are separated by over one bit (1/2 an ambiguity distance). Under one bit the recovery of returns is impossible with the current methodology. This result is

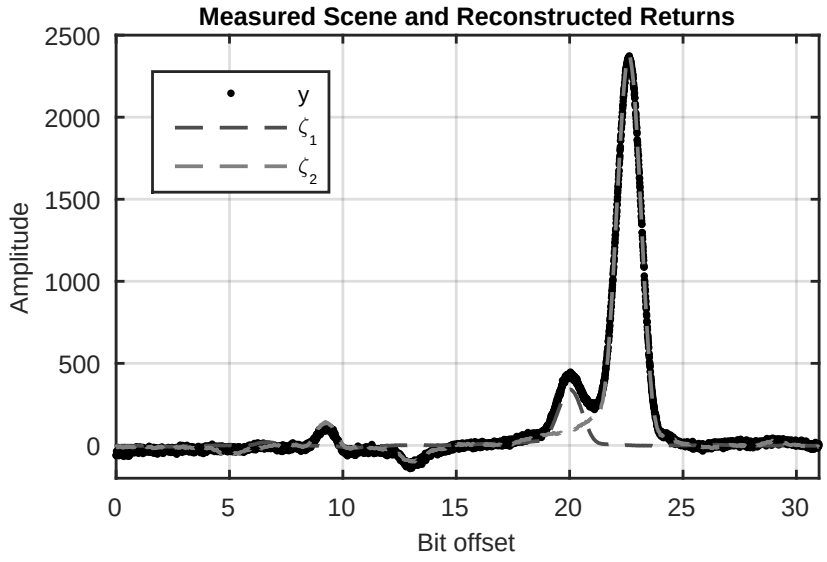


Figure 4.5: The measured signal, y, contains two propagation paths, therefore the measured signal is the sum of two signals, ζ_1 and ζ_2, which are overlaid on the measured.

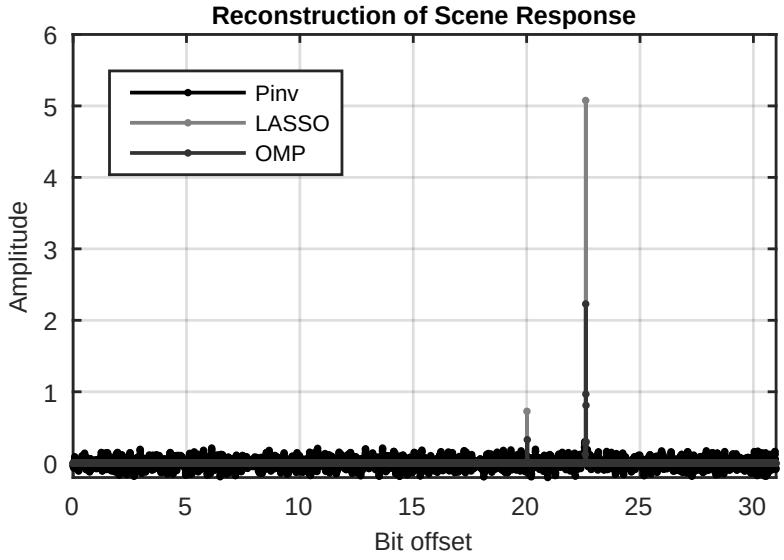


Figure 4.6: Results of matrix inversion using Pseudo Inverse (Pinv), LASSO and Orthogonal Matching Pursuits (OMP) techniques for resolving multi-path interference on the measured data from figure 4.5.

consistent with the maximum rank of the convolution matrix H from equation 4.18.

4.6.3 Two Return Recovery

To test the amplitude recovery a translucent sheet was placed 2.1 m in front of a wall with "TIME OF FLIGHT" printed on the wall. For testing the correct depth recovery a garden gnome is placed 2.22 m behind a translucent sheet and the distance error is compared to the distance error when the ToF camera operates in AMCW mode. The ground truth distance is measured by removing the translucent sheet. The application of both amplitude and depth recovery is tested with a scene of a coffee cup and translucent unicorn. The error of the measured distance of a garden gnome behind a translucent sheet is in figure 4.8. Using OMP sparse deconvolution the distance error is significantly decreased from over 2 m in most parts of the garden gnome to less than 0.1 m. However OMP does not always converge to the correct result and salt and pepper noise is present in the depth image. The root square mean error of the measured distance over the entire image for the full field image is 0.954 m, and for the coded ToF distance image is 0.145 m.

The results of measuring the amplitude of the reflection from a translucent sheet and the amplitude behind the translucent sheet are shown in figure 4.9. The text "TIME OF FLIGHT" is readable in the background. The results in figure 4.9 and figure 4.8 show that the amplitude and distance of multiple propagation paths can be recovered for an entire scene.

Another application of imaging using a translucent unicorn, is in figure 4.10. The distance and amplitude for the front and back of the unicorn are correctly reconstructed. However some pixels on the edge of the unicorn where specular reflections are present are not correctly recovered.

4.6.4 Three Return Recovery

To test recovery recover of three propagation paths a scene was constructed of two translucent sheets in front of a wall. There were 1.97 m between the translucent sheets and 1.81 m between the back sheet and object.

The results of expanding the method to three propagation paths is demonstrated in figure 4.11. The location and amplitude of each return is reasonably estimated using OMP. Theoretically more returns can be recovered, however as demonstrated in figure 4.7 each return must be separated by one bit and

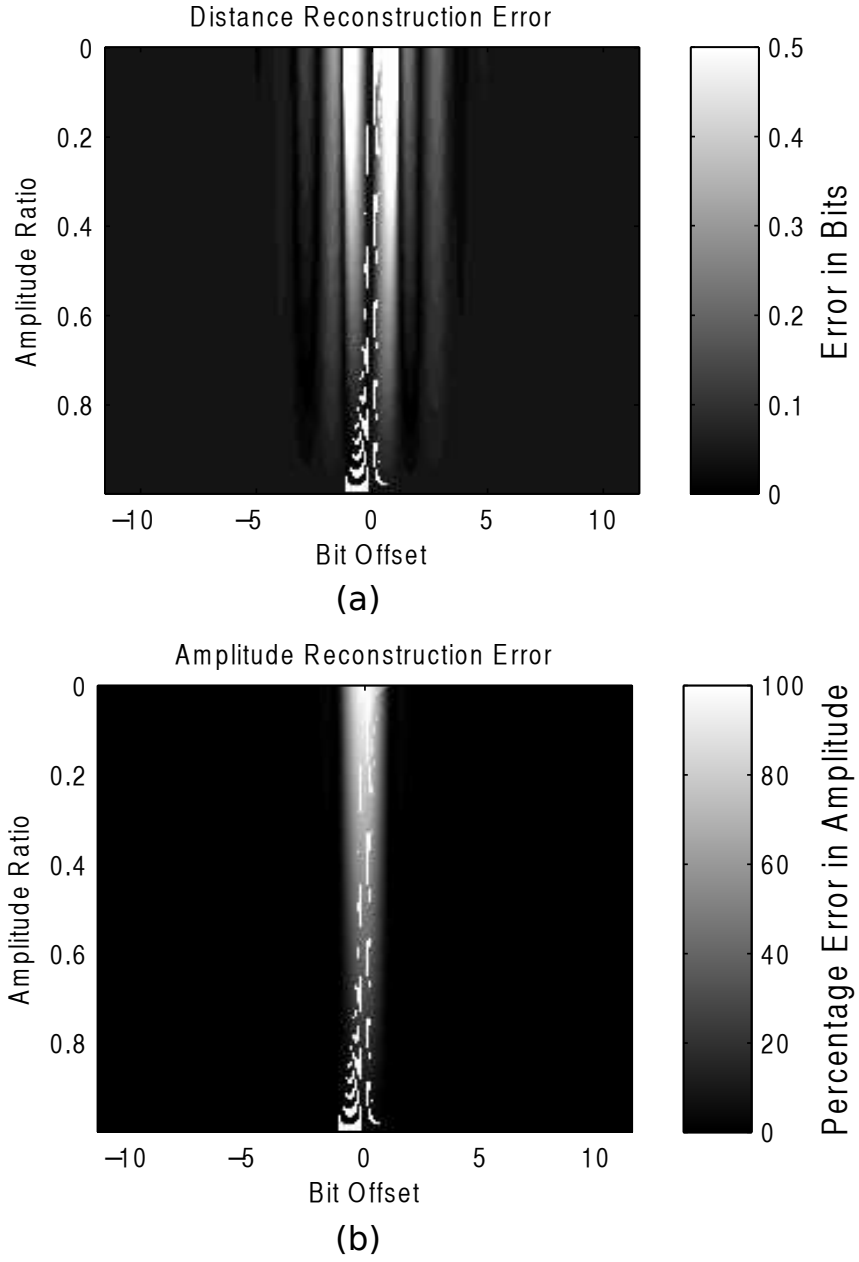


Figure 4.7: (a) Error in distance recovery in wavelengths for different path separations, in bit offset, and ratios of return amplitudes. (b) Percentage error in amplitude recovery for different path separations and ratios of return amplitudes.

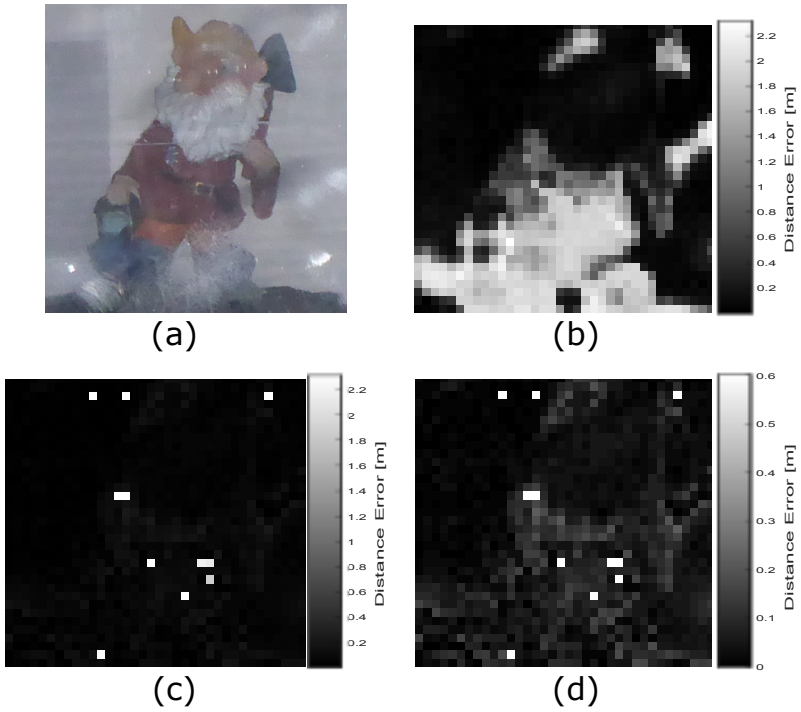


Figure 4.8: Comparison of the distance error of measuring a garden gnome behind a translucent sheet using full field and coded time-of-flight range measurements. (a) Photograph of the garden gnome, (b) distance error using full field, (c) distance error using coded time-of-flight, (d) same as (c) but with a reduced error scale.

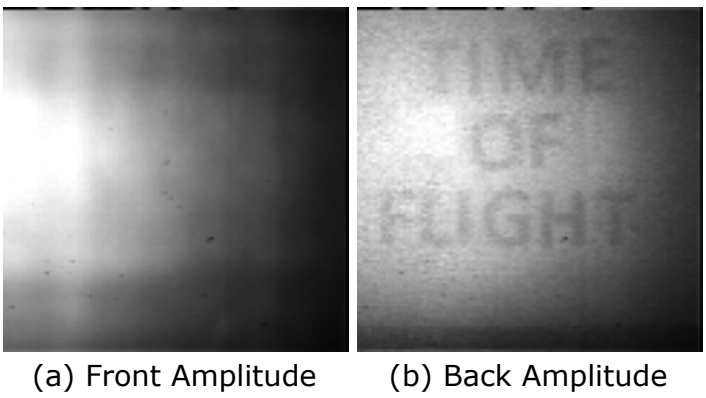
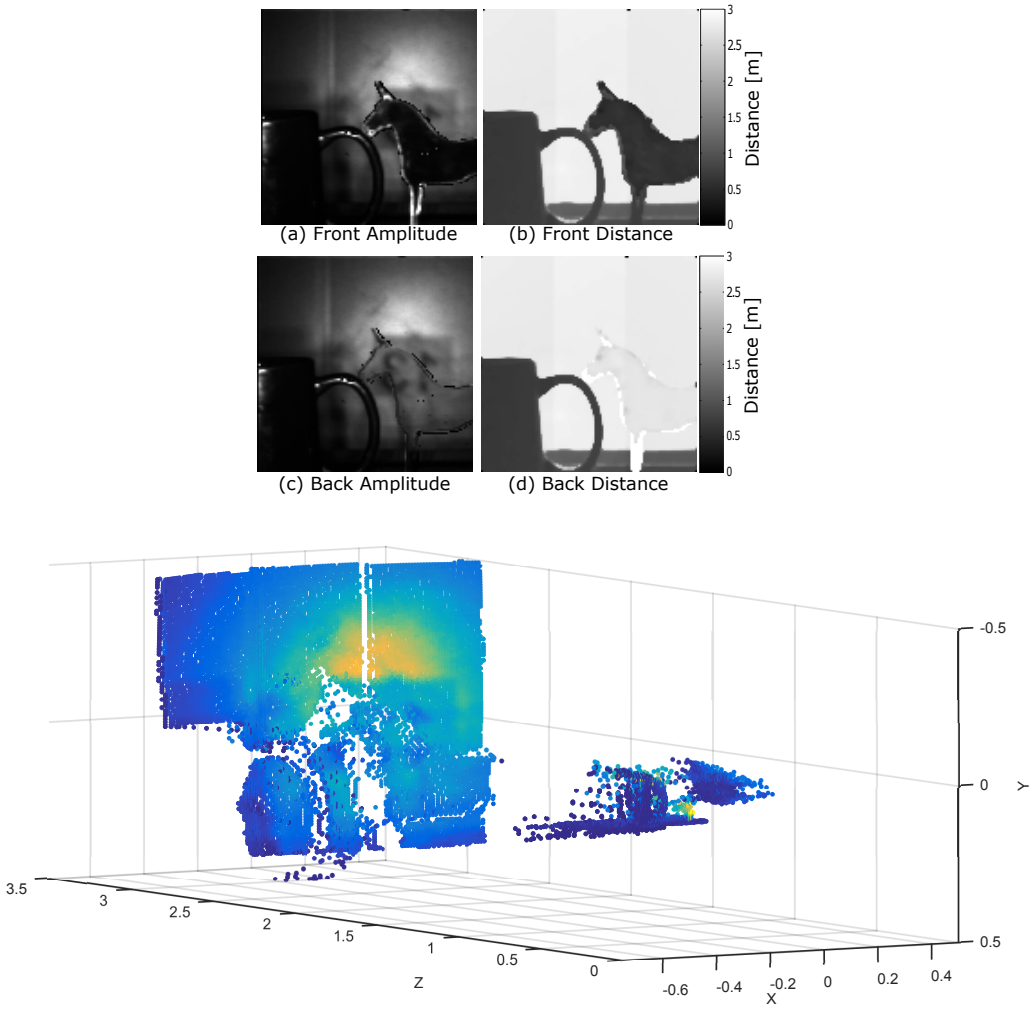


Figure 4.9: Recovered amplitudes of return from translucent sheet which is shown in (a), and behind translucent sheet in (b). The text "TIME OF FLIGHT" is readable in (b).



(e) Point cloud from combining all measurements

Figure 4.10: Reconstructed distance and amplitude of a translucent unicorn, (a) the amplitude of the shortest propagation path, (b) the distance of the shortest propagation path, (c) the amplitude of the longest propagation path, (d) the long propagation path distance, (e) point cloud of the combined results, showing the amplitude and distance of the unicorn and background.

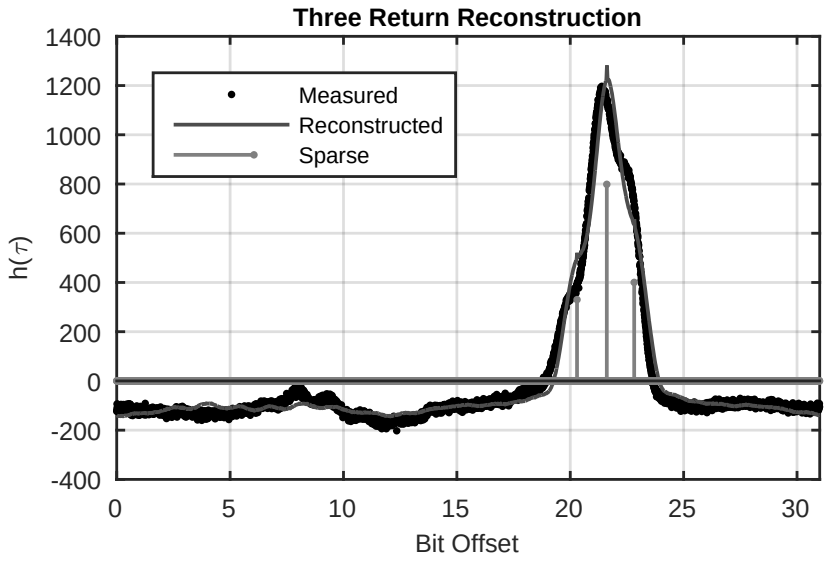


Figure 4.11: Results of using OMP to resolve three propagation paths of light using a time-of-flight range camera.

the amplitude decreases by $1/d^2$. Therefore practical measurements of more returns is difficult due to the low SNR ratio.

4.7 Discussion

There are four variables that contribute to the success of resolving multi-path interference: the relative amplitudes of each return a_1/a_2, the distance between each return $d_1 - d_2$, the number of samples on the correlation waveform N, and the signal to noise ratio (SNR) of each sample. This section has investigated the first two variables and how they affect the multi-path interference restoration. If the path length is separated by over one bit (1/2 an ambiguity distance) then the distance and amplitude of the multiple propagation paths can be recovered. This method fails when non-sparse multi-path interference is present, for example in corners and other diffuse inter-reflections. As the maximum modulation frequency increases the minimum separable distance also decreases.

One issue with using binary sequences is the computation of depth and amplitude from the measured correlation signal. With a square wave the Fourier transform is ideal as the phase and amplitude are part of the transform. Future work is to find a basis to compute depth when using an arbitrary binary sequence.

Chapter 5

Frequency Coding

Time-of-flight (ToF) range cameras currently measure distance by illuminating the scene with amplitude modulated light, at frequencys, nominally from 1 MHz to over 150 MHz, and the phase of the reflected light's modulation envelope encodes the distance travelled. In this chapter sampling over multiple modulation frequencies is investigated. Sampling over multiple frequencies has become more realistic because of the improvement in the high end of the modulation frequency of ToF range cameras. The increase in frequency bandwidth enables the application of radar image processing to ToF range cameras. Traditionally ToF range cameras sample over the phase of the reference signal. In this chapter sampling over both the phase and modulation frequency, and just the frequency are explored to resolve errors caused by multi-path interference and phase wrapping.

5.1 Background

5.1.1 Radar

Radar systems fall into two broad categories: pulsed and continuous wave (CW). Pulsed radar systems emit a short pulse and measure the time between the emission and reflected signal returning. The range resolution of pulsed radar systems is related to the pulse width time: the shorter the pulse width time the higher the range resolution. The SNR of pulsed radar is proportional to the power emitted, but it is difficult to put the same amount of energy into narrower pulses. These two reasons make accurate and low noise pulsed radar systems difficult to build. CW radars overcome these limitations by continuously transmitting into the scene. By modulating the carrier additional

information is encoded allowing for the range to objects to be measured. Three modulation schemes commonly employed are amplitude, frequency and phase modulation.

Amplitude and phase modulation are the operating principle behind AMCW ToF range cameras. Frequency modulating a CW radar such that the frequency changes linearly with time is known as frequency modulated continuous wave (FMCW), and was originally developed for altimeters for aircraft in the 1930s. Today FMCW radar is used when high bandwidth and low power range measurements are required. FMCW radar requires a small number of components which are in-expensive.

FMCW radar emits a linear chirp of frequencies and mixes the reflected signal with the current emission. The difference in frequency between the current emission and reflection encodes the distance travelled. To process the measured mixed signal the inverse discrete Fourier transform (IDFT) is used to convert from the frequency domain to the time domain. However the range resolution using the IDFT is directly dependent on the bandwidth of emitted frequencies. Spectral estimation techniques can be used to improve the range resolution.

One modification to FMCW is called stepped frequency continuous wave (SFCW), where instead of linearly sweeping the frequency each frequency is measured separately. SFCW radar compared to FMCW has the advantage of better SNR for each measurement and improved linearity of the selected frequencies, however it takes longer to perform a depth measurement. SFCW is common in ground penetrating radar where the SNR is important and there are more relaxed measurement times.

5.1.2 Time-of-Flight Range Cameras

Currently ToF range cameras are operated in an amplitude modulated continuous wave (AMCW) mode, in which the phase shift of the amplitude modulation encodes the distance. Equation 2.3 can be rewritten to include the modulation frequency as a variable. The combination of sampling over frequency and phase is

$$h(\tau, \omega) = a \cos\left(\omega\tau + \frac{2\omega d}{c}\right). \tag{5.1}$$

Thus sampling across two variables with a ToF camera is possible. Figure 5.1 is an illustration of the measured data from a ToF camera of sampling over

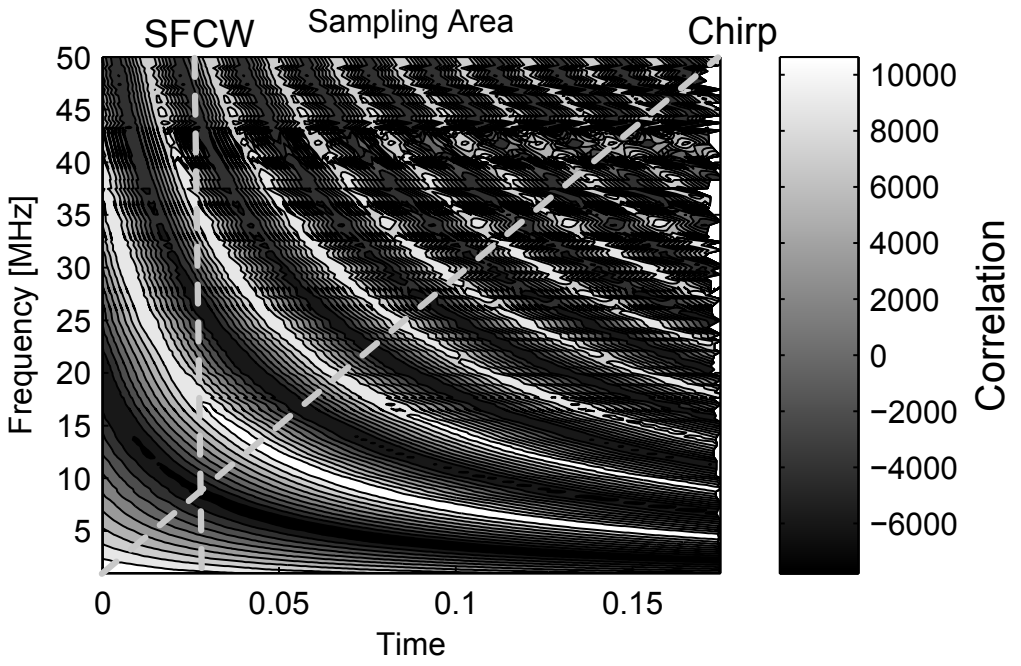


Figure 5.1: The sampling space of $h(\omega, \tau)$ in equation 5.1, with the frequency and time for the correlation signal measured with a PMD19k camera. At higher frequencies nonlinearities caused by the camera electronics become observable. The SFCW and chirp sampling lines are marked. The problem is to optimally sample over phase and frequency to resolve measurement errors in ToF range cameras.

frequency and time. The problem faced is how to sample over $h(\tau, \omega)$ most effectively to acquire depth while being resistant to measurement errors.

5.2 Phase and Frequency Sampling

Phase and frequency measurement samples the entire $h(\tau, \omega)$ area with N phase steps per frequency, and M frequencies. An introduction to the techniques that use multiple modulation frequencies to resolve multi-path interference and phase unwrapping can be found in chapter 2. Bhandari *et al.* (2014) sampled M frequencies to resolve K propagation paths, this work is included herein as it was a collaboration.

5.2.1 Multi-path Interference

Multiple propagation paths from the light source to pixel cause measurement errors in AMCW ToF range cameras. The forward models of how to describe multi-path interference is discussed in chapter 2. Bhandari *et al.* (2014) as-

sumes a sparse number of discrete propagation paths. Each measurement at a single frequency, ω measures the phase, ϕ, and amplitude, a. When multiple measurements are performed at evenly spaced frequencies the measurement can be represented as a complex vector, $\mathbf{z}$, for M frequencies.

$$\mathbf{z} = [\zeta_1, \zeta_2, \ldots, \zeta_M], \tag{5.2}$$

where each complex measurement ζ is the sum of K propagation paths

$$\zeta = \sum_{k=1}^{K} a_k e^{j\phi_k}. \tag{5.3}$$

The measured vector is constructed from

$$\mathbf{z} = \Phi\mathbf{g} + \epsilon. \tag{5.4}$$

The amplitude is encoded in the column vector $\mathbf{g}$, and ϵ is zero mean independent and identically distributed (i.i.d) Gaussian noise. The phase information is encoded in the matrix, Φ, which is an M by K matrix. Where there are M measurement frequencies and K discrete propagation paths. Φ is identified as a Vandermonde matrix, defined as

$$\Phi = \begin{pmatrix} e^{j\phi_1} & e^{j\phi_2} & \ldots & e^{j\phi_K} \\ e^{j2\phi_1} & e^{j2\phi_2} & \ldots & e^{j2\phi_K} \\ \vdots & \vdots & \ddots & \vdots \\ e^{jM\phi_1} & e^{jM\phi_2} & \ldots & e^{jM\phi_K} \end{pmatrix}. \tag{5.5}$$

The amplitude vector is $\mathbf{g} = [\mathbf{a_1}, \cdots, \mathbf{a_K}]^\top$. To recover the unknowns the similarity between Φ and an oversampled N by L discrete Fourier transform matrix, Ψ, where the elements $\Psi(n, l) = \exp(jml/L)$. If $L >> K$, the discretization of Ψ is small enough to assume that the columns of Φ are contained in Ψ. If the vector $\mathbf{u}$ whose elements are zero except for the K reflection amplitudes so therefore $\mathbf{z} = \Psi\mathbf{u}$. With this problem formulation sparse deconvolution methods can be applied by expressing the problem as

$$||\mathbf{z} - \Psi\mathbf{u}||_{\ell_\mathbf{2}} < \varepsilon_\mathbf{0} \text{ such that } ||\mathbf{u}||_{\ell_\mathbf{0}} = \mathbf{K}. \tag{5.6}$$

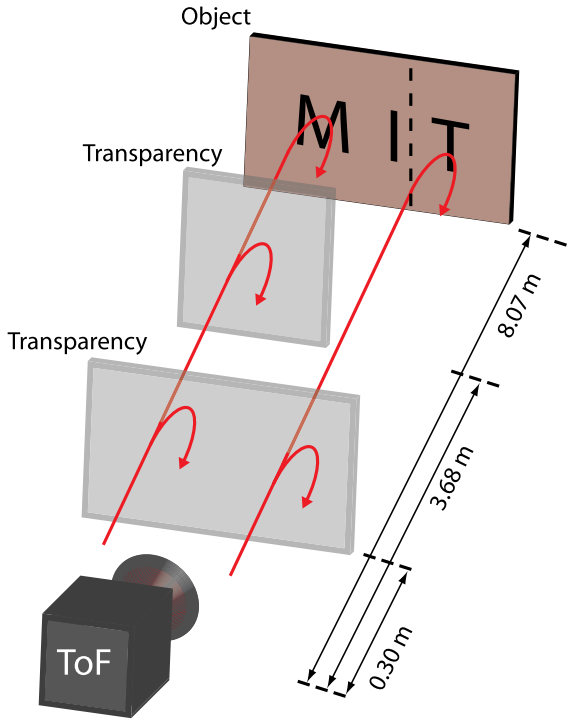


Figure 5.2: Experimental set up of resolving three propagation paths using measurements taken at multiple modulation frequencies.

Bhandari *et al.* (2014) used Orthogonal Matching Pursuits (OMP) to solve for the phases and amplitudes of K returns. To successfully solve equation 5.6 with $K = 3$, for three propagation paths, a minimum of five frequencies is required to find the six unknown variables.

5.2.2 Experimental Setup, Results and Discussion

Resolving Three Returns

Two experiments are conducted to test the OMP method of resolving multi-path interference. The first experiment is to test resolving multiple propagation paths, by imaging through two translucent sheets with the text "MIT" printed in the background. A diagram of the set up is in figure 5.2. A total of 77 modulation frequencies, spaced 0.7937 MHz apart were acquired and $K = 3$ returns is assumed.

A point cloud of the resolved propagation paths resolving using the OMP method is plotted in figure 5.3. The individual path lengths and their corresponded amplitude is plotted in figure 5.4. The text "MIT" is readable on the back sheet in figure 5.3, and in figure 5.4 the "T" is readable in figure 5.4e,

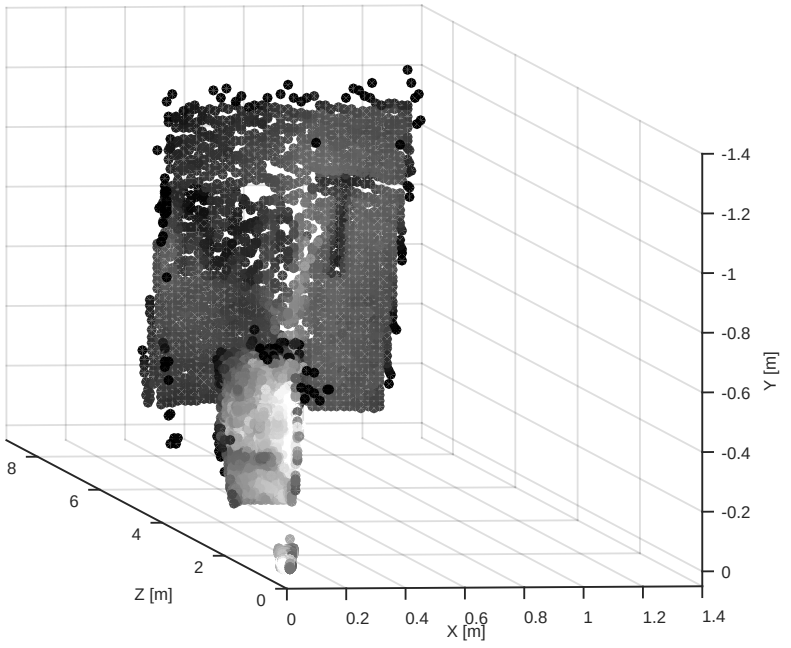


Figure 5.3: Point cloud of recovered distances and amplitudes using the OMP method for resolving multi-path interference. This is the result of the setup illustrated in figure 5.2, the front translucent sheet is much smaller because it is so much closer to the camera than the second translucent sheet.

and the "MI" in figure 5.4h. The amplitude and distance measurement for the original AMCW measurements are corrupted by the multi-path interference, and neither the text or distance is recovered.

Comparison

The second experiment compares the OMP method to other methods of resolving multi-path interference. The coded, mixed pixel, and OMP multi-path interference restoration techniques are compared. The 12 bit sequence [0, 1, 0, 0, 1, 0, 0, 1, 1, 1, 0, 1] is used for the Coded ToF method with T_c =1/(100 MHz), and 1728 measurements are taken on the correlation signal. The multi-path interference removal technique described by Godbaz (2012) is used with measurements taken at 25 MHz and 50 MHz. Godbaz (2012) assumed two propagation paths and used a two dimensional lookup table to correct multi-path interference at over 30 fps. For the OMP method eight frequencies are captured at 6.25, 12.5, 18.75, 25, 31.25, 37.5, 43.75, and 50 MHz. A garden gnome is placed 2.22 m behind a translucent sheet. The ground true was obtained by

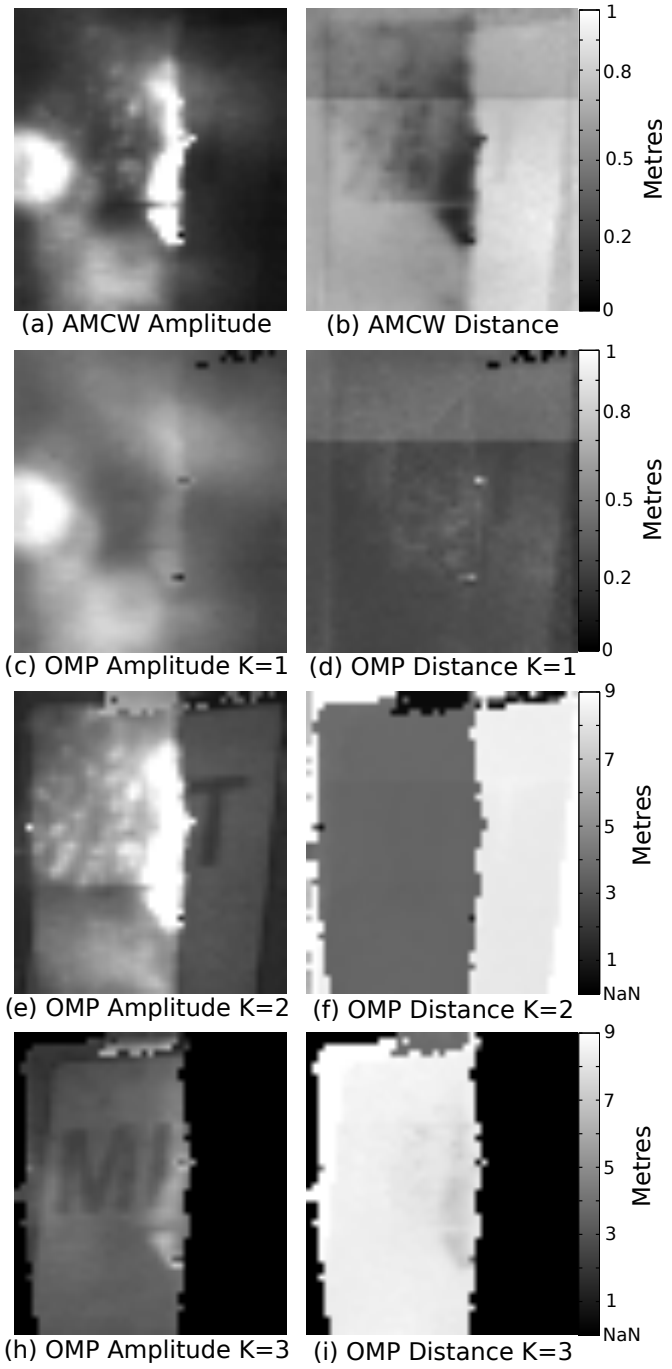


Figure 5.4: The AMCW amplitude (a), and the AMCW distance (b) is compared to the three measured propagation paths using the OMP methods, (c), (e), and (h) are the measured amplitude of the propagation paths, from shortest path length to the longest, (d), (f), and (i) are the recovered distant. The distance scales are different because of the large range measured over.

measuring the scene without the translucent sheet present. The error in the multi-path interference restoration is compared to the AMCW measurement at 50 MHz.

The results of OMP for resolving multi-path interference is compared against coded ToF, a method described in chapter 4, and mixed pixel a method described by Godbaz (2012), in figure 5.5. The AMCW distance error is plotted in figure 5.5b and compared to the OMP error with the same scale in figure 5.5c. Notice how the AMCW distance error is correlated with the colour on the garden gnome. For better compassion the OMP error, coded error, and mixed pixel error are plotted on the same scale in figures 5.5d, e, and f respectively. The RMSE of the AMCW distance is 0.954, OMP 0.214, coded 0.145, and mixed pixel 0.118. The OMP method suffers from noise introduced by the non-linear changes in amplitude and phase between modulation frequencies, which explains the higher RMSE value compared to the other methods. The ToF range camera acts as a low pass filter, higher order frequencies are attenuated, the frequency response of the PMD19k3 range camera is plotted in figure 3.11. At low frequencies the correlation signal is more triangular and becomes more sinusoidal at higher frequencies. A calibration for the phase and amplitude over frequency is required for practical implementation. The OMP method requires over 15 raw samples, while coded ToF generally takes over 2000 raw samples. The sampling constraints of coded ToF and OMP make it impossible for 30 fps of depth to be obtained for both methods. Both OMP and coded ToF methods can resolve three or more propagation paths, as demonstrated in figures 5.3 and 4.11, while mixed pixel can only resolve two propagation paths.

5.3 Frequency Sampling

By sampling at a constant phase of $\tau = 0$ the measured correlation signal is

$$h(\omega) = a \cos\left(\frac{2\omega d}{c}\right). \tag{5.7}$$

Equation 5.7 is identical to stepped frequency continuous wave (SFCW) modulation for radar (Charvat, 2014), given the phase offset, τ is constant. The frequency is incremented between each raw frame and the frequency of $h(\omega)$ encodes the depth. The measured data from equation 5.7 is real, where as in some radar systems the real and imaginary components are measured. In

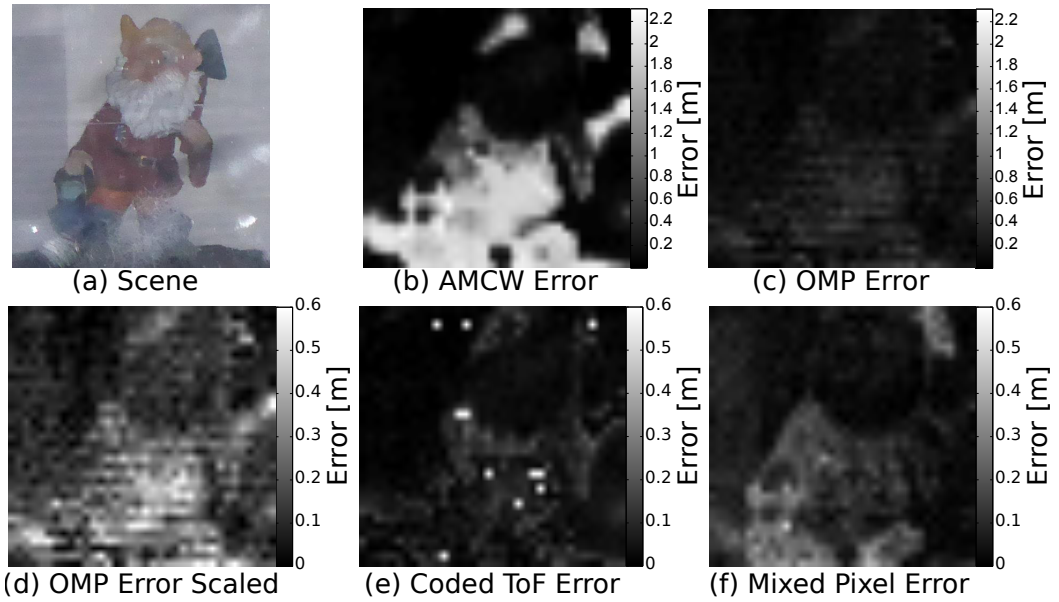


Figure 5.5: Comparison of methods to resolve multi-path interference of a garden gnome 2.22 m behind a translucent sheet, as photographed in (a), with AMCW in (b), OMP in (c) and (d), with coded ToF in (e) and with mixed pixel in (f).

both cases the distance is encoded by the frequency content of the signal. Normally the discrete Fourier transform (DFT) is taken to measure the frequency, however the bandwidth of a ToF camera is limited, therefore other spectral estimation techniques are used. The relationship between the estimated frequency, ω_{est}, and the distance is

$$\omega_{est} = \frac{BWd}{2c}, \tag{5.8}$$

where BW is the bandwidth of measured frequencies. Example raw data from the frequency sampling is plotted in figure 5.6. The closer object contains a lower frequency than the object that is further away.

5.3.1 Accuracy

The accuracy of the SFCW measurement depends on the resolution of the frequency estimation. Traditionally the DFT has been employed by the radar community. However the minimal resolvable distance with the DFT is

$$\Delta R = \frac{c}{2BW}. \tag{5.9}$$

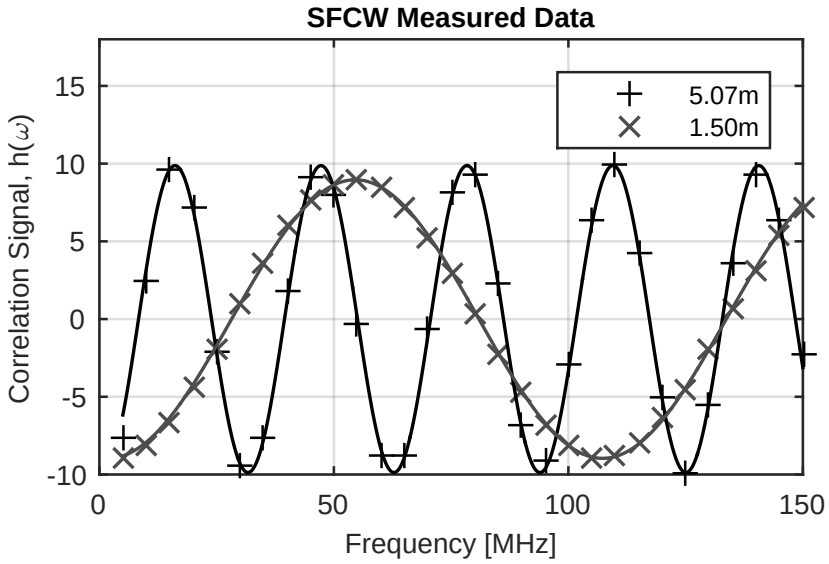


Figure 5.6: Raw data measured from stepped frequency continuous wave sampling. The frequency of the measured signal encodes the distance: the higher the frequency the further away the object.

Currently ToF cameras typically operate up to 150 MHz, therefore the distance quantization is 1 m. The use of modern spectral estimation techniques allows for better range resolution (Kay and Marple Jr, 1981).

The selection of the spectrum estimation is important for the quality of distance calculation. Modern approaches to spectrum estimation fall into three categories: parametric model-based, nonparametric, and sub-space models. Non-parametric methods make very few hypothesis on the input signal, while parametric methods work on priori assumptions about the signal. A special case of parametric methods is that of a signal composed of a finite number set of sinusoidal functions, which is common in signal and image processing. The process that generates the measured spectrum is well known, and a small number of sinusoidal signals is present in the measured signal. Naturally parametric and sub-space methods can improve the spectral resolution and fidelity for solving equation 5.7.

5.3.2 Multi-path Interference

When multiple propagation paths are present multiple frequencies are present in the correlation signal,

$$h(\omega) = a_1 \cos\left(\frac{2\omega d_1}{c}\right) + \int_{d_1}^{\infty} a_k \cos\left(\frac{2\omega d_k}{c}\right) dk. \tag{5.10}$$

The different propagation paths show up at different frequencies. The direct return has the lowest frequency, as the direct return is the shortest propagation path. If the spectral estimation technique can resolve the frequencies then multi-path interference has no effect on the measurement. For a sparse forward model the measured signal is the sum of discrete returns. The measurement across frequency for a discrete number of returns is modified from equation 2.19, as the measured data is now real, thus

$$h(\omega) = \sum_{k=1}^{K} a_k \cos\left(\frac{2\omega d_k}{c}\right). \tag{5.11}$$

5.3.3 Harmonic Aliasing

Harmonic aliasing is one source of measurement error in AMCW range cameras. Since square waves are used for the modulation signal there are higher order odd harmonics present in the correlation signal

$$h(\omega) = \sum_{\upsilon=1,3,5,\ldots}^{\Upsilon} \frac{a}{\upsilon^2} \cos\left(\frac{2\upsilon\omega d}{c}\right). \tag{5.12}$$

With SFCW the higher order harmonics appear as higher frequencies in the spectrum than the fundamental, therefore they are naturally separated from the distance estimate by the spectral estimation technique and do not cause error in the measurements.

5.3.4 Phase Wrapping

ToF range cameras when operating in AMCW mode suffer from phase wrapping and this is becoming more of an issue as the modulation frequencies increase. SFCW radar's ambiguity distance is caused by frequency aliasing and the maximum unique resolvable distance is

$$d_{amg} = \frac{c}{2\Delta f}, \tag{5.13}$$

where Δf is the separation between measured frequencies. With SFCW measurements the frequency separation is normally smaller than the AMCW modulation frequency, therefore when operating in SFCW mode it is less likely to encounter ambiguous distances.

5.3.5 Experimental Setup, Results and Discussion

The PMD19k3 camera described in chapter 3 has inadequate bandwidth for sufficient demonstration, therefore a proprietary camera with a maximum modulation frequency of 150 MHz was used for this experimentation. The fixed pattern noise and amplitude decay with frequency are both removed by calibration as the SFCW measurements are sensitive to these offsets.

AMCW verse SFCW

For the SFCW mode a scene with a maximum distance of 8 m is imaged, sampling from 5 MHz to 150 MHz in increments of 5 MHz. The scene is measured with and without a translucent sheet placed 1.1 m in front of the camera. This scene is designed to exhibit phase wrapping and multi-path interference. The SFCW measurement is compared to an AMCW measurement captured with nine raw frames at 20 MHz. The spectrum is estimated with the auto-regression pmusic method (Marple, 1987) as implemented in Matlab with an order of 5 and length of 1024.

Results comparing the AMCW and SFCW distance and amplitude measurements with and without a translucent sheet are shown in figure 5.7. The SFCW measurements can resolve the distance and amplitude of objects behind a translucent sheet, as demonstrated in figure 5.7(h), while the AMCW measurements cannot, as shown in figure 5.7(d). The SFCW measurement fails to measure the distance of dark objects behind the translucent sheet as these measurements are in the noise floor and the pixel values are discarded. A slice of measured distance for both SFCW and AMCW are plotted in figures 5.8 and 5.9. The AMCW measurement suffers from phase wrapping after 7.5 m, where as the SFCW measurement has an ambiguity distance of 30 m. The same slice is plotted when a translucent sheet is placed in front of the scene in figure 5.9. The AMCW measurement is corrupted by the multi-path interference, where as the SFCW measurement is less corrupted, and able to separate both returns. A box plot of the measured distance error of the pixels behind the translucent sheet for both AMCW and SFCW is in figure 5.10. The SFCW measurements have a lower mean error, and the variance of error is significantly reduced compared to AMCW. Referring to table 5.1, the SFCW measurements have much lower median error than the AMCW, and the spread in error as measured by the interquartile range is reduced by 22 times. This result shows SFCW modulation is resilient to multi-path interference and phase wrapping.

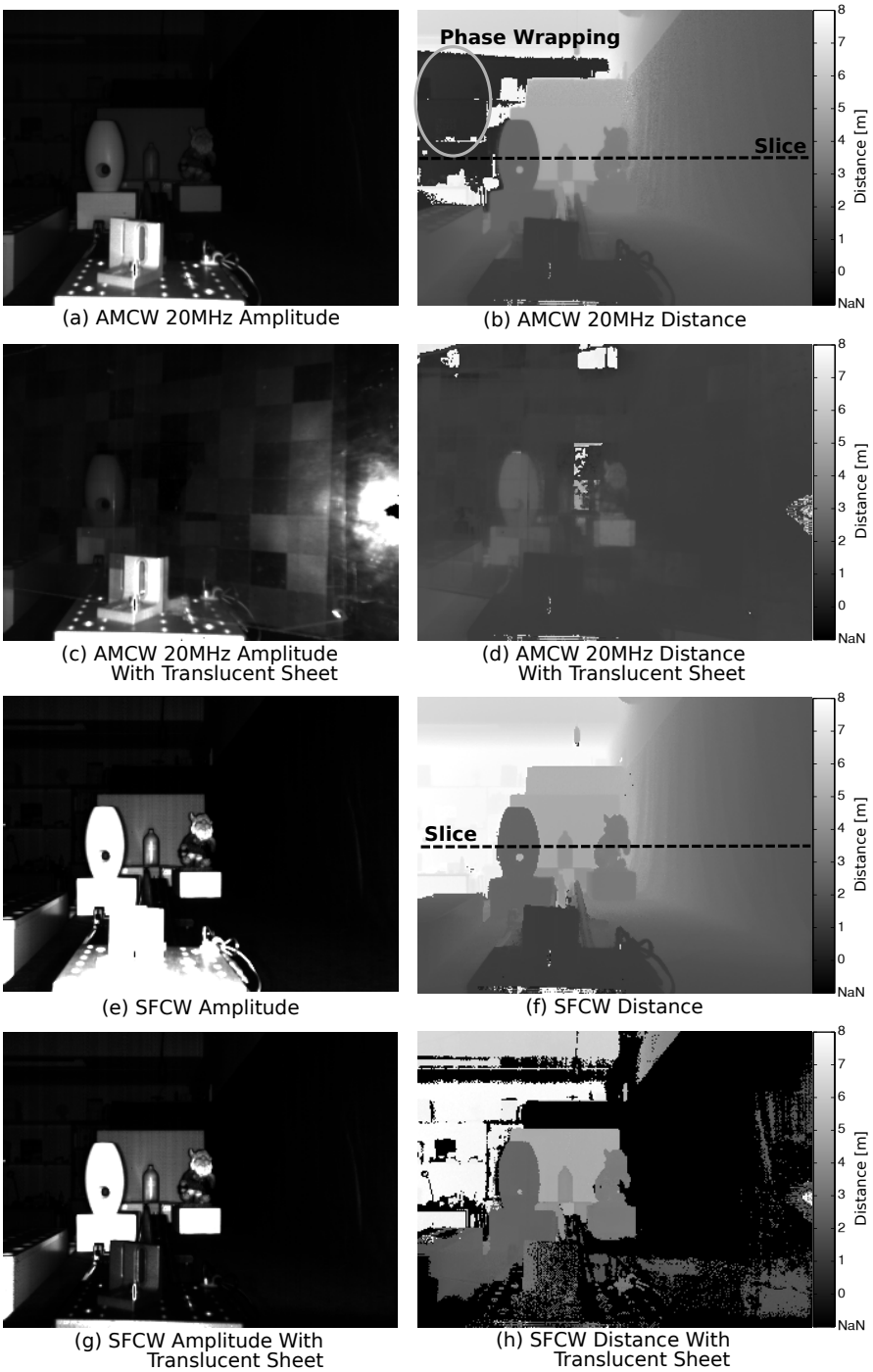


Figure 5.7: Comparison between AMCW measurements in (a), (b), without a translucent sheet and (c), (d) with a translucent sheet. These are compared to SFCW measurements in (e), (f) without a translucent sheet and (g), (h) with a translucent sheet.

Table 5.1: Quartiles of distance errors through a translucent sheet.

Quartile	AMCW	SFCW
25	-4.0 m	-0.18 m
50	-1.3 m	-0.06 m
75	0.0 m	0.0 m

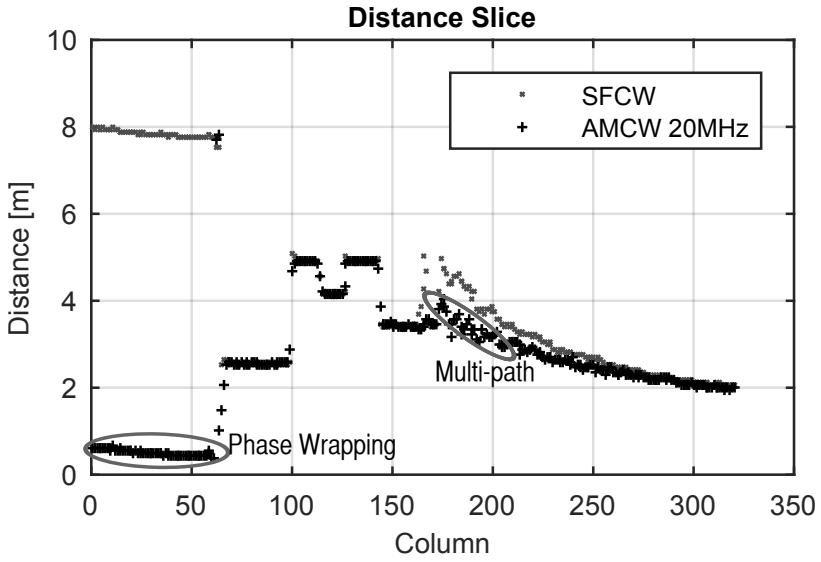


Figure 5.8: Distance slice comparing AMCW and SFCW without a translucent sheet as measured in figure 5.7 (b) and (f). The AMCW measurement suffers from phase wrapping and multi-path interference, while the SFCW does not.

SFCW Accuracy

The accuracy of the SFCW distance measurement for a single return is tested by moving a white sheet 1.1 m in increments of 0.05 m on a translation stage. The ground truth distance is known as the translation stage moves the object in a controlled and precise fashion. At each location the frequencies from 5 MHz up to 150 MHz in increments of 5 MHz are measured. The measured frequencies are down sampled to investigate the effect of the number of samples. The depth at each location is calculated using the pmusic method. The measured distance using SFCW is compared against an AMCW measurement of nine raw frames at 50 MHz. One hundred depth measurements are taken to compute the variance of each measurement.

The results of the distance accuracy of SFCW, with a bandwidth of 150 MHz and 10 frequency steps, is compared to AMCW, at 50 MHz with 9 phase steps, for a single return is plotted in figure 5.11. For a single return AMCW is cur-

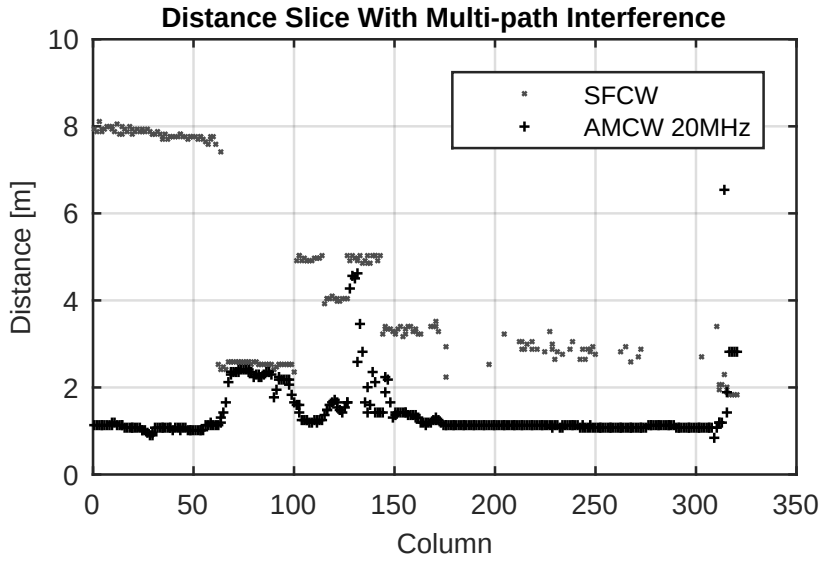


Figure 5.9: Distance slice comparing AMCW and SFCW with a translucent sheet present as measured in figure 5.7 (d) and (h).

rently more accurate than SFCW, with the SFCW oscillating around zero with a maximum error 0.02 m.

Figure 5.12 is a plot comparing the distance error with a different number of samples (number of frequencies sampled over). With more samples the accuracy increases, and with fewer than six samples the accuracy significantly decreases. Currently most ToF range cameras use four raw frames to compute depth, so measuring six or seven for real time depth measurement is possible. In figure 5.12 the reason for the sudden jump in distance error at 1.3 m with six samples is because the spectral estimation is sensitive to the signals phase, and at 1.3 m the measurement at 150 MHz phase wraps which changes the phase of the measured signal at that point.

In SFCW mode the depth can be measured with a similar number of measurements as AMCW, however AMCW ToF range cameras suffer from scene dependent measurement errors caused by phase wrapping and multi-path interference. The results show that the accuracy of SFCW is resilient to phase wrapping and multi-path interference. However, the accuracy of SFCW is dependent on the frequency bandwidth. If the trend of higher operating frequencies continues then the accuracy of operating in SFCW mode will improve and become a viable measurement technique.

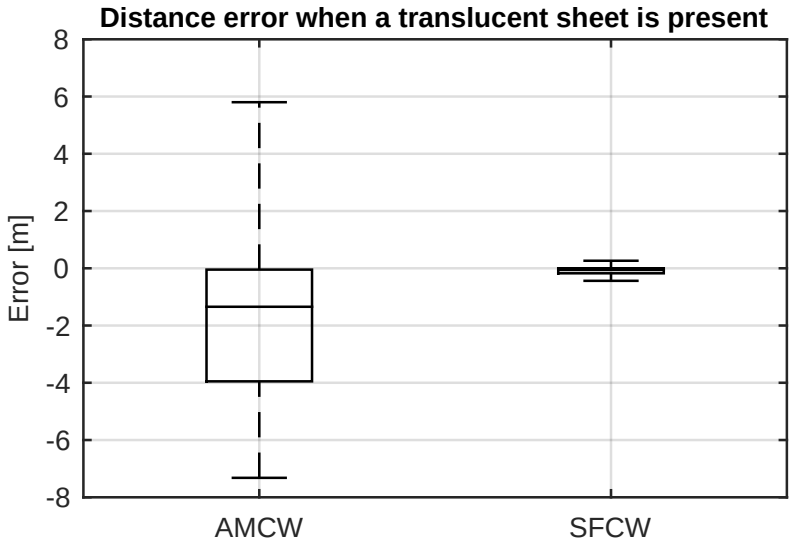


Figure 5.10: Box plot of the distance measurement error of pixels behind a translucent sheet for both the AMCW and SFCW measurements from figure 5.7 (d) and (h). The central mark is the median, the edges of the box are the 25th and 75th percentiles, the whiskers extend to the most extreme datapoints and the outliers have been removed.

5.4 Chirp Sampling

Sampling in phase and frequency simultaneously is possible by changing ω and τ together between raw frames. By setting $\tau = \omega/2\pi$ samples down the diagonal of figure 5.1 can be obtained. The measured signal is

$$h(\omega) = a \cos\left(\frac{\omega^2}{2\pi} + \frac{2\omega d}{c}\right). \tag{5.14}$$

Equation 5.14 is the equation for a linear chirp, with the distance encoded by the initial frequency ω_0 given by

$$d = \frac{c\omega_0}{4\pi}. \tag{5.15}$$

There are two unknown parameters, the amplitude and initial frequency. When multi-path interference is present multiple superimposed chirp signals are present in the measured signal. Superimposed chirp estimation is an area of research (Saha and Kay, 2002).

Figure 5.13 plots the output of the PMD19k3 camera measuring both frequency and time together as described in equation 5.14. Two distances are measured at 0.85 m and 3.63 m. As expected the farther object has a higher

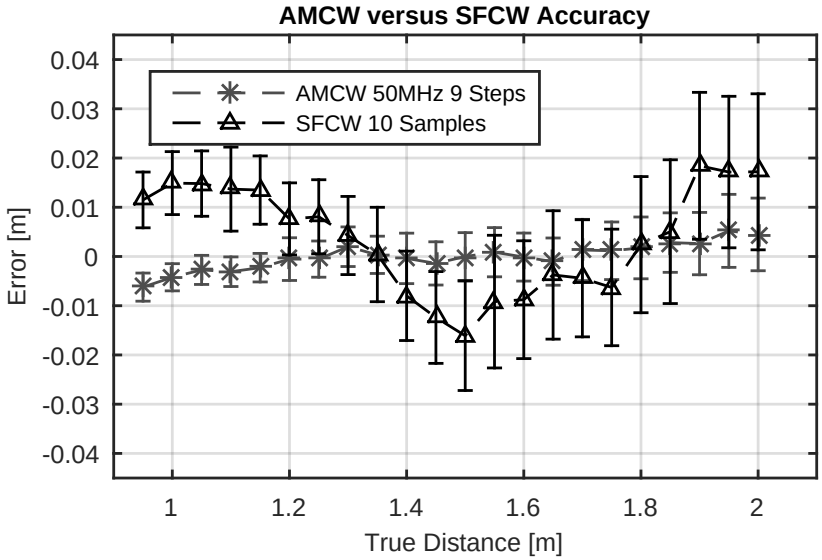


Figure 5.11: SFCW versus AMCW distance accuracy over 1.1 m. The SFCW samples over 150 MHz, in increments of 15 MHz for 10 samples, compared to the nine phase steps taken with the 50 MHz AMCW measurement. The error bars are plus and minus one standard deviation.

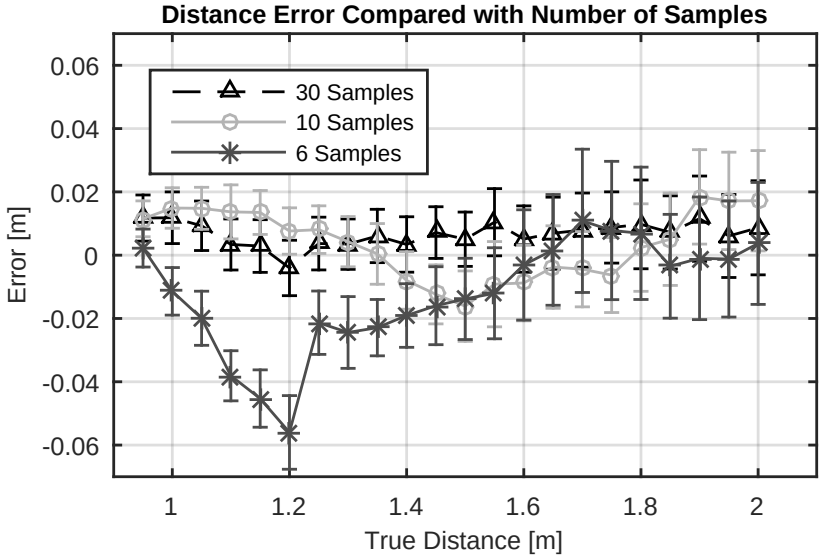


Figure 5.12: The accuracy of measured distance with the number of samples (number of frequencies measured) for the SFCW modulation scheme. The error bars are plus and minus one standard deviation.

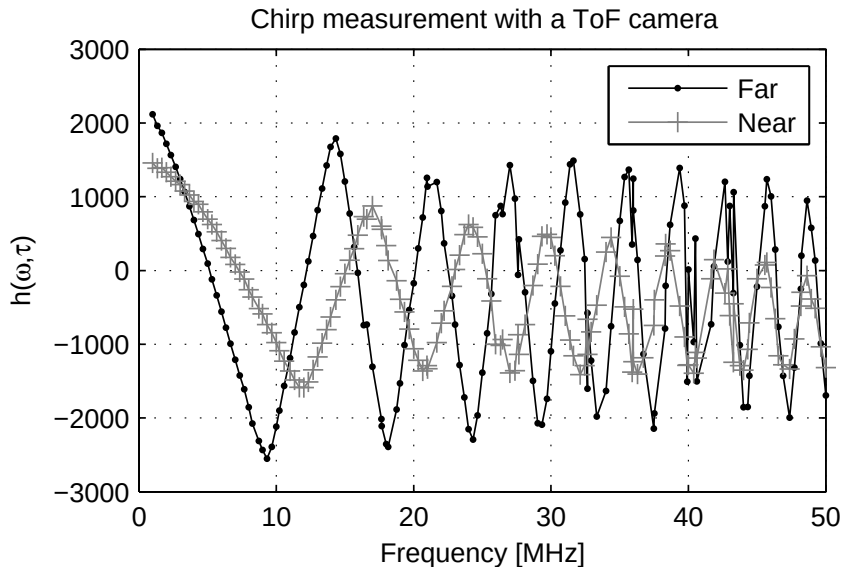


Figure 5.13: Measurements with a time-of-flight range camera with $\omega/2\pi = \tau$ for objects at 0.85 m and 3.63 m. The initial chirp frequency encodes the distance of the object, with closer objects having a lower initial frequency.

initial frequency.

The described chip sampling method to measure distance was not explored further in this thesis. Further work is required to implement chirp estimation algorithms and investigate the properties of this method when multi-path interference is present.

5.5 Comparison and Discussion

In this chapter sampling over frequency and phase has been investigated. Figure 5.1 shows the sampling space in phase and frequency. Sampling over phase is the current method of AMCW operation and it suffers measurement errors from multi-path interference, phase wrapping, and harmonic aliasing.

Sampling over frequency instead of phase is analogous to SFCW radar. The disadvantage is the accuracy of depth measurement depends of the bandwidth of frequencies over which it samples. The trend of increased frequency bandwidth of ToF range cameras is expected to continue, make operating in SFCW mode more appealing.

Sampling over frequency and phase has been used to resolve multi-path interference. The accuracy of the distance measurement does not depend on the frequency bandwidth. However, the minimal resolvable difference between

propagation paths is dependent on the maximum frequency. The minimum number of samples is 15 for five frequencies and three phase steps per frequency. This is compared to sampling over just frequency, the accuracy of the distance measurement depends on the number of samples and frequency bandwidth. The frequency bandwidth also determines the minimum resolvable separation between propagation paths. The advantage of sampling over just frequency is a minimum of six samples is required for the depth measurement. The constraints placed on the number of samples will determine the best sampling method.

Sampling over frequency and phase simultaneously is possible. The measured signal is a chirp. There is no closed form solution for estimating the chirp parameters therefore the usefulness of this method is unknown.

Chapter 6

Time Frequency Duality

Conventional time-of-flight (ToF) range cameras illuminate the scene with an amplitude modulated light source. In chapter 4 replacing the square wave amplitude modulation with a binary sequence, a technique named coded ToF, was explored. In chapter 5 sampling over multiple modulation frequencies, which is similar to stepped frequency continuous wave radar, was investigated. In this chapter a transform between the measurements taken with a binary sequence and a frequency sweep is developed. The transform demonstrates a duality between data collected with a frequency sweep of modulation frequencies and coded ToF. The advantages of collecting data with each method is investigated in this chapter.

6.1 Background and Theory

Here the transform between the measured cross-correlation and a frequency sweep and vice versa is developed. Recall from chapter 4, that the binary sequence, p_γ, that replaces the illumination and reference signal is Γ bits long.

$$p_\gamma \in \{0, 1\} \quad \gamma = 0, 1, .., \Gamma - 1. \tag{6.1}$$

Each bit of the binary sequence has a transmission length of T_c seconds. The index γ of the binary sequence can be expressed as a function of time, mapping the binary sequence from the discrete bit index to the continuous variable time.

$$\gamma(t) = \left\lfloor \frac{t}{T_c} \right\rfloor \bmod \Gamma - 1. \tag{6.2}$$

The simplest and shortest binary sequence is [0 1]. The binary sequences transmission is repetitive, so [0 1] is identical to square wave modulation. The bit transmission time T_c sets the maximum frequency possible in the binary sequence. Higher order odd harmonics do exist, as explained for a square wave in chapter 2, but due to their limited amplitude they are currently ignored. The maximum frequency is

$$\omega = \frac{\pi}{T_c}. \tag{6.3}$$

One basis to represent the binary sequence is with the Fourier transform. The binary sequence contains a discrete number of evenly spaced frequencies, namely

$$P(f) = \mathcal{F}_t[p(t)](f), \tag{6.4}$$

where $\mathcal{F}$ is the Fourier transform. The mathematical forward model of the ToF range camera in the frequency domain is therefore,

$$H(f) = \mathcal{F}_t[p(t) \star ap(t - t_0)]. \tag{6.5}$$

By invoking the correlation and Fourier shift theorems, a time delay is seen to become a phase offset in the frequency domain, that is,

$$H(f) = aP(f)P^*(f)e^{-jm\omega t_0}, \tag{6.6}$$

where $P^*(f)$ is the complex conjugate. The ToF camera's hardware is band-limited which limits the maximum frequency possible in the measured cross-correlation, see figure 3.11. It is well known that a square wave, the most basic form of a binary sequence, can be represented by a infinite sum of discrete odd decaying sinusoids, and when filtered through the camera hardware the output is a finite number of sinusoids. The same is true of binary sequences; they can be accurately represented by a discrete number of sinusoids. By taking the discrete Fourier transform (DFT) over the sampled auto-correlation, $h(\tau)$, it can be expressed as a sum of M frequencies, where each frequency has relative

amplitude, α, and phase, ϕ.

$$H(f) = a \sum_{m=0}^{M} \alpha_m e^{j\phi_m}, \tag{6.7}$$

$$h(\tau) = a \ \mathcal{F}_f^{-1}[H(f)](\tau), \tag{6.8}$$

where the DC offset of equation 6.7 is included when $m = 0$. The correlation signal is composed of M discrete frequencies and Γ bits. From equation 6.3 the number of frequencies that can adequately represent the correlation signal is $M_{max} = \Gamma$. To satisfy the Nyquist-Shannon sampling theorem $2M$ samples are required to recover all frequencies. One bit in the binary sequence has a transmission time of $1/T_c$ seconds, therefore the frequencies of the DFT are at

$$\omega_m = \frac{2\pi}{T_c}\frac{m}{\Gamma} = \omega\frac{2m}{\Gamma}. \tag{6.9}$$

The individual frequencies in equation 6.7 can be decomposed into a complex measurement vector $\mathbf{z_{syn}}$. In chapter 5 the complex measurement vector was used to resolve multi-path interference in ToF range cameras, motivating the following representation.

$$\mathbf{z}_{syn} = [b, a_1\alpha_1 e^{j\phi_1}, \ldots, a_M\alpha_M e^{j\phi_M}]^\top. \tag{6.10}$$

The amplitude, α, of each frequency is inherent to each binary sequence, and therefore can be removed with calibration, leaving the actual amplitude, a, of each frequency. Equation 6.10 demonstrates how a measured correlation signal can be transformed into a measured frequency sweep using the DFT and calibrating for the spectral power of the binary sequence. The significance of the transformation in equation 6.10 is that the processing techniques developed to operate on multiple frequencies can now be applied to measurements recorded with coded ToF.

The inverse transform is possible, taking a frequency sweep and transforming it to a cross-correlation. The inverse transform is applied to a measured frequency sweep of M evenly sampled frequencies. Given a vector of frequency measurements $\mathbf{z}$ at frequencies ω such that

$$\mathbf{z} = [b, a_1 e^{j\phi_1}, \ldots, a_M e^{j\phi_M}]^\top. \tag{6.11}$$

Taking the inverse discrete Fourier transform (IDFT) over the complex values a synthetic correlation signal can be generated.

$$\mathbf{h_{syn}} = \mathcal{F}^{-1}(\mathbf{z}) \tag{6.12}$$

The length of $\mathbf{z}$ determines the length of $\mathbf{h_{syn}}$. The elements of $\mathbf{z}$ can be 0, but the frequencies must be evenly spaced. Zero padding $\mathbf{z}$ generates a longer cross-correlation, however, no more information is added.

Equation 6.11 is the dual of equation 4.9 by the DFT. Data measured in the time domain can be transformed to the frequency domain and vice versa. Techniques designed for multi-path interference removal and phase wrapping in one domain can now be applied to data collected in the other. In this chapter the transformation is explored with the advantages of each measurement domain investigated.

6.1.1 Propagation of Uncertainty

Each measurement of the correlation function contains the true measurement and noise. It is assumed that in a ToF range camera the noise sources are additive white noise and are independent between samples. Schoukens and Renneboog (1986) showed if these conditions are met then the scaled covariance matrix C is

$$C = \frac{2}{N}\sigma_{h_\tau}^2 I, \tag{6.13}$$

where there are N samples, I is the identity matrix, and $\sigma_{h_\tau}^2$ is the variance of each sample.

If the real and imaginary components of the DFT are independent then the variance of the amplitude and phase of each frequency is

$$\sigma_a^2 = \frac{1}{N}\sigma_{h_\tau}^2, \tag{6.14}$$

$$\sigma_\phi^2 = \frac{2}{Na^2}\sigma_{h_\tau}^2. \tag{6.15}$$

To compare the variance of measurements between domains there are two factors: number of samples, and the power of the spectrum. For a frequency sweep all the power is concentrated in one frequency but generally only three or four samples are taken per frequency. For a binary sequence the spectral power is spread across multiple frequencies, however more samples are taken.

6.2 Testing Duality

There are two experimental sections. In the first section the duality between a frequency sweep and using a binary sequence is tested. The second experimental section demonstrates that methods for resolving multi-path interference using a binary sequence and a frequency sweep produce the same results.

6.2.1 Experimental Setup

Three experiments are conducted to test the duality between a frequency sweep and a binary sequence. The first tests the duality and the rest explore the advantages of measuring in each domain.

The first experiment images a flat white foam board sheet. Both the binary sequence [0 0 1 0 1 0 1 1] with a bit time of 1/(60 MHz), and the matching frequencies of 3.75 MHz increasing in increments of 3.75 MHz up to 30 MHz are measured. The binary sequence was selected because of its short length, and the corresponding frequency content of the binary sequence. The corresponding frequencies were calculated using equation 6.9. A flat white foam board sheet was imaged for this experiment. The measured frequencies and amplitudes are calibrated using previously collected data. To statistically compare the frequency sweep and the synthetically generated version a two paired t-test is applied to the amplitude and phase of each frequency, with 100 measurements per frequency. Where the null hypothesis is the means are equal and the alternative hypothesis is the means are different. The null hypothesis is rejected at the 1% significance level.

In the second experiment compares the jitter between a binary sequence and a frequency sweep is compared. A scene with multiple objects at different distances and varying reflectance was imaged. The binary sequence [0 0 0 1 0 1 1 1 0 1] with a bit time of 1/(155.56 MHz) was compared to the three frequencies 7.78 MHz, 23.33 MHz and 38.89 MHz. For a fair comparison the integration period of each raw frame is identical, therefore the photon shot noise is the same. The only difference in the noise sources between captures should be the jitter.

In the last experiment the phase linearity between a frequency sweep and a binary sequence is explored. A flat diffuse white cardboard sheet is imaged, and the frequencies swept from 1 MHz in increments of 1 MHz up to 63 MHz. The frequency sweep is compared to a broad spectrum 100 bit binary sequence.

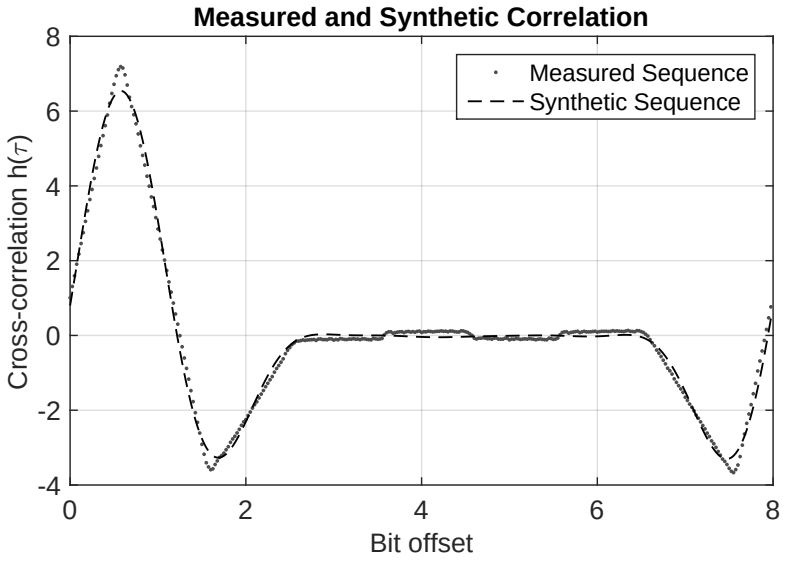


Figure 6.1: Measured correlation signal using a binary sequence and the synthetically generated version from a frequency sweep.

6.2.2 Results and Discussion

Duality

The measured cross correlation of the binary sequence [0 0 1 0 1 0 1 1], and the synthetically generated version from six frequencies is plotted in figure 6.1. The distance travelled and the reflection amplitude is encoded by the correlation peaks phase offset and magnitude respectively. The peak location and magnitude visually match for the frequency sweep and binary sequence, for further analysis the error between the synthetic and measured and plus and minus one standard deviation is plotted in figure 6.2, and overlaid with the measured correlation. The error is systematic over the correlation signal and occurs where the synthetic correlation is smooth compared to the measured. This difference is present for the same reason that the correlation for a square wave is triangular rather than sinusoidal, which is small higher order harmonics are present in the correlation signal. This is demonstrated by comparing the spectrum of the measured and synthetic, as plotted in figure 6.3. The measured correlation spectrum contains more frequencies than the synthetically generated. With the current frequencies used to generate the spectrum it will only be an approximation and not identical.

The measured phase and amplitude of the six frequencies are plotted on an Argand diagram in figure 6.4 and compared to the synthetically generated measurements using a binary sequence. The means and standard deviations

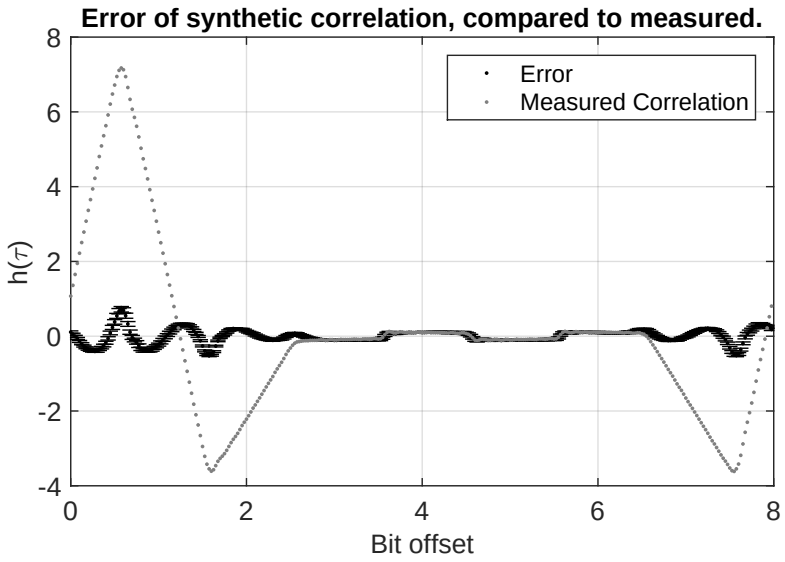


Figure 6.2: The measured correlation and the error between the measured and synthetic in figure 6.1 plotted with plus and minus one standard deviation.

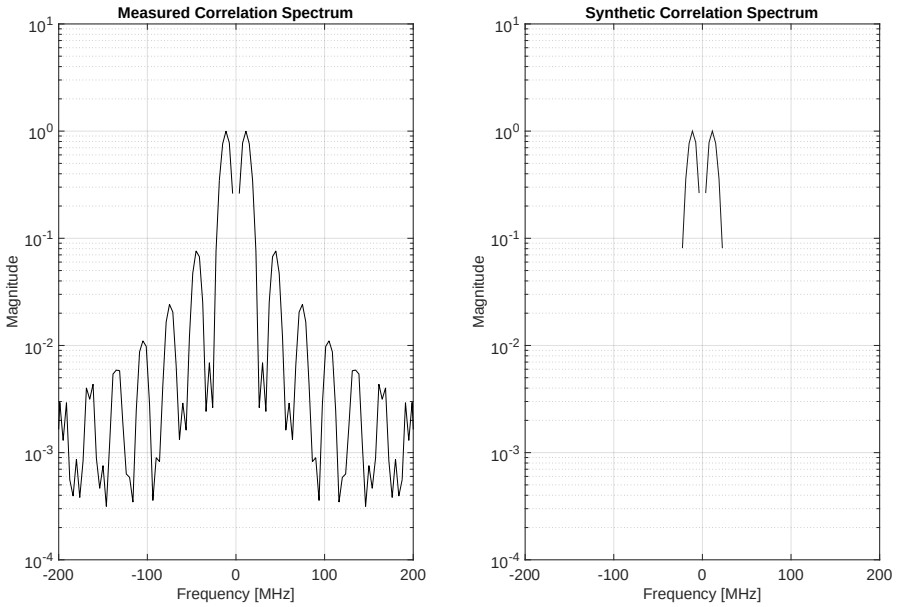


Figure 6.3: Comparison of spectrum of measured correlation and synthetically generated version. The measured spectrum has a number of higher frequency lower amplitude than the synthetic version. These additional frequencies result in the error measured in figure 6.2.

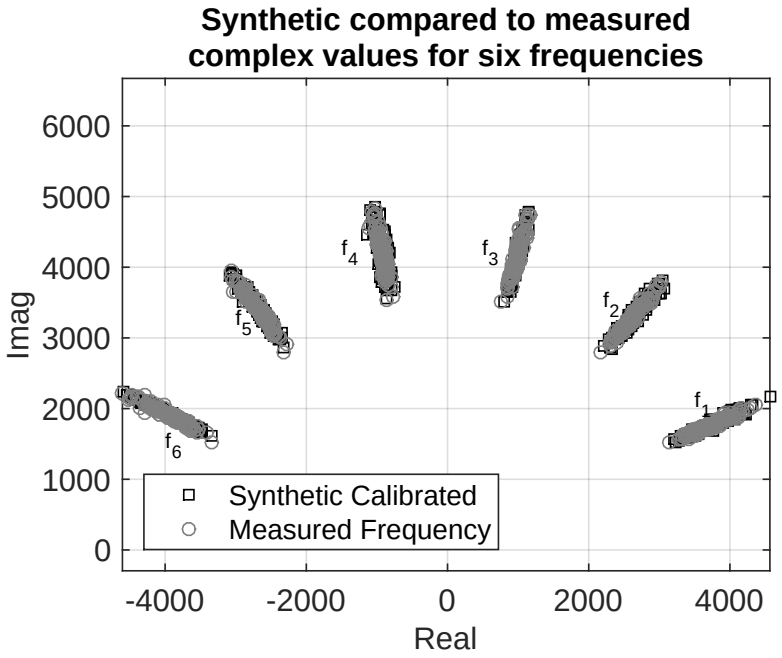


Figure 6.4: Measured complex values for a frequency sweep and synthetically generated version from a binary sequence for six frequencies.

of the amplitude and phase of each measurement is tabulated in table 6.1. Comparison of a single frequency, $f = 11.25$ MHz, for 100 measurements is plotted in figure 6.5, as comparing the results in figure 6.4 is impossible due to significant overlap. The null hypothesis can not be rejected at the 1% confidence level for all six tests comparing the measured phase to the synthetically generated phase. With (P=0.0368, 0.0797, 0.1406, 0.4510, 0.0182, 0.7508) for the six respective frequencies. The null hypothesis can not be rejected at the 1% confidence level for the six amplitude measurements with (P=0.6914, 0.8922, 0.9732, 0.6802, 0.5092, 0.4579). The measured and synthetic phase and amplitude values after calibration can not reject the null hypothesis, that the means are equal, at the 1% confidence level. This indicates that the measured phase and amplitude for the frequency sweep and synthetic version are the same. There is a difference between the generated correlation from six frequencies and the measured due to the missing frequencies. The affect of the differences is explored in section 6.3 when multi-path interference correction methods are applied to measured and synthetic correlation signals.

Jitter

Jitter limits the minimum phase variance achievable, the more jitter present the higher the minimum phase variance. Comparison between the frequency

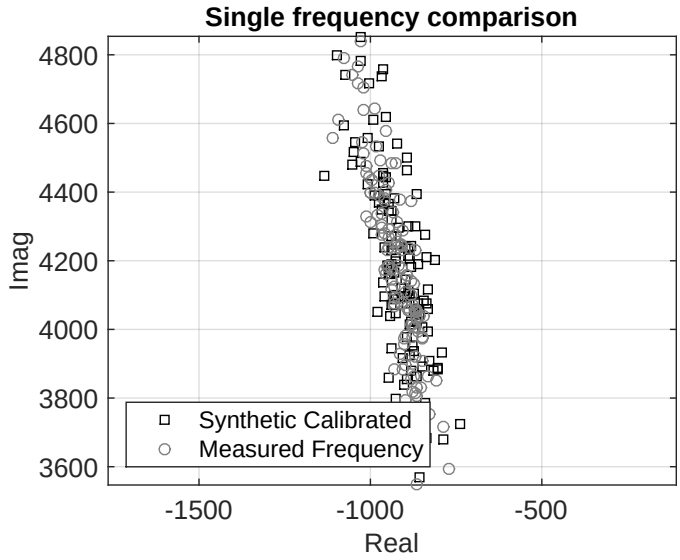


Figure 6.5: Measured 100 complex values for a single frequency comparing measured and synthetically generated values.

Table 6.1: Phase standard deviations for a frequency sweep and synthetic version for 100 measurements.

	Measured Phase [rad]	**Synthetic Phase [rad]**
f_1	0.442 ± 0.007	0.442 ± 0.007
f_2	0.898 ± 0.007	0.896 ± 0.009
f_3	1.338 ± 0.006	1.337 ± 0.005
f_4	1.788 ± 0.007	1.786 ± 0.010
f_5	2.238 ± 0.007	2.237 ± 0.007
f_6	2.691 ± 0.008	2.689 ± 0.003

sweep and binary sequence phase variance versus amplitude is shown in figures 6.6 and 6.7. The minimum phase variance increases with frequency for the frequency sweep, and the minimum phase variance is similar for each synthetic frequency from the binary sequence. The minimum phase variance for the two higher synthetic frequencies in the binary sequence is lower than the frequency sweep. This result indicates that jitter in these higher frequencies has been reduced by using a binary sequence.

The modulation frequency for the PMD19k camera is generated by a phase lock loop (PLL) in the FPGA. The PLL uses the reference clock as an input to a voltage controlled oscillator (VCO), where the VCO is down sampled to generate each modulation frequency. The jitter period, J_{PER}, at the VCO is constant therefore it is constant between modulation frequencies. As the frequency decreases the phase noise decreases. There should be a linear rela-

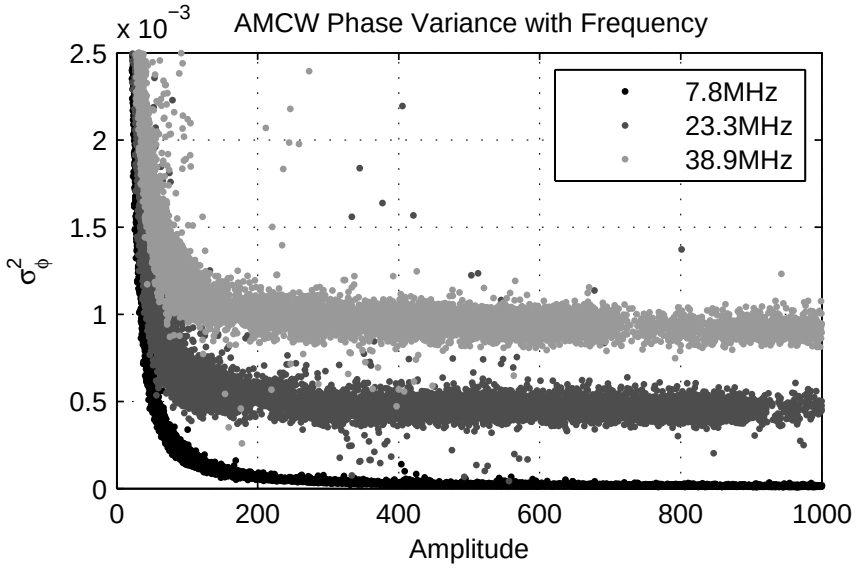


Figure 6.6: The change in phase variance with amplitude for varying modulation frequencies.

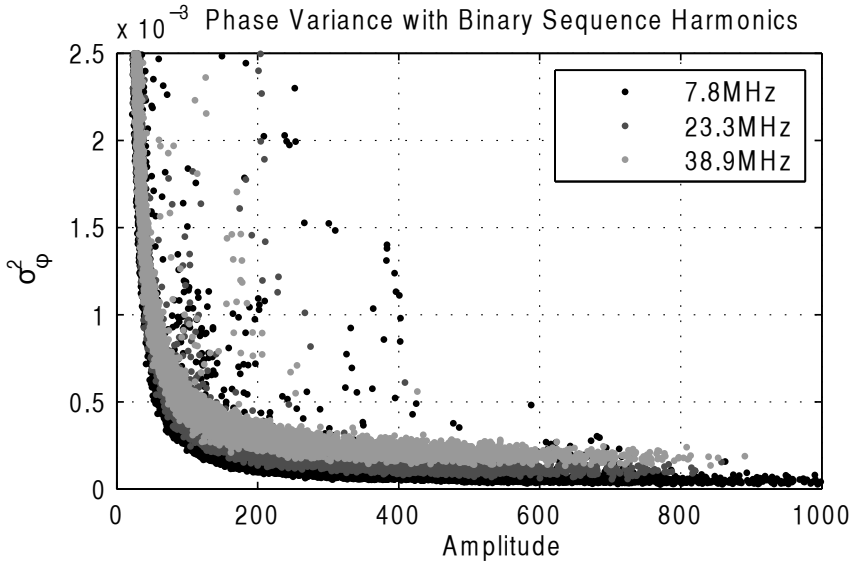


Figure 6.7: The change in phase variance with amplitude for different harmonics of a binary sequence.

tionship between modulation frequency and the variance limit achieved due to jitter.

$$\sigma^2_{\phi_{min}} \propto \omega. \tag{6.16}$$

Equation 6.16 is demonstrated in figure 6.6. There is a linear increase in the minimum variance with a linear increase in modulation frequency.

When using a binary sequence the bit period, T_c, changes during the sequence. Instead the average period T_{avg} is the average expected period over the sequence. This will give an estimate of the expected jitter for each harmonic when using a binary sequence. However in a binary sequence each edge does not contribute evenly to each harmonic therefore there are differences in jitter between each frequency as shown in figure 6.7.

Ideally the relationship between modulation frequency and measured phase is linear when no multi-path interference is present. However, the camera hardware acts as a low pass filter with an amplitude and phase response. The phase response is linear below a cut-off frequency. Using a binary sequence the linear phase response region can be extended. The error in phase from a linear fit is plotted as a function of frequency in figure 6.8. The phase response is linear between 0 and 30 Mhz for the frequency sweep. Between 30 Mhz and 50 Mhz there is an increase in error and above 52 MHz the phase response is no longer linear. The error in phase for the synthetic frequencies does not depend on frequency, extending the linear operating region of the camera. This is significant because the OMP method for resolving multi-path interference presented in chapter 5 is dependent on a linear amplitude and phase response. Collecting data with a binary sequence is easier as less calibration is required.

6.3 Multi-path Interference Restoration

Multi-path interference is a problem in ToF range cameras, and the previous chapters have described a number of techniques developed to resolve multi-path interference. In this section data captured in both domains is transformed to resolve multi-path interference with the technique developed for the other domain.

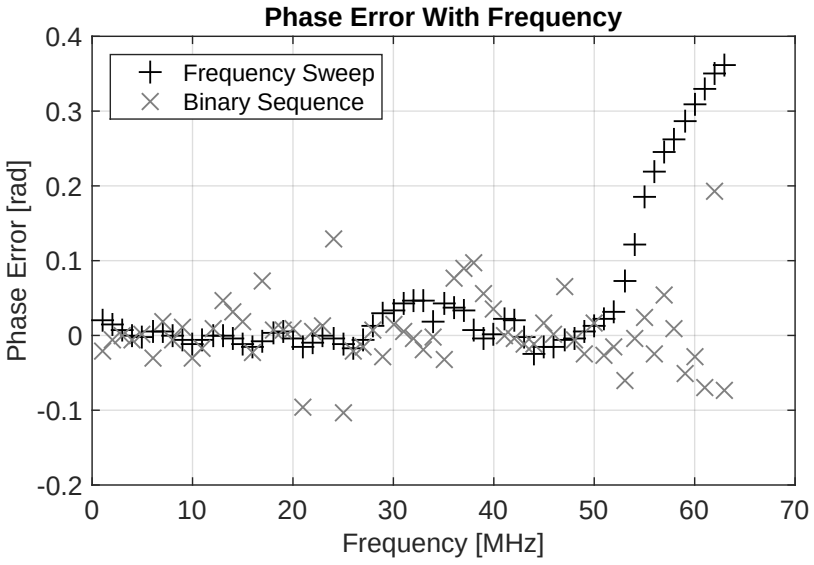


Figure 6.8: Phase error with modulation frequency comparing a frequency sweep and a binary sequence.

6.3.1 Experimental Setup

To test the duality between a binary sequence and a frequency sweep for resolving multi-path interference a translucent sheet was placed 2.5 m in front of a diffuse white foam board. The 23bit binary sequence [0 0 1 0 0 1 1 0 1 0 1 1 0 1 0 1 0 1 1 1 0 0 0] was used because it has the smallest variance in the power spectrum over the desired frequencies, as described by equation 4.20. The corresponding frequency sweep starts at 4.35 MHz, increases in increments of 4.35 MHz up to 56.52 MHz. The technique presented in chapter 4 for resolving multi-path interference using sparse deconvolution on the measured cross-correlation is compared to the OMP technique presented in chapter 5 using measurements at multiple modulation frequencies. Data collected in each domain is transformed to the other and the corresponding multi-path interference restoration technique is tested. Each method for resolving multi-path interference outputs four variables, the amplitudes and distances of the two propagation paths. The paired t-test is applied to the distance outputs of each method, and the ratio of the amplitudes. The null hypothesis is the means are equal, and the alternative hypothesis is the means are different. The null hypothesis is rejected at the 1% significance level.

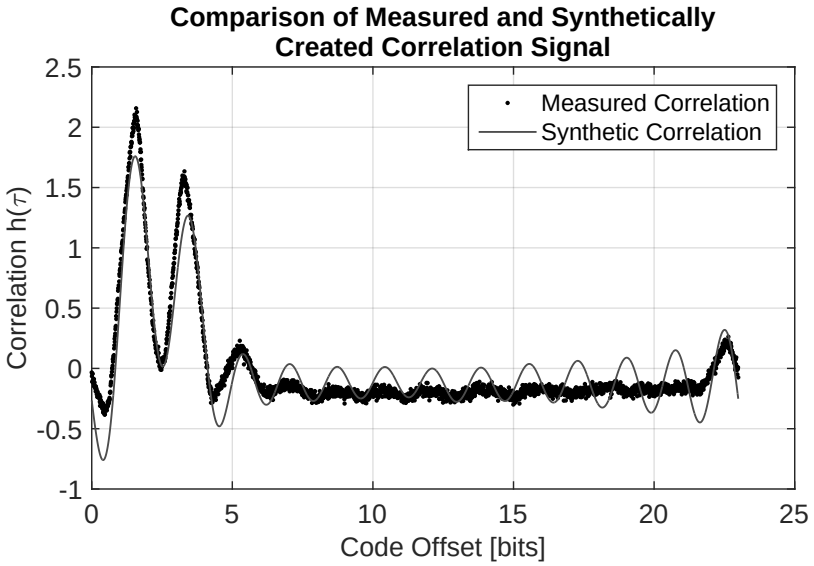


Figure 6.9: Comparison between the measured correlation function using a binary sequence and the synthetically generated correlation function using a frequency sweep when multi-path interference is present. The frequency sweep is from 4.35 MHz to 56.52 MHz while the code supports up to 65.22 Mhz, this explains some of the differences.

6.3.2 Results and Discussion

The measured correlation and the synthetically generated correlation when two propagation paths are present is plotted in figure 6.9. The location and size of the largest peaks encode the distance and amplitude of each return. The ringing in the frequency sweep is caused by reduced spectral content as the higher order harmonics caused by using square waves are missing.

Table 6.2 compares the restoration of the multi-path interference between the frequency sweep and binary sequence and the synthetically generated versions. The measured value is measured 100 times to estimate the uncertainty of the measurement, with the plus and minus being one standard deviation of the measurement. The purpose of this experiment is to prove the measurements and processing techniques are equivalent. The ratio between the two measured amplitudes a_1 and a_2 encodes the relative strength of each return. The ratios are tabulated in table 6.3.

The paired t-test results for the output of the OMP algorithm is as follows: At the 1% significance level the null hypothesis can not be rejected for the measured d_1 with (P=0.0442). While for the measured d_2 the alternative hypothesis is true (P=$2.26e^{-12}$). The t-test for the amplitude ratios the alternative hypothesis is true (P=0.0023). While the null hypothesis has been

Table 6.2: Measured amplitude and phase of each return for the coded ToF and OMP multi-path interference restoration techniques.

Data	a_1	d_1[m]	a_2	d_2[m]
Measured Frequency	313.9 ± 24.2	0.65 ± 0.13	283.2 ± 25.2	3.07 ± 0.13
Synthetic Frequency	303.6 ± 26.8	0.68 ± 0.13	292.9 ± 24.6	3.22 ± 0.15
Measured Sequence	312.3 ± 8.4	0.83 ± 0.07	297.7 ± 6.7	3.10 ± 0.07
Synthetic Sequence	248.9 ± 21.8	0.84 ± 0.26	237.2 ± 18	3.17 ± 0.29

Table 6.3: Amplitude ratios from table 6.2

Data	a_1/a_2
Measured Frequency	0.90 ± 0.025
Synthetic Frequency	0.97 ± 0.038
Measured Sequence	1.05 ± 0.039
Synthetic Sequence	1.06 ± 0.119

rejected for some of the measurements all the output results, tabulated in table 6.2, are within two standard deviations of each other.

The paired t-test results for the output of the coded ToF algorithm is as follows: At the 1% significance level the d_1 and d_2 measurements can not reject the null hypothesis with (P=0.822) and (P=0.0207) respectively. The amplitude ratio can not reject the null hypothesis with (P=0.6488). Indicating the recovered distance and amplitude ratio is the same between the measured cross-correlation and the synthetically generated version. This is despite the systematic differences in the cross-correlation as plotted in figure 6.2.

The results presented in tables 6.2 and 6.3 with the combined paired t-testing strongly indicates measurements taken either with a frequency sweep or with a binary sequence can be transformed into the alternative and the output of subsequent processing is the same.

6.4 Discussion

In this chapter the duality between data collected with a binary sequence and a frequency sweep was experimentally verified. The results for resolving multi-path using sparse deconvolution on the correlation signal and OMP on the measured complex data produce the same results. The correlation signal can be described by its frequency content, once the Nyquist sampling criteria have been satisfied no more information is gained from more samples on the

correlation signal. However, the sparse deconvolution quantizes the depth to the number of measurements. It is possible to interpolate to increase the depth resolution, but this increases the computational complexity of the sparse deconvolution. The OMP method for resolving multi-path interference has a lower computational complexity than the sparse deconvolution. However, the phase linearity with frequency is important for the OMP method for correct results, as demonstrated in figure 5.5, and a binary sequence has improved linearity, as demonstrated in figure 6.8. It is recommended collecting data with a binary sequence and resolving multi-path interference using OMP.

6.4.1 Binary Sequence Advantages

With a binary sequence all the frequencies are measured simultaneously. Therefore the jitter in each of the measurements is approximately equal as shown in figure 6.7. Simultaneous measurements are also advantageous for resolving motion artifacts as the corruption is the same in all frequencies. As shown in figure 6.8 a binary sequence can increase the linear operating region of the camera. Allowing for high operating frequencies with less calibration.

The power dissipation of electronic devices is related to the switching frequency, the faster the switching the more power dissipated. This is true for ToF range cameras, the temperature increases with modulation frequency. Temperature change is problematic for multi-frequency captures as the temperature change can introduce measurement errors. When operating a coded ToF camera the power dissipation is constant over the integration period, keeping the temperature constant between raw frames.

6.4.2 Frequency Sweep Advantages

The advantage of a frequency sweep is any possible selection of modulation frequencies is possible. The selection of frequencies with a binary sequence is limited by the possible sequences. In situations where non-evenly spaced frequencies is required a frequency sweep is ideal. With a frequency sweep only three raw samples per frequency is required. For a binary sequence the Nyquist sampling criteria stipulates the number of samples required, the larger the bandwidth of the sequence the more samples required. With a frequency sweep no power is wasted on unused harmonics.

Chapter 7

Spatial Coding

In the previous chapters the modulation signal of the time-of-flight (ToF) range camera was coded in time and in frequency to resolve multi-path interference. Currently ToF range cameras fully illuminate the field of view of the camera with amplitude modulated light. In this chapter spatially modulating the illumination signal is explored to resolve multi-path interference.

The theory of light transport describes how light travels from its source to the viewer, where often the light takes multiple paths to the viewer. The light transport can be broken down into two components: a direct return which is the shortest path, and a global return which encompasses everything else. Practically direct and global separation is performed by spatially modulating the illumination source. In this chapter direct and global separation techniques for conventional cameras is combined with ToF range cameras to resolve multi-path interference. The background of direct global separation is first reviewed, then is adapted and combined with ToF range imaging such that the direct and global components are separated using Fourier analysis.

The light transport in a scene was originally measured by Seitz *et al.* (2005) by raster scanning a laser point over the entire scene, and measuring the interaction between the illuminated point with all the other points in the scene. The light transport can then be broken down into K discrete returns. Light transport was further developed into two return components, with the direct return being the shortest path length, and the global return containing everything else. Nayar *et al.* (2006) demonstrated a fast method for separating the direct and global components by illuminating the scene with a checkerboard pattern, and phase shifting the pattern between multiple frames. Nayar *et al.* (2006)

proved the direct and global components can be recovered if the spatial frequency of the global return is much lower than the spatial frequency of the projected pattern. Global components of light propagation cause issues for structured light range cameras. The development of structured light patterns to remove the global effects has been an area of research by Gupta *et al.* (2013) and Gu *et al.* (2011). A direct ToF range camera was used to measure the light transport in a scene by Wu *et al.* (2014). Wu's method relied on the direct measurement of the time-of-flight of light, therefore the multi-path interference problem was trivial to solve. Recently O'Toole *et al.* (2014a) applied direct and global seperation to ToF range cameras for transient imaging.

7.1 Direct and Global Component Separation

The purpose of the direct and global separation is to decompose the measured radiance, L, at each pixel into a direct component, L_d, and a global component L_g, namely,

$$L = L_d + L_g. \tag{7.1}$$

7.1.1 Checker Board Illumination Pattern

Consider a scene illuminated with a checkerboard pattern so that patches of the scene are directly illuminated and an even number of patches are not illuminated. In the illuminated regions the pixels measure (Nayar *et al.*, 2006)

$$L_{max} = L_d + \frac{1}{2}L_g. \tag{7.2}$$

As half the scene is illuminated compared to full field illumination, then only half the global is present. When illuminated with the inverse checkerboard pattern, the non-illuminated and illuminated regions invert, and the non-illuminated regions measure

$$L_{min} = \frac{1}{2}L_g, \tag{7.3}$$

under the assumption that the global component has a lower spacial frequency than the projected checkerboard. The set of equations 7.2 and 7.3 are well

conditioned. The direct and global components can be calculated by

$$L_g = 2L_{min}, \tag{7.4}$$

$$L_d = L_{max} - L_{min}. \tag{7.5}$$

Ideally each pixel of the camera would measure the directly illuminated and non-illuminated values, however due to non-ideal projector and camera integration the edges measure a combination of the illuminated and non-illuminated. To overcome this problem Nayar *et al.* (2006) proposed shifting the checker board pattern 25 times, for five shifts in the x direction and five shifts in the y direction. Nayar *et al.* (2006) also discussed the minimum spatial illumination frequency (pattern size) required to perform the separation, assuming that the global function at each point is smooth compared to the illumination frequency. This means that the global component cannot be caused by highly specular inter-reflections.

7.1.2 Sinusoidal Illumination Pattern

Using conventional cameras it is possible to resolve the direct and global components with three frames as shown by Nayar *et al.* (2006). By illuminating the scene with the high frequency sinusoidal pattern,

$$P(x, y) = \sin\left(\frac{2\pi x}{p} + \upsilon + \sin\left(\frac{2\pi y}{q}\right)\right), \tag{7.6}$$

where P is the projected image, x and y are the projector coordinates, and υ is the phase shift of the pattern in pixels per frame in the x dimension. The variables p and q change the frequency of the pattern in the x and y direction respectively. The light source has a maximum amplitude, A, and the projected sinusoidal pattern has a spatial frequency of $f_x = x/p$, cycles/pixel. To perform separation the pattern is phase shifted in the x direction between frames. The radiance each pixel measures over time (each frame) is

$$L(t) = \frac{A \cdot (1 + \sin(\omega_p t + \theta))}{2}, \tag{7.7}$$

where $\omega_p = 2\pi f_x \upsilon$ is the angular frequency of the projected pattern. Since the global return is spatially smooth its contribution to the radiance is constant between projected images, therefore it is constant with time. The measured

radiance over time is therefore

$$I(t) = L_d \frac{(1 + \sin(\omega_p t + \theta))}{2} + \frac{L_g}{2}, \tag{7.8}$$

$$= \frac{L_d}{2} \sin(\omega_p t + \theta) + \frac{L_d + L_g}{2}. \tag{7.9}$$

By expansion of the trigonometric functions and rearrangement, equation 7.9 becomes

$$I(t) = \eta \sin(\omega_p t) + \beta \cos(\omega_p t) + \frac{\gamma}{\sqrt{2}}, \tag{7.10}$$

where,

$$\eta = \frac{L_d \cos(\theta)}{2}, \tag{7.11}$$

$$\beta = \frac{L_d \sin(\theta)}{2}, \tag{7.12}$$

$$\gamma = \frac{(L_d + L_g)}{\sqrt{2}}. \tag{7.13}$$

The unknowns η, β, and γ can be extracted by taking the discrete Fourier transform (DFT) over t. The direct and global components are calculated by

$$L_d = 2\sqrt{\eta^2 + \beta^2}, \tag{7.14}$$

$$L_g = \sqrt{2}\gamma - L_d. \tag{7.15}$$

There are three unknowns, θ, L_g, and L_d, therefore three, or more, images are required for a closed form solution.

7.2 Combination of Time-of-Flight and Direct and Global

This section describes the theory of resolving the direct and global components in ToF range images by combining the spatially varying illumination patterns with the temporal varying signal in ToF range imaging. Both checkerboard and sinusoidal patterns are analyzed.

The multi-path interference problem from equation 2.19 can be written in terms of the direct and global components, where the direct return is the shortest path from the light source to sensor and the global return includes all other returns. Thus,

$$\zeta = a_d e^{j\phi_d} + a_g e^{j\phi_g}, \tag{7.16}$$

where a_d and ϕ_d are the amplitude and phase of the direct component respectively, and a_g and ϕ_g are the amplitude and phase of the global component.

When using a high frequency checkerboard pattern as the illumination pattern, equations 7.2 to 7.5 can still be used with complex components if certain assumptions are made. Radiance measurements are positive scalars and always combine constructively, therefore equation 7.2 is always greater than equation 7.3. ToF range cameras measure complex values the maximum and minimum values are estimated by

$$L_{max\_est} \simeq \max |L(t)|, \tag{7.17}$$

$$L_{min\_est} \simeq \min |L(t)|. \tag{7.18}$$

The maximum and minimum complex value can be approximated by the absolute minimum and maximum. The approximations in equations 7.17 and 7.18 are valid provided that

$$|L_g + L_d| > |L_g|. \tag{7.19}$$

The constraint in equation 7.19 is broken down into two constraints that are visualized in figure 7.1.

$$|L_g| << |L_d|, \tag{7.20}$$

or

$$|\phi_g - \phi_d| \leq \pi/2. \tag{7.21}$$

The direct and global components add on the complex plane, as demonstrated in figure 7.1. It is possible that the direct and global add together destructively, if this occurs then equations 7.17 and 7.18 can be invalid. To guarantee

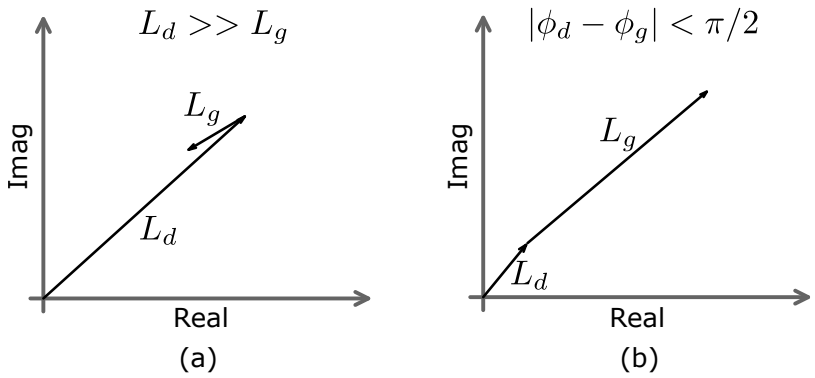


Figure 7.1: The direct and global components are vectors on the complex plane. To satisfy the constraint that $|L_g + L_d| > |L_g|$ results in (a) or (b) being true.

equations 7.17 and 7.18 are valid then either the direct amplitude is much greater than the global amplitude, so even if they add together destructively the summation is always greater than the global by itself, this results in the constraint in equation 7.20. Or the direct and global add constructively, resulting in equation 7.21.

Separation with no constraints and the minimum number of raw frames is desired, ideally allowing for 30 fps of depth data free from corruptions caused by multi-path interference. The general form can be expressed as the combination of equation 7.8 and equation 2.4, leading to

$$I(t,\tau) = \frac{h_d(\tau)}{2} \cdot L(t) + \frac{h_g(\tau)}{2}, \tag{7.22}$$

where h_d and h_g are the correlation signals for the ToF range camera for the direct and global components, and $L(t)$ is the projected pattern with time. It is desired to find a function $L(t)$ that minimizes the number of samples required to separate out the components, and is robust to noise. Ideally no additional constraints would be placed on the unknown variables. If we can only detect multi-path interference then we still obtain a valuable result, as no data is better than wrong data, but correction of multi-path interference is the ideal outcome.

The amplitude and modulation frequency are the two parameters of a ToF range camera that can be modified. Multiple modulation frequencies can be

included into equation 7.22. The resulting equations become:

$$\phi(\omega) = \frac{2\omega d}{c}, \tag{7.23}$$

$$h(\tau, \omega) = \frac{a}{2}\cos(\omega\tau + \phi(\omega)), \tag{7.24}$$

$$I(t, \tau, \omega) = \frac{h_d(\tau, \omega)}{2} \cdot L(t) + \frac{h_g(\tau, \omega)}{2}. \tag{7.25}$$

A function $L(t)$ that makes the problem well conditioned so the inverse is possible is desired. This is discussed further in chapter 8.

A form of $L(t)$ which we will show leads to a closed form solution is a cosine wave physically realised by the projection of the high frequency sinusoidal pattern from equation 7.7,

$$L(t) = \frac{1 + \cos(\omega_p t - \theta)}{2}, \tag{7.26}$$

and its frequency by

$$\omega_p t = l\omega\tau, \tag{7.27}$$

where l is a positive integer. Equation 7.22 becomes

$$I(\tau) = \frac{a_d}{4}\Big(\cos((l-1)\omega\tau + \theta - \phi_d) + \cos((l+1)\omega\tau + \theta + \phi_d)\Big) + \frac{a_d\cos(\omega\tau + \phi_d) + a_g\cos(\omega\tau + \phi_g)}{2}. \tag{7.28}$$

Equation 7.28 is the result of mixing the ToF phase shifting modulation signal with the phase shifting projected sinusoidal pattern. Due to the mixing the direct return is shifted into the higher frequencies, and with Fourier analysis over τ the amplitude and phase can be recovered. A value for l is desired such that the problem is invertible in the least number of samples of τ. In table 7.1 is enumerated for integer values of l if the problem is solvable using Fourier analysis and the number of samples required.

With $l = 3$ a closed form solution is present with the fewest samples. With α_k and φ_k being the amplitude and phase at the k^{th} Fourier frequency and

Table 7.1: Summary of solvability of equation 7.28 using Fourier analysis for different values of l. The number of samples is calculated by the Nyquist frequency.

l	0	1	2	3	4	5
Solvable	No	No	No	Yes	Yes	Yes
Samples	-	-	-	9	11	13

$l \geq 3$ then

$$\theta = \frac{\varphi_{l-1} + \varphi_{l+1}}{2}, \tag{7.29}$$

$$\phi_d = \frac{\varphi_{l+1} - \varphi_{l-1}}{2}, \tag{7.30}$$

$$a_d = 2(\alpha_{l-1} + \alpha_{l+1}), \tag{7.31}$$

and,

$$\frac{a_g}{2}\exp(-j\phi_g) = \alpha_1 \exp(-j\varphi_1) - \frac{a_d}{2}\exp(-j\phi_d). \tag{7.32}$$

There are no additional constraints on the direct and global components, only the already existing constraint that the global component is bandlimited in the spatial Fourier domain.

7.2.1 Error Analysis

In the analysis above we neglected the well recognized fact that unwanted harmonics add error to ToF range cameras (Payne *et al.*, 2008). We now consider the harmonics and their impact on the direct and global separation using sinusoidal spatial modulation. These unwanted harmonics arise from the common hardware implementation that utilizes square waveform modulation driven by digital circuitry, and by nonlinearities in the hardware, as discussed in chapter 2. The correlation function is therefore a triangle waveform and higher order odd harmonics are present in the signal. Under ideal square wave

temporal modulation the odd harmonics have the following effect,

$$h(\tau) = \sum_{\upsilon=1,3,5,\ldots}^{\Upsilon} \frac{a}{2\upsilon^2} \cos(\upsilon\omega\tau + \upsilon\phi), \tag{7.33}$$

$$I(\tau) = \left[\sum_{\upsilon=1,3,5,\ldots}^{\Upsilon} \frac{a_d}{2\upsilon^2} \cos(\upsilon\omega\tau + \upsilon\phi_d) \right] \cos(l\omega\tau - \theta) + \frac{h_g(\tau) + h_d(\tau)}{2}, \tag{7.34}$$

$$I(\tau) = \left[\sum_{\upsilon=1,3,5,\ldots}^{\Upsilon} \left(\frac{a_d}{4\upsilon^2} \cos((\upsilon+l)\omega\tau + \upsilon\phi_d - \theta) + \frac{a_d}{4\upsilon^2} \cos((\upsilon-l)\omega - \upsilon\phi_d - \theta) \right) \right] + \frac{h_g(\tau) + h_d(\tau)}{2}. \tag{7.35}$$

The third harmonic of the triangular modulation signal is mixed onto the DC and sixth harmonic of the new correlation. In the case where nine samples are used the sixth harmonic is aliased onto the third harmonic which is not used to compute the direct component. However the fifth harmonic is mixed onto the second and eighth harmonics. Then the eighth harmonic is aliased onto the first harmonic. This causes errors in the direct and global computation. The presence of the third harmonic does not cause problems but the fifth one does. The fifth harmonic has a relative amplitude of 1/25 compared to the first, therefore the fifth has limited effect.

7.2.2 Propagation of Uncertainty

Each measurement of the correlation function contains the true measurement and noise. In ToF range cameras this consists of photon shot noise, dark current, ADC quantization and jitter. It is assumed that the resulting noise is white and independent between samples. Schoukens and Renneboog showed if these conditions are met then the scaled covariance matrix C for the discrete Fourier transform is (Schoukens and Renneboog, 1986)

$$C = \frac{2}{N} \sigma_{h_\tau}^2 I, \tag{7.36}$$

where N is the number of samples, I is the identity matrix and $\sigma_{h_\tau}^2$ is the variance of the additive white noise.

Assuming the real and imaginary components are independent then it can be shown that the variance of the amplitude a and phase ϕ for each frequency of the DFT is

$$\sigma_a^2 = \frac{1}{N}\sigma_{h_\tau}^2, \tag{7.37}$$

$$\sigma_\phi^2 = \frac{2}{Na^2}\sigma_{h_\tau}^2. \tag{7.38}$$

Equations 7.37 and 7.38 are the generalisation from four samples, given by Frank *et al.* (2009), for N samples. The increase in variance in the direct phase and amplitude estimation from the standard ToF measurement is calculated, given that due to the spread of the amplitude across multiple frequencies due to mixing the amplitude is

$$\alpha_{l-1} = \alpha_{l+1} \approx \frac{a_t}{4}, \tag{7.39}$$

where a_t is the total amplitude. Using the propagation of uncertainty through equations 7.30 and 7.31 the variance compared to the total can be computed as

$$\sigma_{a_d}^2 = \frac{8}{N}\sigma_{h_\tau}^2, \tag{7.40}$$

$$\sigma_{\phi_d}^2 = \frac{16}{Na_t^2}\sigma_{h_\tau}^2. \tag{7.41}$$

The variance of the direct phase and amplitude estimation increases by eight times compared to using the standard time of flight technique. The variance in the global component will be greater because the global is calculated from the direct.

7.3 Experimental Setup, Results and Discussion

A variety of scenes are imaged to test the performance of the proposed direct and global separation methods. The scenes imaged are a corner, wax and peppers, a checkerboard on a flat sheet, a translucent sheet, and a mannequin. Each scene is designed to cause one form of multi-path interference. The corner causes inter-reflections, the wax and peppers cause subsurface scattering, and imaging a checkerboard exposes lens flare. In each scene the checkerboard and

sinusoidal direct and global separation methods are used. The checkerboard method requires 100 raw frames, which provides 25 depth frames, the sinusoidal method uses nine raw frames.

Multi-path interference caused by inter-reflections in corners has been difficult to resolve. ToF camera manufacturers have advised against imaging corners (Mesa Imaging, 2010), and previous research into using multiple modulation frequency measurements have failed (Kadambi *et al.*, 2013). The only methods that have correctly recovered a corner are other direct and global separation methods (O'Toole *et al.*, 2014a) and intensive computational techniques that make assumptions about the scene (Jiménez *et al.*, 2012).

The results of the direct and global separation with a projected checkerboard pattern and a sinusoidal pattern are presented for a variety of test cases.

7.3.1 Corner

To test the accuracy of the direct and global separation on complex data the ground truth distance is required. To acquire the ground truth of the internal corner, each side is illuminated with small patches with the opposing side removed. The illumination of small patches and removal of scattering objects means that we can assume that the measured return is the direct return. The purpose of this test is to demonstrate that the direct measurements are more accurate than the full field, and to statistically show this. The average of 100 ground truth phase images is computed to obtain a single reference image. The reference image is subtracted from each of the one hundred full field, checkerboard, and the sinusoidal phase images. The root mean squared error (RMSE) for each phase frame is calculated. Significance testing is performed by the Wilcoxon signed rank paired test for difference in the median of the RMSE between the different methods.

In figure 7.2 we compare the measured phase of one row in the corner when using full-field illumination, the recovered direct distance component and the ground truth distance measurement. The direct component is consistent with the ground truth and align within three standard deviations of each other. The measured phase in the corner when using full field illumination is consistent with results shown by Fuchs (2010).

The mean and standard deviation of the phase RMSE over 100 frames for the three measurement techniques is tabulated in table 7.2. Each RMS value is reported is separated by more than three standard deviations. The result of

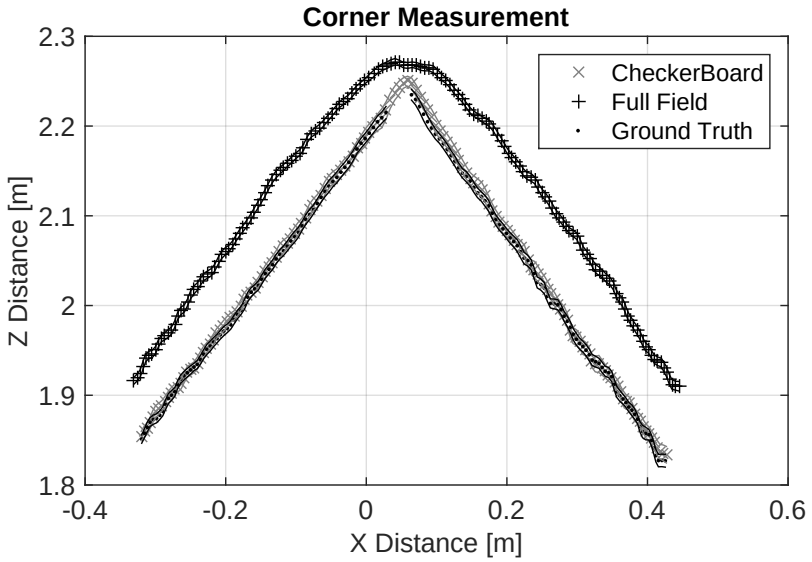


Figure 7.2: Comparison of measured distance when using full field illumination and the reconstructed direct component using checkerboard patterns compared to the ground truth, with plus and minus three standard deviations plotted around each measurement.

the Wilcoxon signed rank test between full field and checkerboard illumination, and between checkerboard and sinusoidal illumination, both have p-values that are zero to machine precision, which supports the assertion that the direct and global separation is successful.

Table 7.2: Mean and standard deviation of phase RMSE over 100 frames for full field (FF), checkerboard (CB), and sinusoidal (SW) illumination techniques.

	FF	CB	SW
RMSE Mean	0.0952	0.0153	0.0112
RMSE STD	0.58×10^{-3}	0.11×10^{-3}	0.16×10^{-3}

7.3.2 Albedo Dependence

Previous research has noted an albedo dependence on distance (Lichti *et al.*, 2014), which is highly noticeable when imaging a flat checkerboard. A rectangular checkerboard with 90% contrast between the dark and light regions of the checkerboard is imaged to test direct and global separation to resolve this error. To show the direct distance is statistically significantly different from the full field distance a paired t-test is performed between each distance image. The null hypothesis is the means are equal, and the alternative hypothesis are the means are different.

Figure 7.4 plots the results of the direct and global separation when imaging a black and white checkerboard on a flat sheet. The sinusoidal pattern and checkerboard pattern correct the distance error. The results in figure 7.4 demonstrate the distance error with amplitude in ToF range cameras is due to multi-path interference, and can be corrected with direct and global separation. It has been proposed the albedo dependence is created by lens flare by Mure-Dubois and Hügli (2007), as illustrated in figure 7.3 with inter-reflections in the lens cavity. The reflection scatters inside the lens cavity and mixes with surrounding pixels which becomes noticeable when imaging a flat sheet with varying albedo regions, as shown in figure 7.4c. The different albedo patches are visible in the distance image. Using direct and global separation the albedo is no longer visible in the distance images in figure 7.4e and 7.4h. As the lens scattering is spatially smooth the multi-path interference can be removed by direct and global separation. To get an indication of the improvement a 3D plane is fitted through the [X,Y,Z] points for all three measurements. The RMSE is taken of the difference between the fitted plane and measured points. The full field RMSE is 6.2 mm, the checkerboard direct is 5.3 mm and the sinusoidal pattern is 4.8 mm. The improvement in RMSE is limited by the increase in noise as discussed in section 7.2.2. The null hypothesis is rejected at the 1% confidence level, (P=0), for both the checkerboard and sinusoidal illumination patterns, indicating the direct and global separation significantly changes the image and the differences in the images plotted in figure 7.4 are not due to noise.

Other techniques have been developed to remove the albedo dependent error. Lindner *et al.* (2010) demonstrated a distance error correlated with amplitude and attempted a correction with calibration. Lichti *et al.* (2014) also used scattering models to perform depth calibration based on amplitude. The advantage of using direct and global separation is no calibration or scene preparation are required.

7.3.3 Subsurface Scattering

A scene with two candles and five peppers is imaged to test the performance of direct and global separation to resolve multi-path interference caused by sub-surface scattering. There has been limited work into measuring sub-surface scattering and resolving the resulting multi-path interference in ToF range cameras, while measuring sub-surface scattering has been investigated with traditional direct and global separation. Subsurface scattering is present in

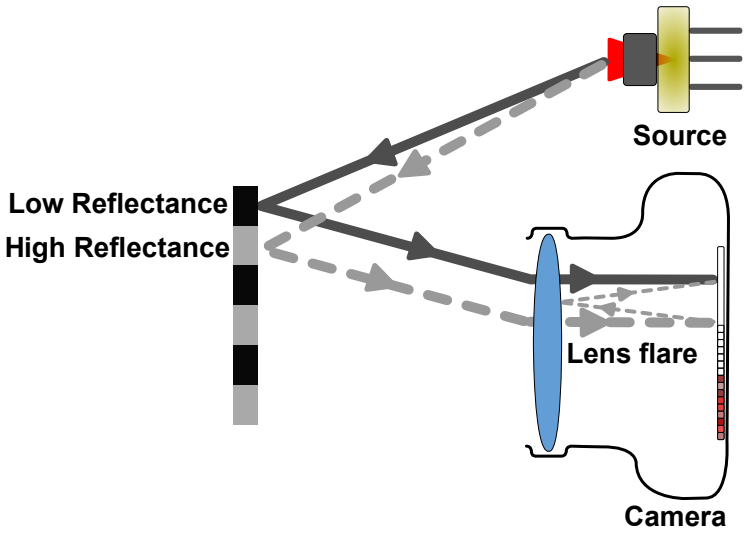


Figure 7.3: Proposed cause of the albedo dependent error. Due to scattering inside the optical cavity of the camera objects at the same distance but different albedo appear at different distances.

measuring meat and fruit products, resolving subsurface scattering allows for application of ToF range imaging in more industrial applications.

Subsurface scattering is a type of multi-path interference caused by light penetrating the object and scattering within the object and exiting at a different location. Materials such as wax, milk, tissue (human and animal), and fruit all exhibit sub-surface scattering (Jensen *et al.*, 2001). Figure 7.5 shows the results of applying direct and global separation on a scene containing objects that exhibit subsurface scattering. From the global amplitude images multi-path interference caused by subsurface scattering is present in the wax and peppers. Subsurface scattering is wavelength dependent, and with the 660 nm laser diode used, the red peppers exhibit more subsurface scattering than the other coloured peppers. Figure 7.6 clearly demonstrates the difference in distance and amplitude measurements between full field illumination and the direct component. In the direct image the candles are now cylindrical, and the peppers distances are not amplitude dependent. The application of ToF range camera measurements of subsurface scattering measurements is an area of further research. The amplitude and depth of scattering is encoded in the global return, and this information may enable new future applications in fruit grading or medical imaging devices.

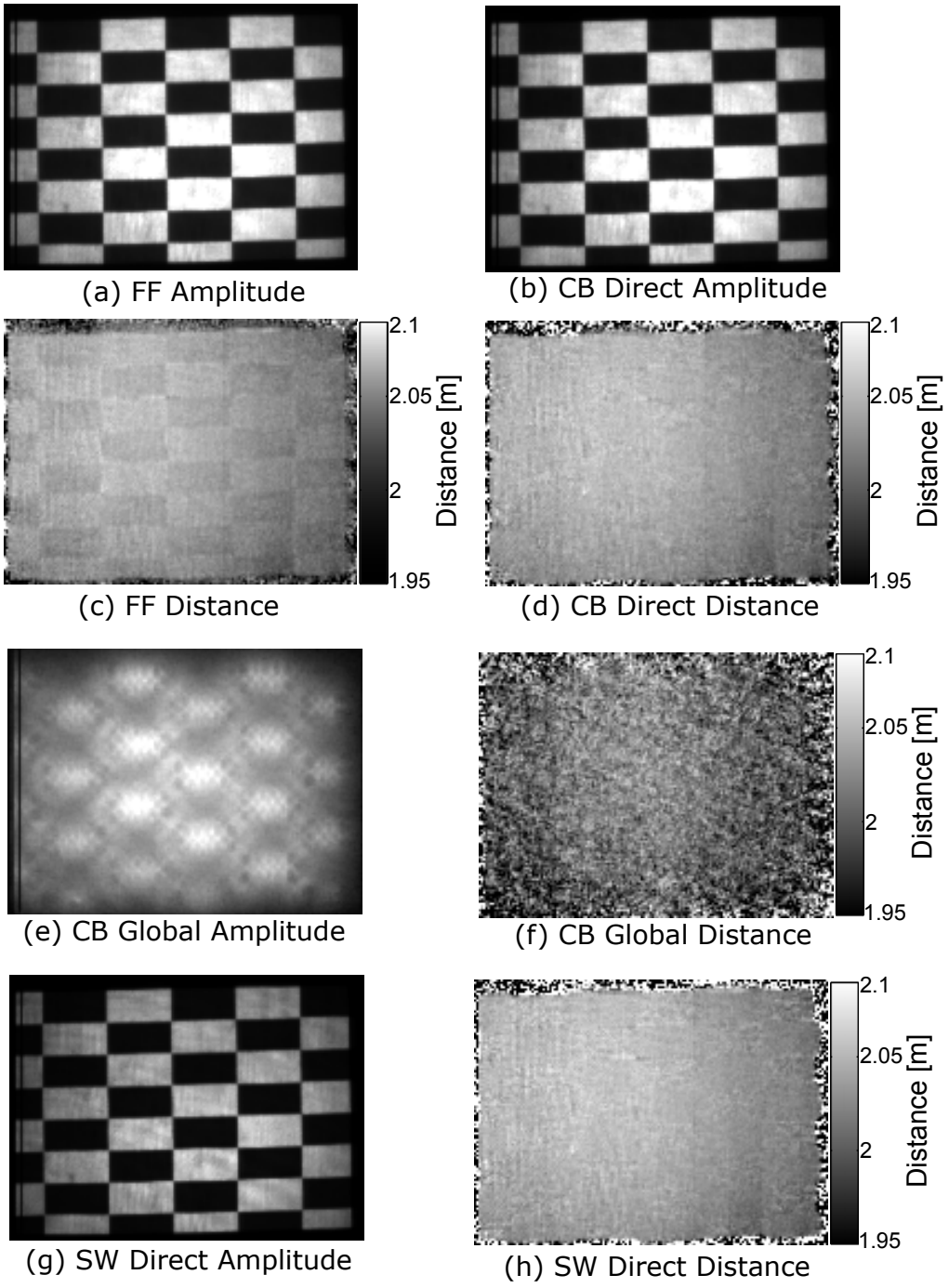


Figure 7.4: The direct and global separation methods remove the distance dependence on albedo. The measured full field (FF) amplitude and distance is in (a) and (c) respectively, the checkerboard (CB) direct amplitude and distance in (b) and (d). The checkerboard global amplitude and distance is in (e) and (f), with the sinusoidal (SW) direct amplitude and distance in (g) and (h).

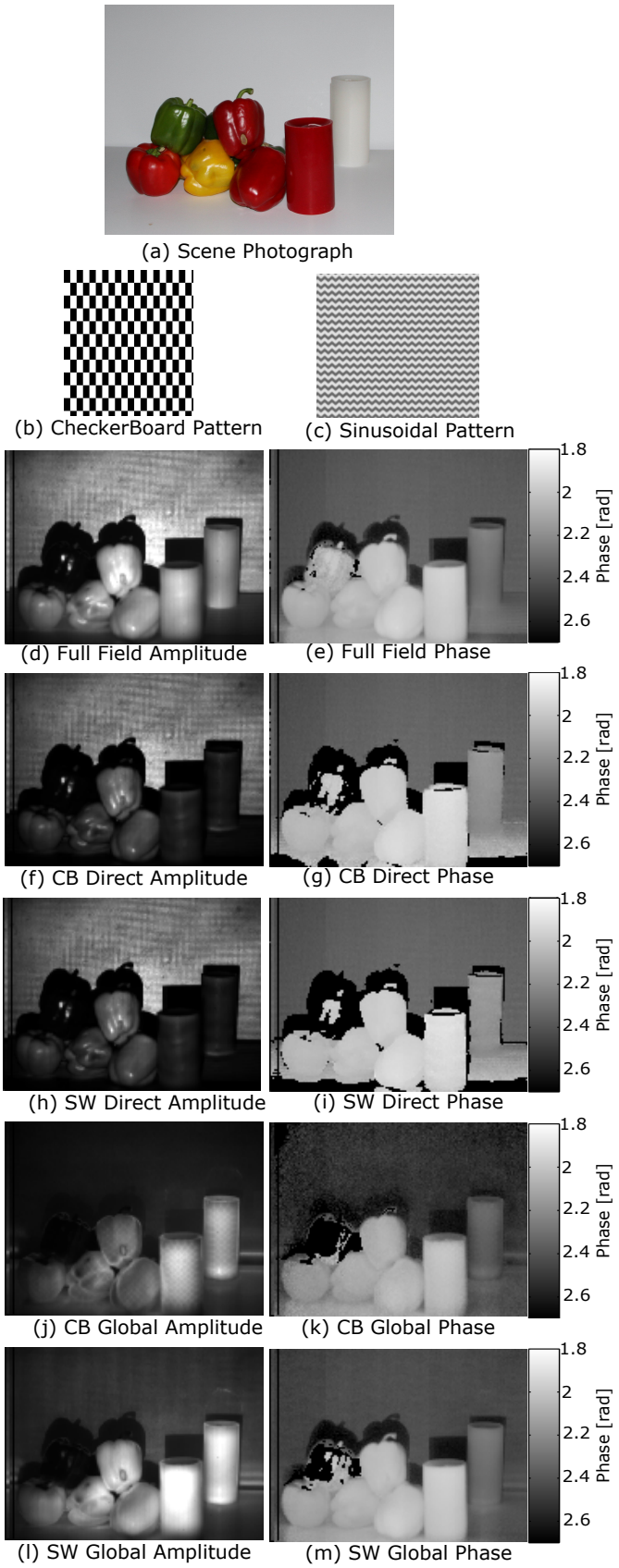


Figure 7.5: Direct and global measurements for checkerboard and sinusoidal pattern with a scene that suffers from subsurface scattering. A photograph of the scene is in (a), with the two projected patterns in (b) and (c). The full field (FF) results are in (d) and (e), the checkerboard (CB) results in (f) to (g) and the sinusoidal (SW) results in (h) to (m).

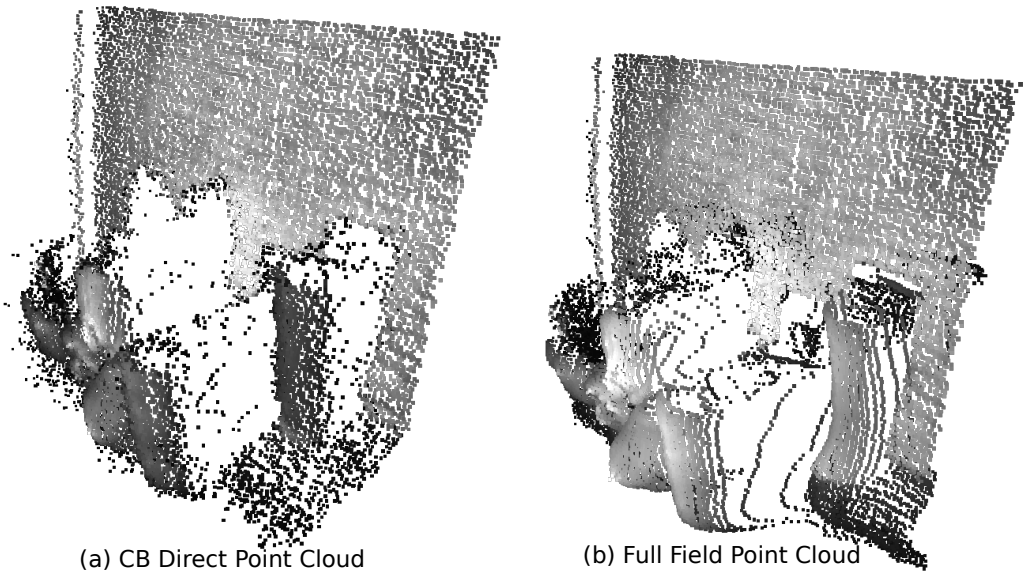


Figure 7.6: Point cloud comparing the measured checkerboard (CB) direct distance and full field (FF) distance. Each point encodes the measured amplitude. The distance to the peppers in the full field point cloud is dependent on the measured amplitude while this is not the case for the checkerboard point cloud. The candles are correctly measured in (a) while highly distorted in (b).

7.3.4 Pattern Size

To test the effect of pattern size on the direct and global separation a mannequin and spray paint can are imaged. The mannequin has varying size locations for inter-reflections to occur, and is a real world demonstration of inter-reflections of every day objects. The scene is imaged with three different sized checkerboard patterns, with each pattern phase shifted 25 times. The different sized patterns are displayed in figure 7.7.

The pattern size projected has an effect on the reconstructed results. Direct and global separation works on the constraint that the global component has to have a lower spatial frequency than the projected pattern. A demonstration of the effect of the checkerboard pattern size are shown in figure 7.7 for the direct and global amplitude. The global amplitude indicates where multi-path interference is present, and as the checkerboard grid size increase the global amplitude decreases in some regions. The inter-reflections on the mannequin's eye sockets and nose are not detected at the larger pattern sizes. The projected pattern size is limited by the camera's sensor's spatial resolution, the theoretical maximum is a one to one correspondence between the checkerboard pattern and each pixel of the imaging sensor. The smaller the projected pattern the higher the spatial frequency the global return image can contain so

the better the multi-path interference removal. However, in practice there is mis-alignment between the projector and camera and edge of the checkerboard pattern is spread across multiple pixels. With a smaller pattern there are more edges, which are partially illuminated, therefore more phase shifted patterns are required. When 25 patterns are used the number of edges is not an issue, but it becomes more of an issue when fewer patterns are used. For practical implementation in 30 fps range cameras the least number of patterns to image is required.

7.3.5 Translucent Sheet

Translucent objects introduce multi-path interference, as demonstrated in chapters 4 and 5. The results of direct and global separation, using a scene with a translucent sheet, are shown in figure 7.8. The direct and global separation fails to resolve the multi-path interference caused by the translucent sheet. The forward model is violated as the amplitude of the front and back returns are of the same spatial frequency as the projected pattern and not spatially smooth as required, therefore the direct and global separation fails. Another issue is the light source and camera are not collocated therefore the illuminated pixel on the translucent sheet does not illuminate the background.

7.4 Comparison

The presented method of direct and global separation fails on specular reflections and mixed pixels because both forms of multi-path interference violate the global spatial frequency constraint placed on the global component. Compared to the use of binary sequences with ToF range cameras, known as coded ToF, presented in chapter 4 where specular reflections and mixed pixels can be resolved. The multiple frequency methods presented in chapter 5 also can resolve specular reflections and mixed pixels. Direct and global separation can resolve multi-path interference caused by inter-reflection in a corner, in contrast to other methods can not.

The sinusoidal direct and global separation uses nine raw frames for a closed form solution to resolve multi-path interference. The previous closed form solution by Godbaz *et al.* (2012) required phase and amplitude at four different frequencies, the minimum number of raw frames would be twelve. Other non-closed form solutions that measure multiple modulation frequencies include Freedman *et al.* (2014) who uses three modulation frequencies and a

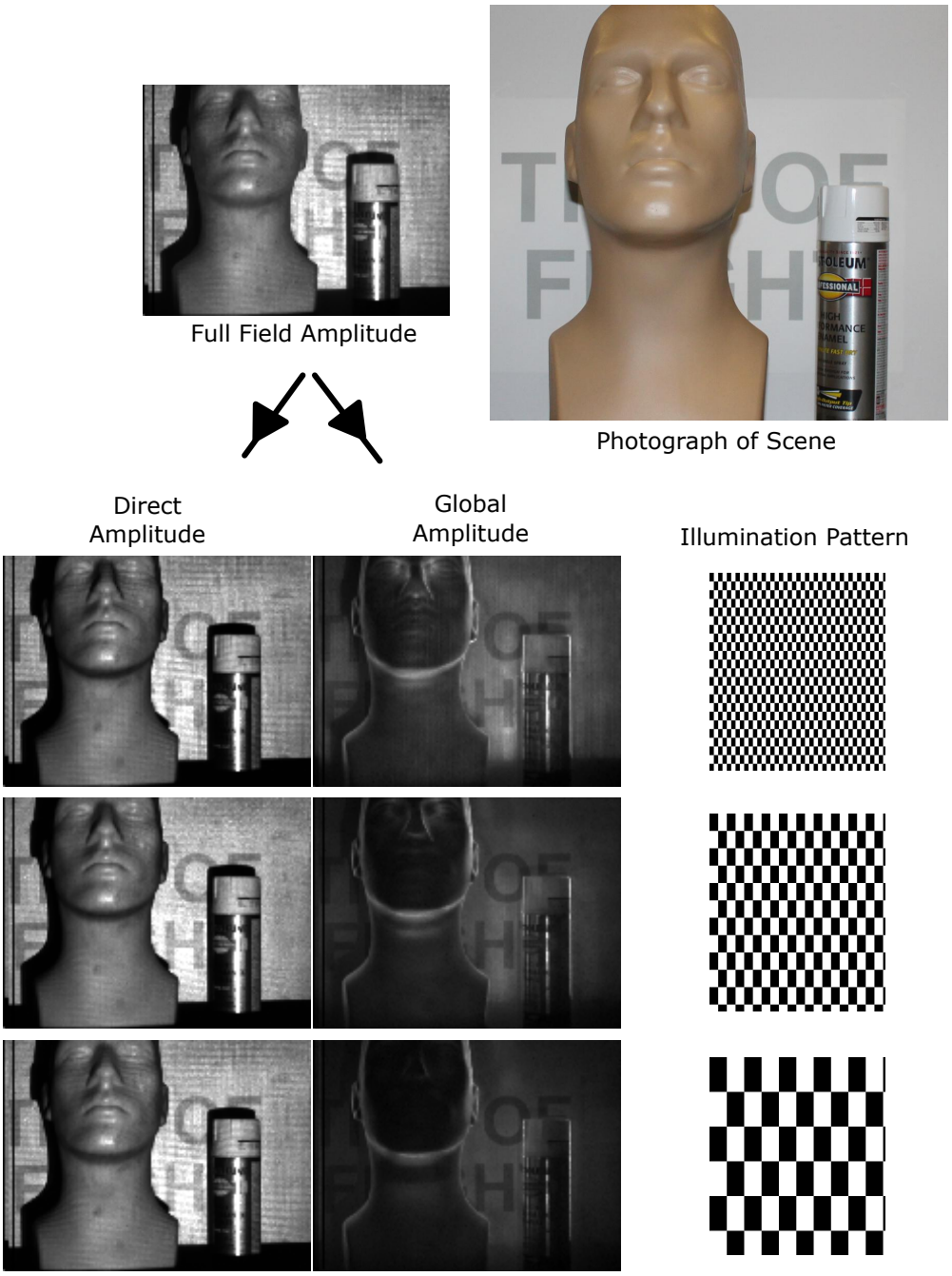


Figure 7.7: Effect of grid size on the direct and global amplitude. With larger patterns the inter-reflections on the mannequin's nose and eye sockets are not detected.

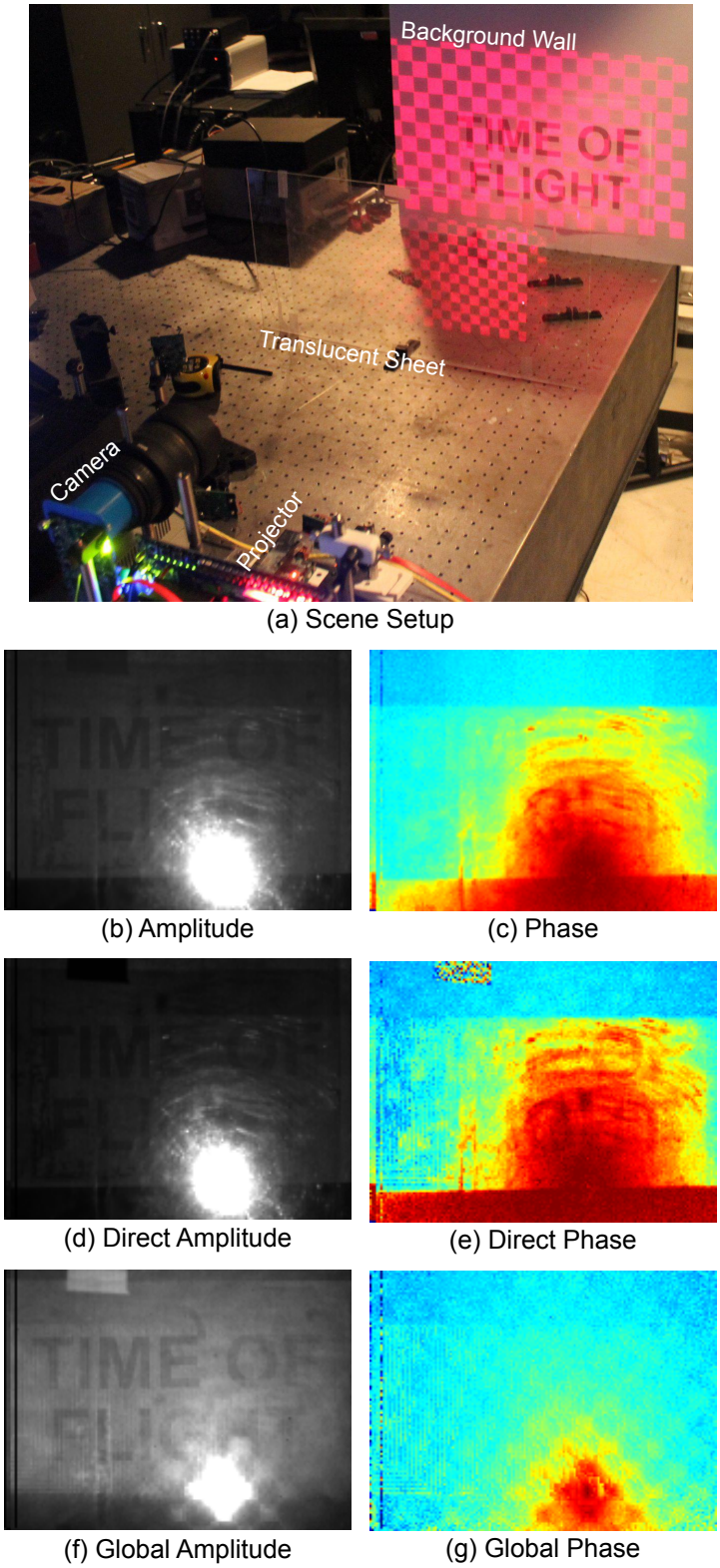


Figure 7.8: Results of direct and global separation on a translucent sheet. Photo of the measured scene in (a), with the full field amplitude and phase in (b) and (c). The resolved direct and global components using a checkerboard illumination pattern is shown from (d) to (g). The direct and global separation is unable to separate the return from the translucent sheet and the return from the background.

compressed lookup table and Dorrington *et al.* who used two modulation frequencies for a minimum of six raw frames. Both these methods have been implemented on 30 fps range cameras. The work presented in chapter 4 used approximately 300 raw frames for accurate depth measurements, making it unsuitable for real-time implementation. Other multi-frequency methods such as Bhandari *et al.* (2014) and Kirmani *et al.* (2013) both require a minimum of five frequencies making them difficult for real time implementation.

Previous methods for resolving measurement errors due to lens flare rely on calibration and estimating the complex point spread function of the mixing in the lens. With the presented method no calibration is required therefore no assumptions are made about the scene.

7.5 Conclusion

Time-of-flight (ToF) range cameras suffer from measurement errors caused by multiple propagation paths. In this chapter a novel method was presented where multi-path interference in ToF range cameras has a closed form solution in nine raw frames. The method combined the phase measurement of ToF range cameras with projected sinusoidal patterns to perform direct and global seperation. The constraint is that the global return is spatially smooth, therefore specular reflections and multi-path interference caused by translucent sheets cannot be resolved. Using the direct and global separation the multi-path interference caused by inter-reflection in a corner, sub-surface scattering in fruit, and lens flare are all correctly removed.

Chapter 8

Combination

The traditional method of ToF range imaging as reviewed in chapter 1 is homodyne AMCW, where the phase is sampled at a single frequency to measure depth. AMCW suffers from errors caused by multi-path interference, and in previous chapters various techniques have been researched to resolve this measurement error. Replacing the square wave with a binary sequence and applying sparse deconvolution to the measured cross-correlation was explored in chapter 4. Sampling over multiple modulation frequencies was presented in chapter 5. In chapter 6 the duality between using binary sequences and sampling multiple modulation frequencies is proven. Using spatially modulated patterns is demonstrated in chapter 7. The possible sampling space is visualized in figure 8.1, with the current demonstrated methods marked. In this chapter combining sampling over phase, frequency and pattern is demonstrated for robust ToF range measurements while trying to minimise the number of raw frames required.

8.1 Theory

The previous image formation models for multi-path interference have been based on assumptions about the type of multi-path interference present. The assumption of sparsity, that a discrete number of propagation paths are present, was enforced in coded ToF, and in sampling multiple modulation frequencies. This assumption works best when the propagation paths are separated by half an ambiguity distance of the highest modulation frequency, in direct and global separation its assumed that the global component is spatially smooth, so no specular inter-reflections are present. These assumptions are only true for certain types of multi-path interference.

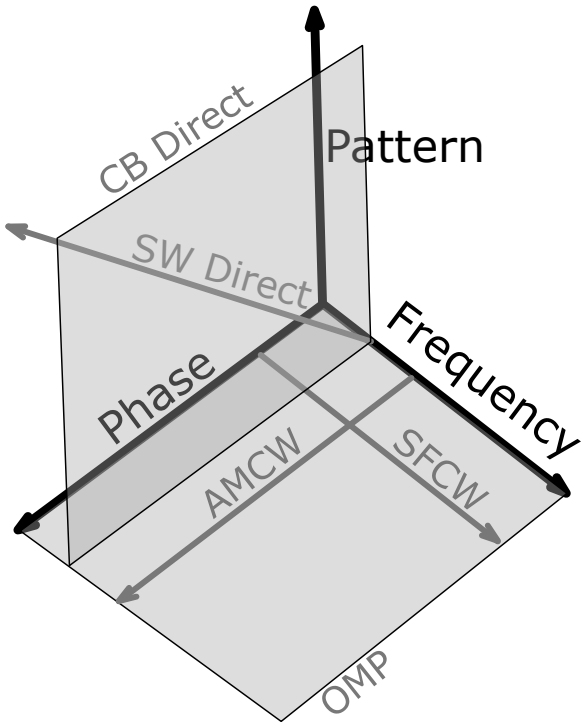


Figure 8.1: Sampling space of a time-of-flight range camera, with various sampling techniques to resolve multi-path interference are marked. AMCW is the traditional range measurement technique reviewed in chapter 1, SFCW and OMP are presented in chapter 5. Checkerboard (CB) and sinusoidal (SW) direct and global separation techniques are presented in chapter 7.

The sample space that is visualised in figure 8.1 is described by equation 7.25. We propose a new image formation model for multi-path interference that is solvable by a combination of coded ToF, presented in chapter 4, and direct and global separation presented in chapter 7. The measured value at each pixel is composed of a direct return, a specular return, and a global return such that

$$L = L_d + L_s + L_g. \tag{8.1}$$

The direct return, L_d is the shortest path length from the light source to pixel. The specular component, L_s, is a discrete propagation path that is longer than the direct path, and the global component encompasses everything else. In the

complex domain equation 8.1 becomes

$$L = a_d e^{j\phi_d} + a_s e^{j\phi_s} + \int_{d_1}^{\infty} a_k e^{j\phi_d} dk. \tag{8.2}$$

When a checkerboard illumination pattern is used then the measured value in the non-illuminated (dark) and illuminated regions are

$$L_{dark} = \alpha L_s + L_g/2, \tag{8.3}$$

and,

$$L_{ill} = L_d + \beta L_s + L_g/2, \tag{8.4}$$

where α and β describe the coupling between the specular component and the surrounding pixels. These variables are highly scene dependent. It would appear that there are more unknowns than equations, however the additional constraint that the L_d and L_s components are sparse, and further that L_s is only one path length then α or β must be 0 or 1.

Coded ToF described in chapter 4 is combined with direct and global separation using checkerboard illumination patterns. When using a checkerboard illumination pattern a function is required for the maximum and minimum of the reflected signal, with the choice of the binary signal determining the measured cross-correlation. The minimum and the maximum over all the images, I, can be defined as the numerical integral of the absolute measured value

$$L_{max} = \max_I \sum_{n=0}^{N} |h_I(\tau_n)|, \tag{8.5}$$

$$L_{min} = \min_I \sum_{n=0}^{N} |h_I(\tau_n)|. \tag{8.6}$$

The equations 8.5 and 8.6 are analogous to equations 7.2 and 7.3. Additional constraints placed on equations 7.2 and 7.3 in chapter 7 as the complex measurements can add together destructively. This additional constraint can be avoided by placing a constraint on the binary sequence so that the cross-correlation adds together constructively and never destructively, written mathematically as

$$|h(\tau + t_d) + h(\tau + t_g)| \geq |h(\tau + t_g)|, \tag{8.7}$$

where t_d and t_g are the phase offset due to the direct and global components. Referring back to figure 4.2 both the maximum length sequence and barker sequences met the constraint in equation 8.7 provided the background signal b is greater than the negative offset of the correlation, which is the case for the PMD19k3 readout process is true.

To combine the image formation presented in equations 8.3 and 8.4 with the coded direct and global separation a further constraint is required on α and β. In the case where $\alpha = 1$ and $\beta = 0$ then the minimum and maximum computation in equations 8.5 and 8.6 can be invalid. The minimum and maximum are calculated correctly, but when the subtraction in equation 8.10 occurs a negative version of the specular components cross-correlation instead of positive version is present in the signal. Therefore for the combined method the constraint that $\alpha = 0$ and $\beta = 1$ is required. With these constraints then

$$L_{max} = L_d + L_s + L_g/2, \tag{8.8}$$

$$L_{min} = L_g/2, \tag{8.9}$$

$$L_{max} - L_{min} = L_d + L_s. \tag{8.10}$$

After the subtraction of the global component sparse deconvolution can be used on the measured cross-correlation to resolve the direct and specular components.

8.1.1 Experimental Setup

Two scenes are imaged to test the combination of coded ToF range imaging with direct and global separation methods. The first scene is a corner constructed out of diffuse white foam board and imaged using an 11 bit maximum length sequence, generated at 100 MHz, and 150 phase steps sampled. Two patterns are used: a checkerboard and its inverse. After the direct and global separation is performed on the cross-correlation the direct component is transformed to a frequency sweep, see chapter 6, with the phase then converted to radial distance.

The second scene is designed to test the combination of multi-path interference removal techniques when both global and specular multi-path interference is present. The additional constraints can limit the scenes that can be resolved using the combination of techniques. In chapters 4 and 5 imaging through translucent sheets was used to demonstrate resolving sparse multi-path interference. However the light source and imaging sensor are not collocated,

therefore illuminated points on the translucent sheet do not correspond to illuminated points behind the translucent sheet on the Z axis, which violates the constraint in equation 8.8. Specular inter-reflections, such as a disco ball inside a corner which is a failure case presented by Nayar *et al.* (2006) would be an ideal test scene. However due to the limited maximum operating frequency of 50 MHz and the constraint that the path lengths should be separated by half an ambiguity distance, which corresponds to 1.5 m, makes it a difficult to create a useful scene. To demonstrate the combination of techniques a corner is constructed out of white foam board and placed 3.1 m in-front of a flat foam board sheet with slits cut out. With the large aperture lens used, the foreground object is out of focus, resulting in mixed pixels between the foreground and background. The scene is imaged with an 11 bit m-sequence generated at 100 MHz, and two spatial patterns are used: a checkerboard and its inverse. The cross-correlation is sampled 88 times, and for the sparse deconvolution this is interpolated up to 3000 so the distance is not quantized by the sparse deconvolution process. The full field and direct components are transformed into a frequency sweep to measure the phase to compute the distance. Four measurements are tested, full field illumination, direct and global separation, spares deconvolution, and the combination.

8.1.2 Results and Discussion

The measured cross-correlation of a single pixel, located near the centre of the corner, and its separation into direct and global components is plotted in figure 8.2. Due to the location of the pixel in the corner the global component has a smaller magnitude and is slightly phase delayed compared to the direct component. Compare this to figure 4.5 where the two returns are separated by a large enough distance that the individual returns are visible. A comparison of the measured distance along a slice is plotted in figure 8.3. The measured direct distance in figure 8.3 is similar to figure 7.2 indicating that the multi-path interference has been removed in the corner.

The next experiment combines the direct and global separation with the sparse deconvolution on the cross-correlation signal to resolve both inter reflections in a corner and mixed pixels. The measured natural logarithm of the amplitude and distance of three of the cases is plotted in figure 8.4. The presented results are the natural logarithm of the amplitude so the entire scene is visible; without this the dynamic range of the amplitude is difficult to present. In the full field amplitude image shown in figure 8.4a reflections of

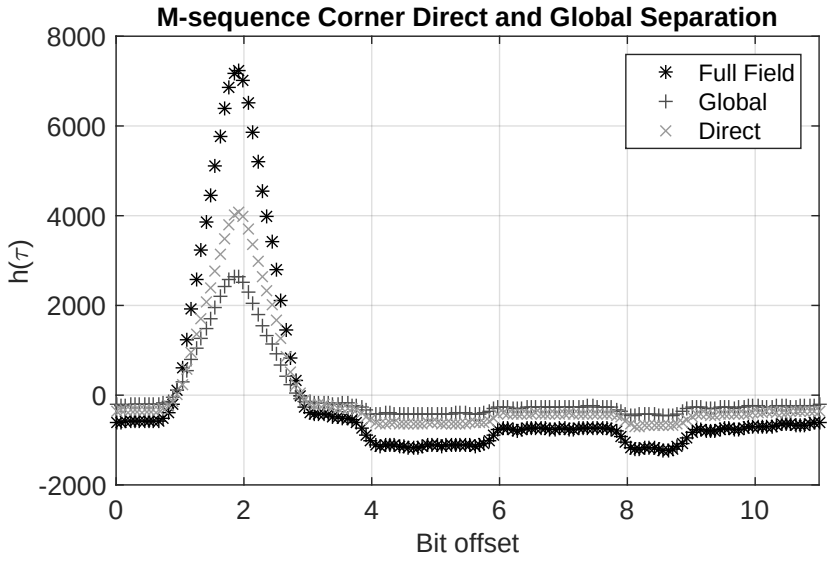


Figure 8.2: Measured cross-correlation of an maximum length sequence in a corner, and its separation into it's direct and global components.

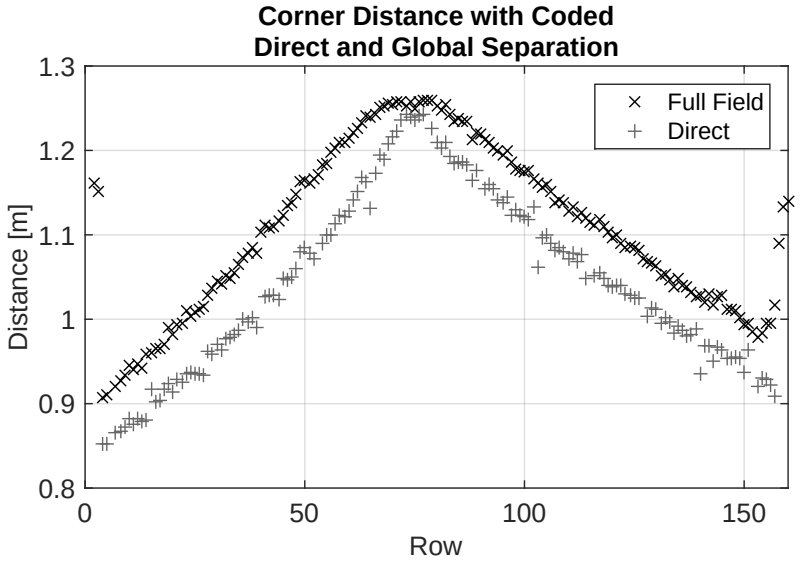


Figure 8.3: Measured distance in a corner using coded ToF techniques with full field illumination and direct and global separation. The direct component is able to correctly measure the distance.

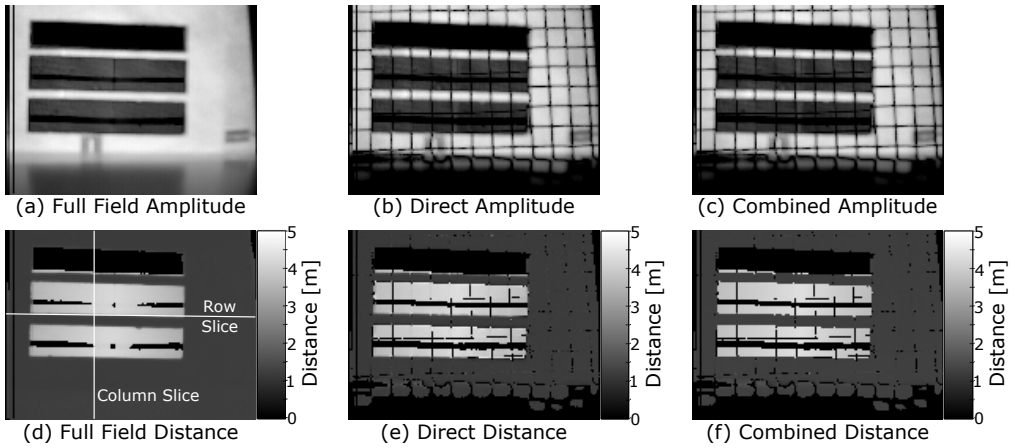


Figure 8.4: Measured amplitude and distance of a corner using full field illumination, direct and global separation, and the combination of techniques. The marked row and column in (d) for the three techniques measurement techniques is plotted in figures 8.6 and 8.7.

the light between the slits and the table it is positioned on is visible. In the other presented methods these pixels have been removed by thresholding the amplitude to remove invalid pixels. In all the images in figure 8.4 there are filtered bands on the corner in the background. These are due to the shadow cast by foreground slits caused by the light source not being collocated with the imaging sensor. The checkerboard pattern is clearly visible in the direct and global separation because two patterns are used to minimize the number of raw frames.

Point clouds of the depth images plotted in figure 8.4d, e and f are plotted in figure 8.5a, b and d. In addition the result of using sparse deconvolution on the measured cross-correlation as detailed in chapter 4 is plotted in figure 8.5c. In figure 8.4a the corner is rounded and there are many mixed pixels between the slits and corner. In the direct image in figure 8.5b the measured distance in the corner is corrected, but there are still a number of mixed pixels between the foreground object and the background corner. The mixed pixels are corrected by sparse deconvolution as shown in figure 8.5c, but the corner is still rounded like in figure 8.5a. The combination of direct and global separation with sparse deconvolution is plotted in figure 8.5d. Both the corner is measured correctly and the mixed pixels are resolved.

Without the ground truth for the corner distance it is difficult to perform quantitative analysis. But it is notable that the corner has gone from rounded to straight, which looks similar as in figure 7.2 where the ground truth has

been measured. The mixed pixels are easily identified and in a perfect imaging system there should be no mixed pixels present. In the full field image there are 699 mixed pixels, in the direct image there are 417 mixed pixels, in the sparse deconvolution no mixed pixels. Finally in the combined approach there are 41 mixed pixels. The combined approach reduces the number of mixed pixels by 94.1% over the full field, and over 90% compared to the direct and global separation, but is not quite as perfect as the sparse deconvolution alone.

To demonstrate further the results of the combination a slice along a single row and column of the imaging sensor as marked in figure 8.4d is plotted in figures 8.6 and 8.7 respectively. The distance along a row of the sensor plotted in figure 8.6 is on a set of pixels on the edge of one of the slits, so a majority of the pixels are mixed between the foreground and background. The full field measurements are corrupted, and no useful depth information is measured. With the direct measurement while it appears to have limited difference from the full field the inter-reflections in the corner have been removed, but the dominant multi-path interference is the mixing between the foreground and background objects. With the combination of direct and global separation and sparse deconvolution the depth of the foreground and the corner are recovered.

8.2 Summary

In this chapter the combination of techniques has been shown to resolve mixed pixels and diffuse inter-reflections. Currently the types of multi-path interference that can be resolved is limited. Ideally the combination would allow for video frame rate range imaging. We have demonstrated the technique using two patterns for direct and global separation. Previous work by Godbaz (2012) used two frequencies to resolve sparse multi-path interference, of which some of the work in this thesis is compared to in figure 5.5. With two frequencies and two patterns and three raw frames per frequency or pattern results in a total of 12 raw frames. Current state of the art cameras use as many as 9 raw frames per depth frame (Bamji *et al.*, 2015), therefore 12 raw frames is in the realm of achievable at 30 depth frames per second. This result is a step towards video frame rate depth measurement that is robust to multi-path interference, resulting in accurate measurements.

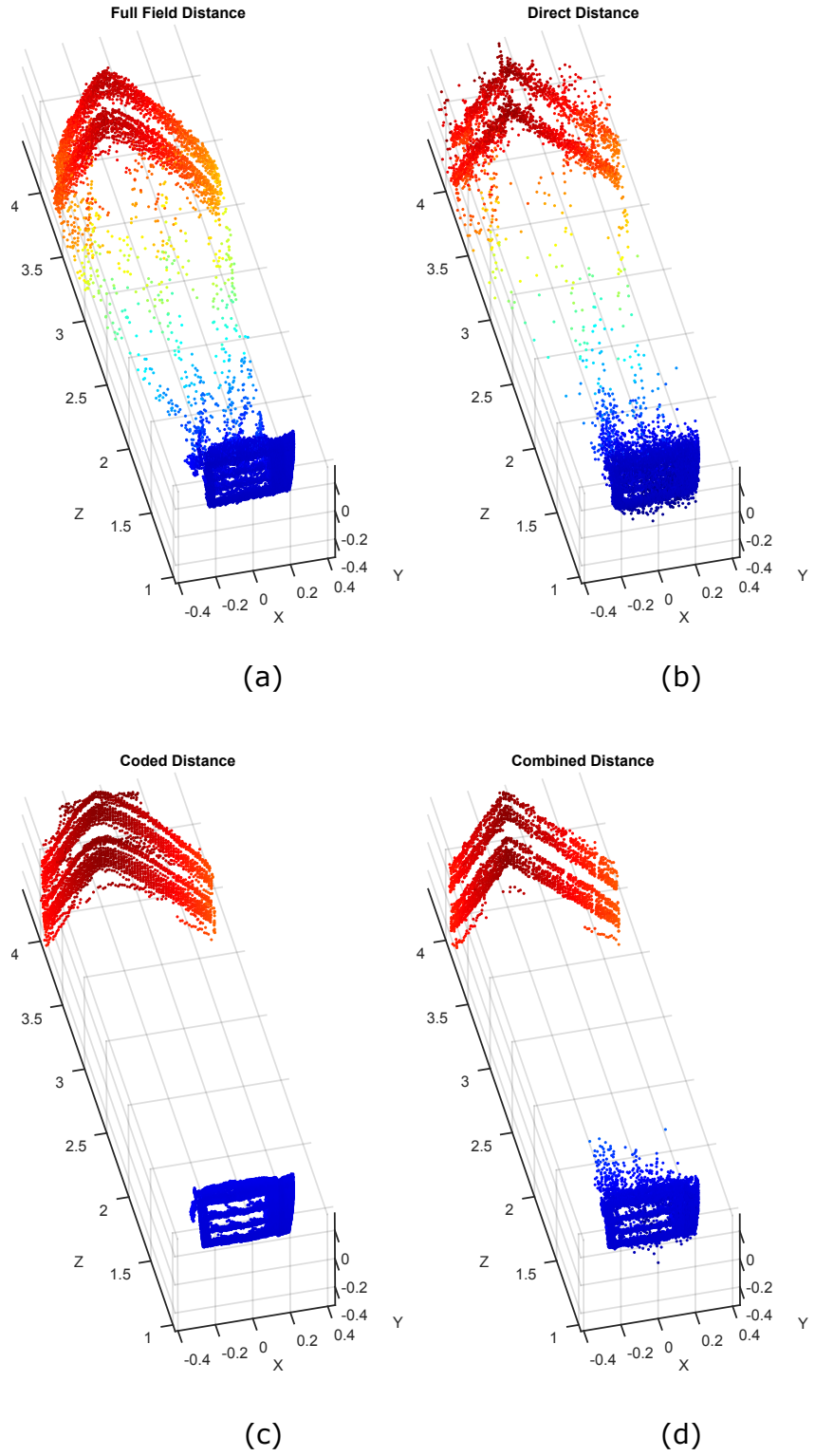


Figure 8.5: Point clouds of the measured distance in figure 8.4. The full field distance in (a), the direct distance in (b), the sparse deconvolution in (c), and the combined in (d). In (a) the corner is measured incorrectly and many mixed pixels between the foreground slits and corner are present. In (b) the corner is correctly measured, but the mixed pixels are still present, in (c) the mixed pixels are resolved but the corner is rounded. While in (d) the mixed pixels are corrected and the corner is correct.

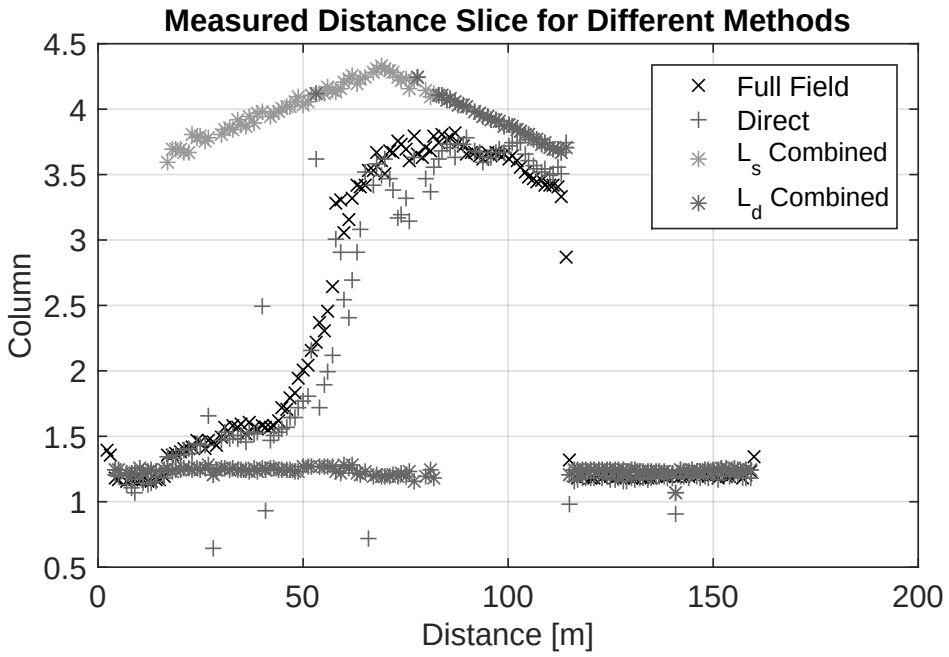


Figure 8.6: Measured distance of a row of pixels in figure 8.4d, for three of the presented measurement techniques. The full field and direct measurements do not resolve the background and foreground resulting in corrupted range information. In contrast, the combined technique can resolve both path lengths of the foreground and the background corner.

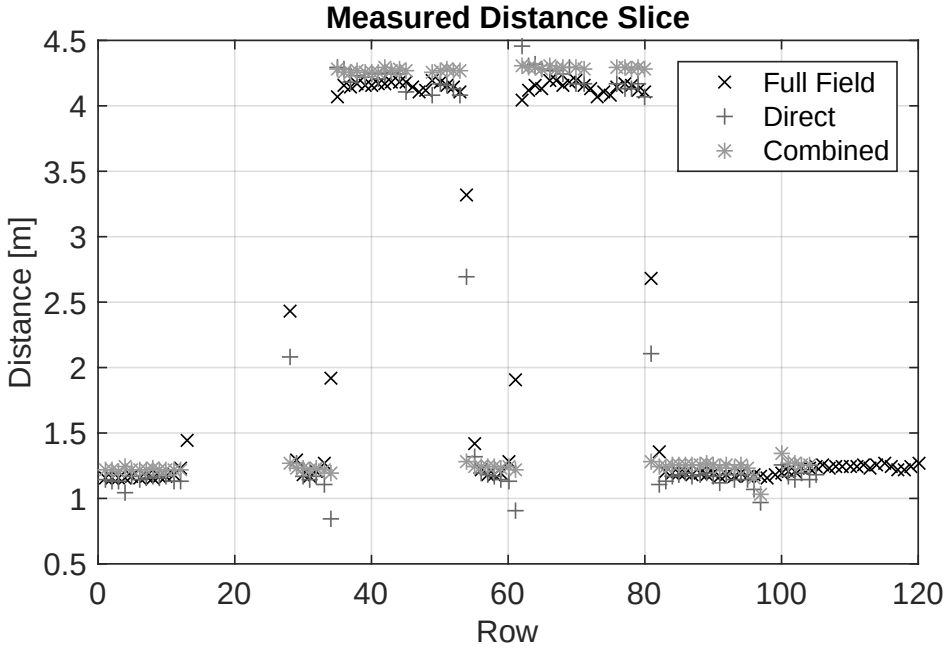


Figure 8.7: Measured column from figure 8.4d for three techniques. The full field and direct distance measurements suffer from mixed pixels between the foreground and background. While the combination of direct with coded ToF resolves the mixed pixels.

Chapter 9

Conclusion and Outlook

Time-of-flight (ToF) range cameras have progressed remarkably over the last ten years. Pixel array sizes have increased from 64 by 64 to 512 by 424. In the same time the modulation frequency range has increased by an order of magnitude from around 20 MHz to over 200 MHz. There has been much activity in the commercial and research space developing range imaging technology and finding novel applications. Microsoft® has acquired two startup companies and sold over 10 million Kinect® ToF range cameras. It is expected that ToF range cameras will continue to appear in a number of consumer applications. A "Killer App" is yet to emerge for range imaging in the consumer market.

One of the limiting factors of ToF range cameras has been error caused by multiple propagation paths from the light source to the sensor an error known as multi-path interference. Multi-path interference can be caused by inter-reflections, subsurface scattering, and bulk diffuse scattering mediums. As multi-path interference is highly scene dependent the error can not be resolved by calibration. This thesis has made significant contributions for realizable methods for resolving multi-path interference in ToF range cameras. Three methods have been developed for resolving multi-path interference.

Ideas from spread spectrum communications has been applied to ToF range cameras to resolve multi-path interference. The cross-correlation function is dependent on the illumination and reference signals, by replacing the current square wave with a binary sequence additional information is embedded in the cross-correlation signal. By applying sparse deconvolution to the measured cross-correlation the amplitude and distance of discrete propagation paths can be resolved. However, when the propagation paths are separated by less than half an ambiguity distance of the highest frequency in the cross-correlation, sparse deconvolution fails to resolve the multi-path interference. The depth is

quantized to the number of samples, therefore a large number of samples (raw frames) are required: in the experimental results over 3000 raw frames were collected. A number of successful examples of resolving multi-path interference by imaging through translucent objects have been presented. The RMSE of imaging a garden gnome through a translucent sheet decreased from 0.954 m using AMCW to 0.145 m using coded ToF.

ToF range cameras are a multi-shot measurement device. Traditionally ToF range cameras perform multiple measurements at different phase offsets of the reference signal to estimate the phase and amplitude of the reflected light. Sampling over multiple modulation frequencies enables resolution of multi-path interference. Three sampling methods have been presented; sampling over phase and frequency, just frequency, and sampling a chirp. Measurements across frequency and phase produce complex data, while measurements across just frequency produce real data. The depth is encoded by the spectrum of the measured signal: longer distances produce a higher frequency. Different propagation paths are encoded at different frequencies, and using frequency estimation techniques the different path lengths can be separated. Harmonic aliasing is also encoded as a higher frequency. The accuracy of the depth measurement when sampling over just frequency is limited by the bandwidth of frequencies sampled over. The resolution of propagation paths is robust over half an ambiguity distance of the highest frequency measured. Sampling over frequency and phase produces complex data, and with this data at least five frequencies are required, with three phase steps per frequency a minimum of 15 raw frames are required. Experimental results show a minimum of six frequency steps are required for sampling over frequency alone. When imaging a scene with a translucent sheet present the SFCW measurements have much lower median error than the AMCW, and the spread in error as measured by the interquartile range is reduced by a factor of 22.

A transform between a binary sequence and frequency sweep has been shown. Measurements taken with one domain can be transformed to the alternative domain. The multi-path interference restoration methods developed for coded ToF and frequency sweeps have been applied to the same data and yield the same results as demonstrated by a paired t-test that can not reject the null hypothesis that the means are equal. The advantages of temporal coding are reduced jitter, temperature stability and an extended linear operating region of the camera. The advantages of frequency coding are control of

frequency selection and no unused spectrum. Temporal coding measurements are more stable, while processing frequency coding requires fewer samples.

Direct and global separation has been applied to ToF range cameras to resolve multi-path interference. Multi-path interference is described as the sum of a direct return, which is the shortest path length, and a global return which encompasses everything else. The direct and global separation techniques were adapted to work on the complex data captured with ToF range cameras and expanded to enhance the modulation and demodulation process of ToF range imaging. In particular the projection of sinusoidal patterns mix with the AMCW illumination signal, a sinusoidal, for a closed form solution to multi-path interference in nine raw frames. The direct and global separation assumes the global return is spatially smooth, therefore multi-path interference caused by specular reflections and translucent sheets cannot currently be resolved using this method. Using the direct and global separation the multi-path interference caused by inter-reflection in a corner, sub-surface scattering in fruit, and lens flare are all correctly recovered. The RMSE in the corner has reduced from 0.0952 m when using full field illumination to 0.0153 m when using checkerboard illumination patterns to 0.0112 m when using sinusoidal illumination patterns. The albedo distance dependence is removed by direct and global separation. The RMSE of imaging a checkerboard, that has a 90% contrast between the dark and light regions, decreases from 6.2 mm using full field illumination to 4.8 mm using sinusoidal illumination patterns. The direct is statistically significantly different from the full field as tested by a paired t-test between the two distance images.

Direct and global separation has been combined with coded ToF, and demonstrated on a scene combining mixed pixels with a corner. The corner is visually corrected and the number of mixed pixels is reduced by 90%.

9.1 Future Work

ToF range cameras are designed to operate at 30 fps of depth information, minimising the number of samples is important to maintain these frame rates. There is further research into minimising the number of raw frames required for the presented techniques.

Measuring light interaction with materials and light propagation in a scene is an area of research in computer graphics. Multi-path interference is a measurement of light interaction with materials, and scene relighting and different

light source wavelength measurements encode further information about light propagation. Techniques from light field photography can be applied to add angle information of the returning light. All this additional information can be used for better understanding of light transport in complicated environments, an active area of research in computer graphics.

Motion is a problem in ToF range cameras. The development of motion correction algorithms for both radial and tangential motion that take advantage of ToF cameras unique properties is an area for future research.

Noise is a problem in all imaging systems, one of the issues glossed over in the presented multi-path interference correction is the effect of noise. There is further work to explore the performance of these methods in the presence of noise. The SNR of ToF range cameras depends on the integration period, optical power emitted, modulation contrast, photon shot noise, readout noise, reset noise, and jitter. Optimizing variables for ideal noise performance is an area for further research.

9.2 Outlook

ToF range cameras have improved significantly in the previous years. As the bandwidth of ToF range cameras continue to improve then the presented methods for resolving multi-path interference will provide improved results. With the increase in bandwidth and availability of cameras novel applications of ToF technology will grow.

The major measurement errors in ToF range cameras, which are multi-path interference, phase wrapping, and harmonic aliasing, now have viable solutions. With the major sources of measurement errors resolved ToF range imaging technology is more appealing against competing range imaging technology. ToF range cameras can now be accurate, low cost, and miniature.

Appendices

Appendix A

Alternative Sampling

Time-of-flight (ToF) range camera technology has applications outside of range imaging. Three novel applications are demostrated in this chapter. This research is not core to resolving multi-path interference, but was undertaken during this research project therefore it is included as an appendix. A ToF range camera is modified to operate in a mode anaglous to continuous wave (CW) radar to measure velocity of objects by the doppler shift in the reflection. The use of multiple light sources for colour range measurements and dynamic relighting is explored. Finally ToF technology is used for fluorescence lifetime imaging (FLI), exploring a low cost alternative to current FLI methods.

A.1 Continuous Wave Modulation

Continuous wave (CW) radars measure the velocity of objects by emitting a continuous signal at a single frequency, normally in the MHz or GHz frequency range. By mixing the reflected signal with the emitted signal the velocity of target objects can be measured. CW radars have been employed since the Second World War, and there is vast literature on their design and applications (Charvat, 2014). With a coherent CW radar the phase of the reflected signal encodes the distance travelled by

$$\phi = \frac{2\pi f_o d}{c}, \tag{A.1}$$

where f_o is the frequency of the emitted signal. A moving object's radial distance (assuming no acceleration) with time is

$$d(t) = d_i + vt, \tag{A.2}$$

where d_i is the initial distance, and v is the velocity. The measured phase over time is therefore

$$\phi(t) = \frac{2\pi f_o(d_i + vt)}{c}. \tag{A.3}$$

The frequency shift in the returning signal encodes the velocity of the target by

$$f_d = \frac{2cv}{f_o}, \tag{A.4}$$

where f_d is the frequency of $\phi(t)$. CW Radars are easy to manufacture and accurately measure the velocity of objects and the CW operation results in good SNR.

A.1.1 Time-of-Flight Continuous Wave Operation

The concept of CW radar can be applied to ToF range cameras. CW radars emit RF spectrum and ToF cameras use visible spectrum. However the amplitude modulation of the ToF camera is in the RF frequency range, allowing radar techniques to be applied. CW radar operation is applied to ToF range cameras by keeping the phase offset, τ, constant. By substituting equation A.2 into equation 2.1 the reflected light is described by

$$s(t) = a(t) \sin\left(\omega t + \frac{2\omega(vt + d_i)}{c}\right). \tag{A.5}$$

The measured amplitude is a function of time because the amplitude depends on distance, which is changing. The amplitude depends on distance, reflectivity (Γ), and surface orientation is given by

$$a = \Gamma \frac{\mathbf{n} \cdot \mathbf{l}}{d^2}, \tag{A.6}$$

where $\mathbf{n}$ is the surface normal, and $\mathbf{l}$ is a unit vector, which together describe the surface orientation. Assuming the surface orientation stays constant during the object's motion, then the amplitude changes as

$$a(t) = \Gamma \frac{\mathbf{n} \cdot \mathbf{l}}{(d_i + vt)^2}. \tag{A.7}$$

The returning light is correlated with the reference signal $g(t)$.

$$g(t) = \sin(\omega t), \tag{A.8}$$

$$h = \int_0^{T_c} a(t)s(t)g(t)dt, \tag{A.9}$$

where T_c is the integration period of the camera. The time an object takes to travel one ambiguity interval (see equation 2.22) is T_v. In the case where $T_c << T_v$ then the correlation of $s(t)$ with $g(t)$ in equation A.9 does not distort the doppler shift in the reflection. The correlation measurement over time, with $\tau = 0$, for a ToF camera is therefore

$$h(t_q) = a(t)\cos\left(\frac{2\omega(vt_q + d_i)}{c}\right). \tag{A.10}$$

The time, t_q, in equation A.10 is quantized to T_c step sizes due to the integration period of the ToF range camera. The integration period of the camera is the main difference between CW radar and the presented method. The spectrum of $h(t_q)$ encodes the velocity of the object and its radiance. From equation A.10 the higher the modulation frequency the better the resolution of v due to more change in $h(t_q)$ for a smaller change in distance. Measuring at a constant phase offset is analogous to CW doppler radar with the frequency encoding the velocity. One difference is that with CW radar and CW ToF (as we propose) is that in CW radar the change in amplitude is not an issue, but it is significant in CW ToF.

A.1.2 Experimental Setup and Results

We capture data using a proprietary ToF range camera which has a maximum modulation frequency of 150 MHz. We test the CW doppler velocimetry by moving a flat white foam board under controlled conditions. The white foam board is placed on a 3.0 m translation stage (Macron Dynamics Inc, Croydon, PA, USA), and the translation stage is moved at varying velocities starting at 0.1 m/s going up to 1.1 m/s in increment of 0.1 m/s. The ToF camera is modulated at 150 MHz, with each raw frame having an integration period of 0.5 ms at a frame rate of 25 fps. The full field image results are calculated from using a zero crossing detection algorithm.

The measurement of one pixel over time for different velocities of the translation stage is plotted in figure A.1. The ambiguity distance with the

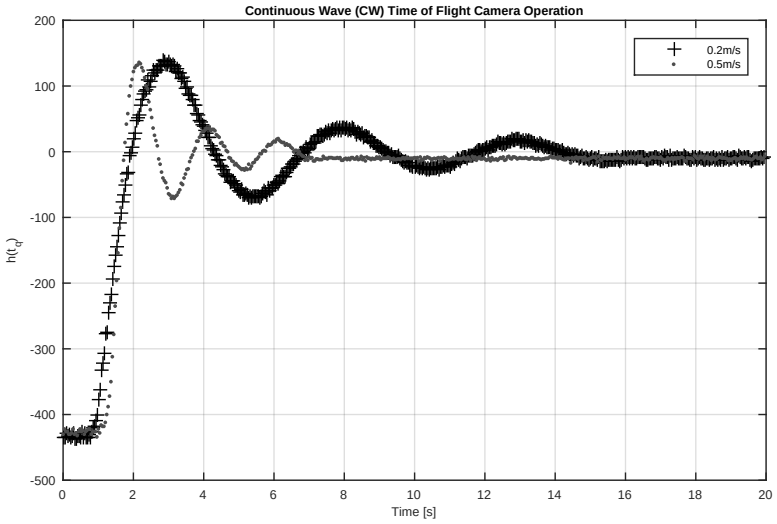


Figure A.1: Measured data in time for an object moving 3 m away from the camera at two different velocities on a translation stage.

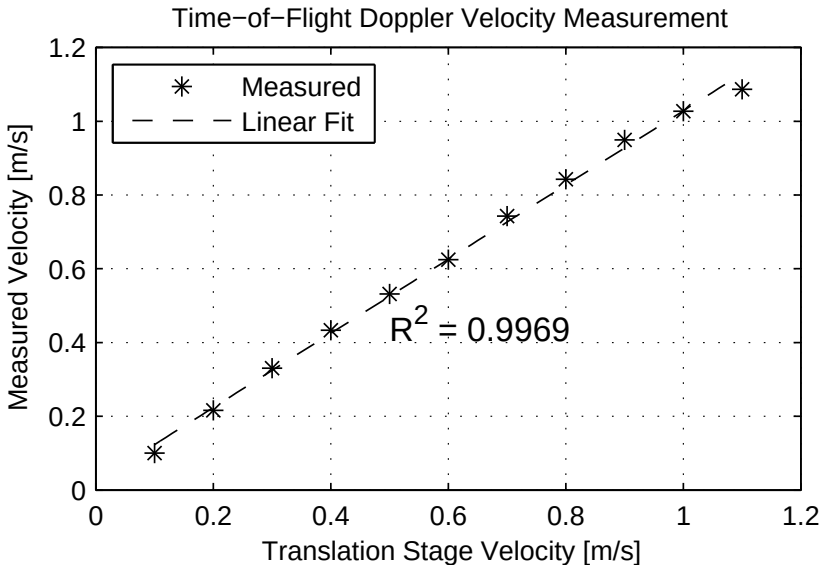


Figure A.2: Results of using time-of-flight range cameras in continuous wave mode to measure velocity.

modulation frequency of 150 MHz is 1 m, therefore three oscillations are observed, and the decay in amplitude is due to increase in distance. As expected the faster moving object's signal contains a larger frequency. The estimated velocity and actual velocity is plotted in figure A.2. An F-test for linearity between the object velocity and the measured velocity is significant ($F = 2893, \rho = 1.33 \times 10^{-12}$) which supports the claim that the velocity estimation is accurate.

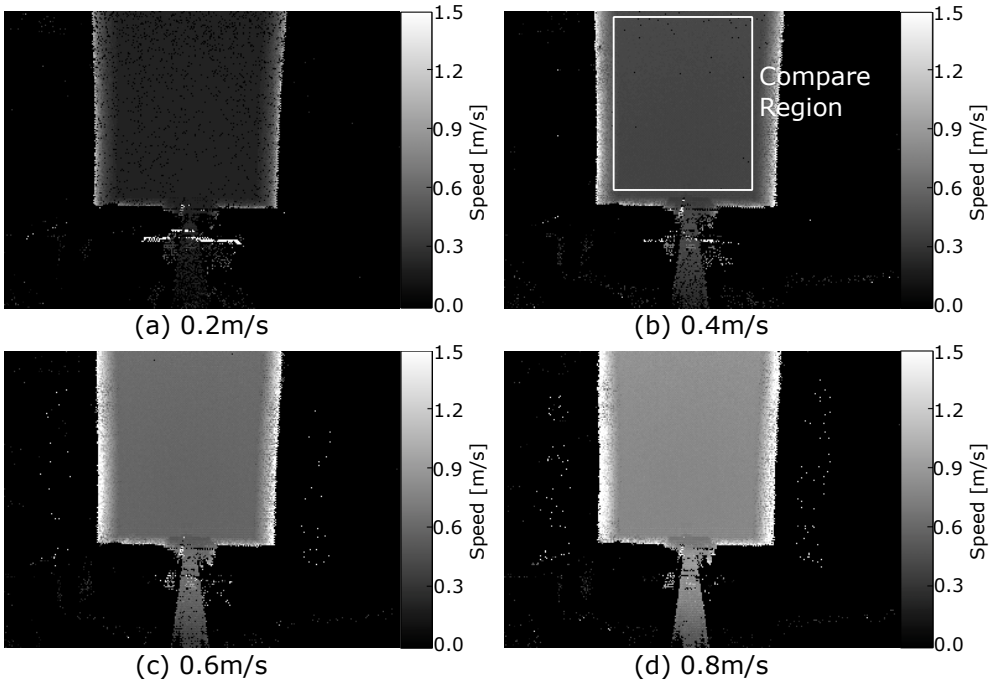


Figure A.3: CW velocity measurement results of moving cardboard sheet at the same distance away from the camera. In (a), (b), (c), and (d) the translation stage is moving at 0.2 m/s, 0.4 m/s, 0.6 m/s and 0.8 m/s respectively. The mean and standard deviation of the pixels in the compare regions are tabulated in table A.1.

An example of the full field results are plotted in figures A.3 and A.4. In figure A.3 the cardboard sheet is moving at different velocities, with the velocity image is taken when the cardboard sheet is the same distance away from the camera. The mean and standard deviation of the region of pixels on the moving cardboard is tabulated in table A.1. The mean is within one standard deviation of the actual translation stage speed.

In figure A.4 the velocity measurement is plotted at different times when the translation stage is moving at a velocity of 1.0 m/s. The velocity measurement of the cardboard sheet is maintained over the two second measurement window. The pixels on the edge of the cardboard sheet measure the velocity incorrectly because the motion in this region is a combination of radial and transverse motion due to the mixing of the foreground and background pixels.

Table A.1: Comparison between translation stage velocity and the measured for a region of 10878 pixels.

Translation Stage Velocity m/s	**Measured Mean m/s**	**Measured σ**
0.1	0.102	0.003
0.2	0.203	0.006
0.3	0.305	0.010
0.4	0.406	0.013
0.5	0.507	0.016
0.6	0.609	0.019
0.7	0.709	0.021
0.8	0.809	0.024
0.9	0.912	0.028
1.0	1.010	0.031
1.1	1.111	0.034

A.1.3 Discussion

Techniques from CW doppler radars have been applied to ToF range cameras to measure velocity by the doppler shift of the reflected signal. Motion causes measurement errors in ToF range cameras. With the presented technique the motion can be measured and in some applications may be more informative than the depth data. The results presented in figure A.2 and table A.1 strongly indicate the velocity can be accurately measured with current ToF hardware. The maximum measurable frequency is set by the integration period of the pixel and its modulation frequency. With the current settings of an integration period of 0.5 ms at a modulation frequency of 150 MHz the maximum velocity measurable is 200 m/s. This is much faster than common applications of ToF applications will encounter, such as computer human interaction. The frame rate of 25 fps is limiting; the low frame rate is because the ToF camera is not designed to operate in CW mode therefore can not operate at a higher frame rate in CW mode. Frame rates of 120 fps are common in ToF cameras.

A.2 Multiple Light Sources

In this section the multiplexing of multiple light sources simultaneously using ToF technology is explored. In conventional light source multiplexing each light source is allocated a sequential time slot during the camera's integration period (Gu *et al.*, 2011), as demonstrated in figure A.5. In communications

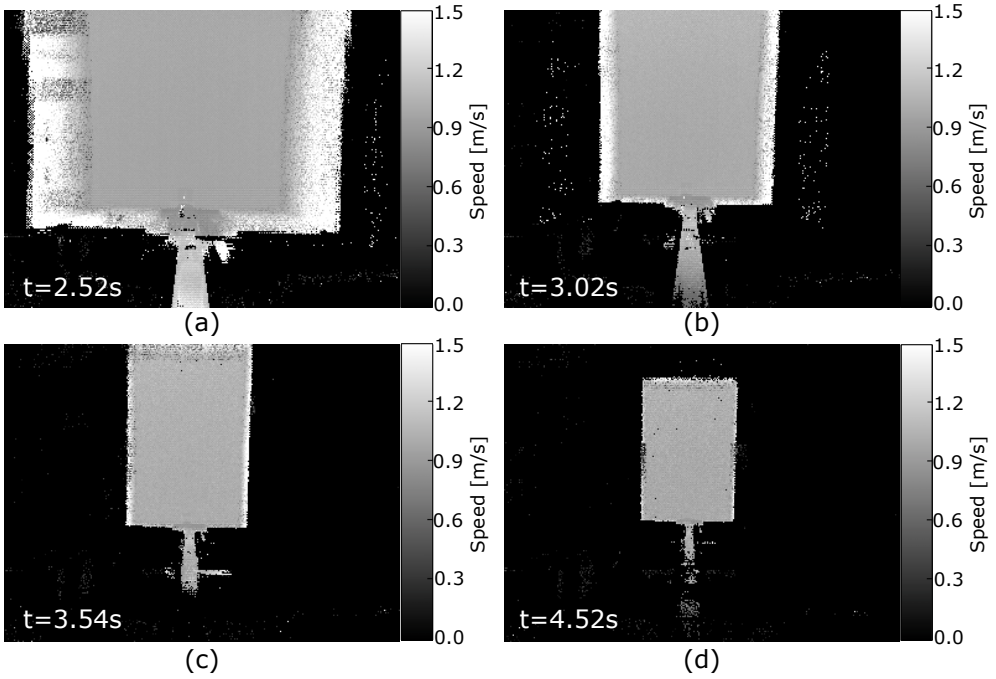


Figure A.4: CW velocity measurements for an object moving 2 m at 1.0 m/s on a translation stage. The time of (a), (b), (c), and (d) is 2.52 s, 3.02 s, 3.54 s, and 4.52 s respectively.

this is called 'time division' multiplexing and it is one method of sharing a communications channel. Motion causes issues for dynamic scenes as discrepancies between each light source are created. With the additional information embedded in a ToF range camera different channel multiplexing techniques are explored. Using ToF technology three methods of operating multiple light sources simultaneously are presented in this section. Multiplexing of multiple light sources is used for dynamic scene relighting, and colour range imaging.

Previous work on using multiple light sources with ToF range cameras has used sequential light multiplexing. Kim *et al.* (2014) used four light sources to apply photometric stereo-based techniques to enhance the measured depth images. Muttayane (2006) investigated using LEDs to perform colour range imaging. True colour was unachievable due to missing visible light wavelengths in the available LEDs. Both applications could benefit from higher frame rates and fewer artefacts by simultaneously multiplexing the light sources.

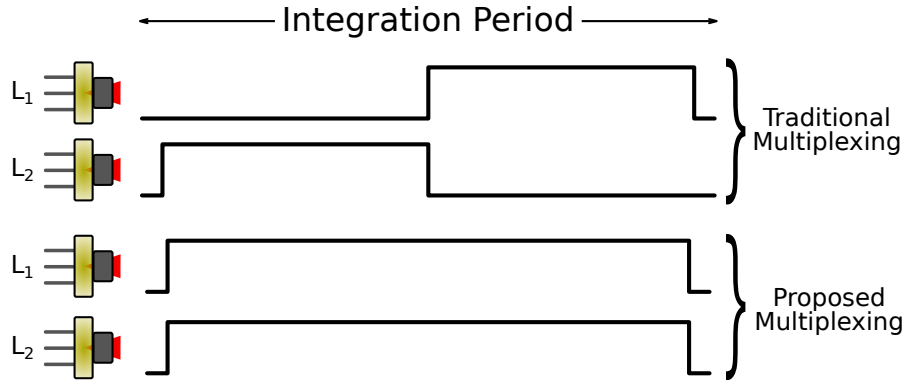


Figure A.5: Difference between traditional light multiplexing and proposed multiplexing using ToF technology.

A.2.1 Theory

Multiplexing multiple light sources by modulating in frequency, phase, and orthogonal spaces is presented. If each light source is modulated at frequencies with a common integer divisor, then for Ξ light sources the reflected light is the sum of all light sources

$$s(t) = \sum_{\xi=1}^{\Xi} a_\xi \cos(\xi\omega\tau + \phi). \tag{A.11}$$

A binary sequence contains multiple modulation frequencies, as demonstrated in chapter 6. By modulating the sensor with a binary sequence which contains the same frequencies then each light source can be demultiplexed. The reference binary sequence, $g(t)$, can be represented by its Fourier components

$$g(t) = \sum_{\xi=1}^{\Xi} \alpha_\xi \cos(\xi\omega\tau), \tag{A.12}$$

where α is the power of the ξ^{th} frequency of the DFT of the binary sequence. The cross-correlation is then

$$h(\tau) = \sum_{\xi=1}^{\Xi} \frac{\alpha_\xi a_\xi}{2} \cos(\xi\omega\tau + \phi). \tag{A.13}$$

If the Nyquist sampling criterion of $N > 2\Xi$ is met, then the amplitude and phase of each light source can be recovered by taking the DFT on the measured cross-correlation. The spectrum of the binary sequence is not flat, therefore the amplitude of each light source is adjusted, and this can be corrected with

calibration. Many multi-path interference removal and phase unwrapping algorithms use measurements taken at multiple modulation frequencies. By multiplexing in frequency all the information is collected simultaneously and further processing can be applied. Binary sequences spread the power across multiple frequencies, with some power being spread into frequencies that are below the noise floor, and this inefficiency is a disadvantage of frequency multiplexing light sources.

The next multiplexing technique is phase modulation. ToF range imaging is a multi-shot measurement, and between each measurement the light source's phase is shifted. Typically four phase shifts are employed for depth measurement (see chapter 2). Phase multiplexing is achieved by phase shifting each light source at a different rate compared to each other. An example of this is using four phase steps the first light sources samples, $[0, \pi/2, \pi, 3\pi/2]$, and the second light source phase shifts at twice the rate resulting in, $[0, \pi, 0, \pi]$, after phase wrapping. If the rates are integer multipliers of ξ then the reflected light is

$$s(t) = \sum_{\xi=1}^{\Xi} a_\xi \sin(\omega \xi t + \phi). \tag{A.14}$$

The sensor modulation and measured cross-correlation is

$$g(t) = \sin(\omega t), \tag{A.15}$$

$$h(\tau) = \sum_{\xi=1}^{\Xi} a_\xi \cos(\omega \xi \tau + \phi). \tag{A.16}$$

Phase encoding is similar to frequency encoding, as each light source shows up as a different frequency in the measured correlation signal, but each light source is amplitude modulated at the same frequency. Phase encoding is advantageous because no power is wasted. An example of the measured correlation with three light sources, phase shifted at different rates, is plotted in figure A.6.

The previous two multiplexing techniques employed the Fourier bases to encode each light source. The sensor and each light source are modulated with different binary sequences. By demodulating the measured correlation using another appropriate transform (which we explain) each light source can be

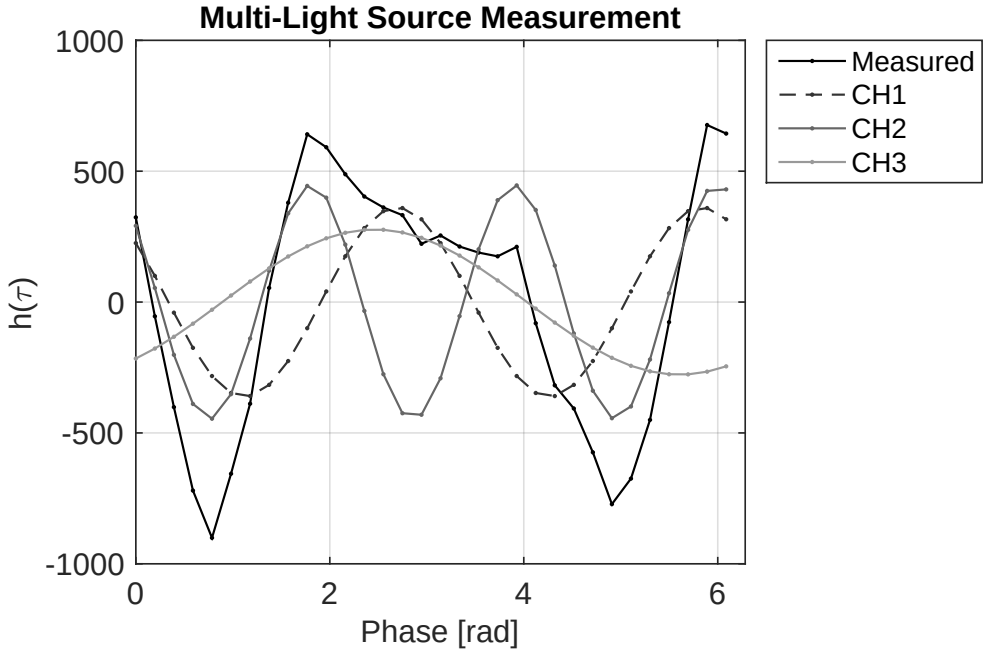


Figure A.6: Three light sources multiplexed using phase multiplexing. The measured cross-correlation is a combination of all three channels (CH1, CH2, and CH3), and is demultiplexed by the discrete Fourier transform.

demultiplexed. In matrix form the measured correlation vector $\mathbf{h}$ is

$$\mathbf{h} = \mathbf{TLx}, \tag{A.17}$$

where $\mathbf{x}$ is the scene response, in this case the reflectance. The matrices $\mathbf{L}$ and $\mathbf{T}$ are the Toeplitz matrices of the illumination and reference binary sequences. Each light source has a unique binary sequence. To demodulate each light source the inverse matrix, $\mathbf{L^{-1}}$, is chosen so that

$$\mathbf{L}^{-1}\mathbf{TL} = \mathbf{I}, \tag{A.18}$$

$$\mathbf{L^{-1}h} = \mathbf{x}. \tag{A.19}$$

The selection of binary sequence for the light sources and sensor is critical. Each light source has to be recoverable and its binary sequence's Toeplitz matrix must be invertible with sufficient rank. The limiting factor is that only binary matrices are possible with the current hardware. An optimization problem was solved to find usable sequences. With the presented method only the amplitude of each light source is recoverable, and if a phase shift is present then the current demodulation fails. Therefore only very slow modulation frequencies are possible.

(a) Left Source (b) Right Source (c) Photograph

Figure A.7: Light multiplexing of scene using two light sources. (a) is the measured amplitude using the left light source, (b) is the measured amplitude using the right light source and (c) is a photograph of the scene. The shadows, specular highlights, and inter-reflections change with the light source.

Three methods have been presented to simultaneously multiplex multiple light sources using ToF range camera technology. Multiplexing in phase is the most efficient with no wasted power. Multiplexing in frequency allows for measuring multiple light sources and frequencies, however some spectral power is wasted. Multiplexing in orthogonal spaces using binary sequences can measure amplitude but not phase and therefore is very limited.

A.2.2 Light Multiplexing

The interaction of light with objects is dependent on the incident angle of light. The shadows and occlusions cast are dependent on the illumination source's location. Multiplexing multiple light sources allows for the different interactions of light with the scene to be observed, with applications to dynamic relighting, photometric stereo, and further analysis of light transport for computer graphics applications.

Figure A.7 is the result of measuring two light sources, one light source located on the left and the other on the right. The specular highlights on the vase and CD are highly dependent on the light sources' locations.

The results of four light sources for dynamically relighting a scene is shown in figure A.8. The shadows cast and specular highlights on the beer glass and golden chalice depend on the light source location. The CD acts as a mirror. When light sources L1 and L2 illuminate the CD, it reflects light onto the playing card. When light sources L3 and L4 illuminate the playing card and the mirror, the image of the playing card is visible in the CD. With four light sources nine samples are required to recover each light sources amplitude and

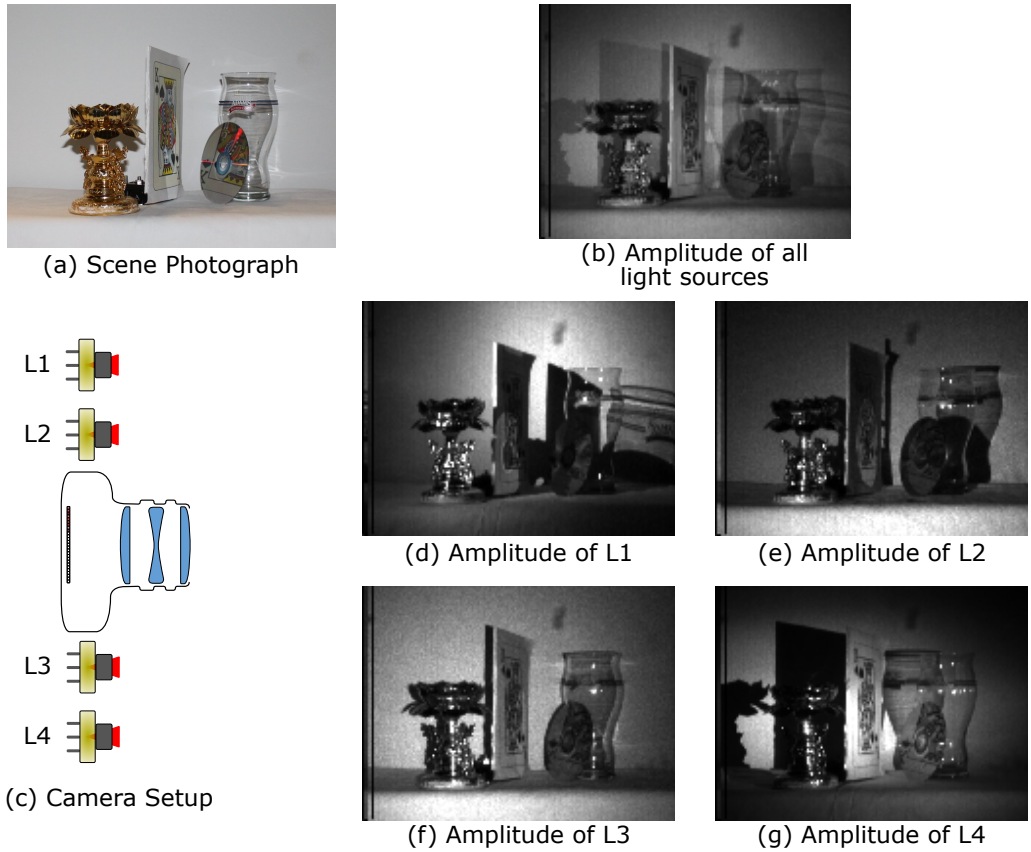


Figure A.8: Relighting of scene using four light sources. (a) is a photograph of the scene, with (b) being the measured amplitude with all light sources illuminating the scene. (c) is a diagram of the camera setup. The individual measured amplitudes are in (d) to (g). The specular reflections and shadows cast are highly light source dependent.

phase, and it is possible to acquire this data at 30 fps. The use of light field photography, shape from shading, photometric stereo and other computational imaging techniques can be used to extract more information from the scene.

A.2.3 Colour Time-of-Flight

ToF range cameras use an active light source to illuminate the scene and measure the amplitude and phase of the reflected light. Hyperspectral imaging measures the electromagnetic spectrum for each pixel and has applications in agriculture, mineralogy, and food processing. By multiplexing multiple light sources of different wavelengths a ToF camera can be transformed into a hyperspectral imaging camera that measures both the amplitude and phase of each light source's wavelength (spectral bandwidth).

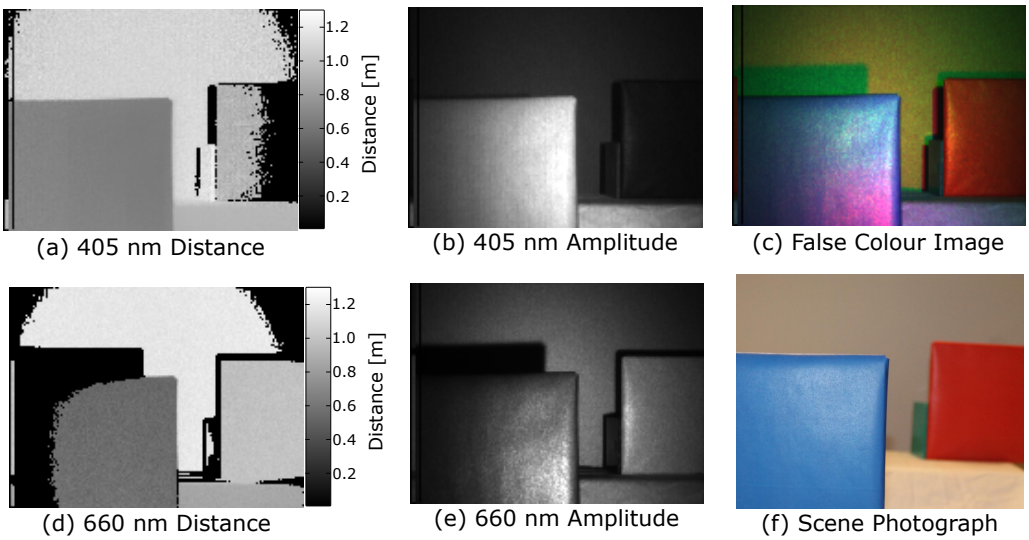


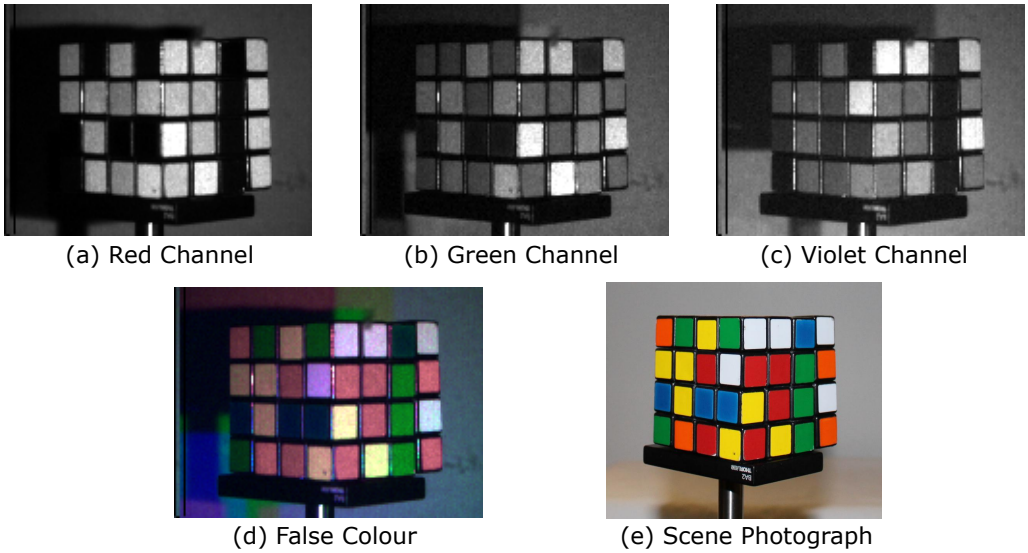
Figure A.9: Combined colour and depth time-of-flight range camera operation. The violet (405 nm) light source distance and amplitude measurements are in (a) and (b). The red (660 nm) light source distance and amplitude measurements are in (d) and (e). The false colour image generated from (b) and (e) is in (c), and the actual scene photograph in (f).

Colour range imaging was performed by illuminating the scene with 405nm, 532nm and 660nm wavelength laser diodes. The amplitude from each light source determines the colour and the phase encodes the depth. An example of phase and amplitude measurements with a 405 nm and 660 nm laser diodes is in figure A.9. The scene photo does not include many depth clues while the depth and colour response can be measured with a modified ToF range camera.

Examples of real-time colour time of flight range imaging is shown in figure A.10. The individual colour channels are measured and the combined result allows for colour range imaging. The light sources are not collocated therefore occlusions are present in the colour images causing the colour artifacts.

A.2.4 Discussion

More information than just relighting and colour could be computed from the results as the interaction of light with the scene is angle and wavelength dependent. It is expected that multiple light sources can be used to produce a light field instead of multiple imaging sensors, and is advantageous as light sources are cheaper than imaging sensors.



(a) Red Channel (b) Green Channel (c) Violet Channel

(d) False Colour (e) Scene Photograph

Figure A.10: Real time colour measurement with time-of-flight range imaging camera of a Rubik's cube. The measured amplitude of each light source is in (a) to (c). The combined colour image in (d) with a photograph of the scene in (e).

There are a number of issues with the illumination of color ToF range imaging. The laser diodes have a narrow wavelength emission. Previous work on colour ToF range imaging using LEDs has shown gaps in the illuminating wavelength (Muttayane, 2006). The produced colour images are only sensitive to select wavelengths. A more pressing issue is that at 405nm a significant number of objects fluoresce. The fluorescent life time of the sample will change the measured returning phase. Currently no solid state green laser diode exists. Green lasers are created from an IR source that is frequency doubled using a crystal. The frequency doubler crystal makes green laser diodes unstable if amplitude modulated, therefore limiting the modulation frequency to the kilohertz range. The quantum efficiency of silicon is a function of wavelength, with the efficiency decreasing above 450nm. The performance of the 405nm light source is limited by the quantum efficiency of the silicon sensor.

The larger picture colour ToF imaging is the equivalent to hyperspectral imaging where the amplitude and phase are measured. There are many applications of hyperspectral imaging from agriculture, eye care, food processing, fluorescence lifetime imaging and mineralogy. Fluorescence lifetime imaging is currently the application that benefits most from measuring both the phase and amplitude. Other applications are possible and this advantage of phase information is to be explored further.

A.3 Fluorescence Life Time Imaging

Fluorescence is caused by an electron being excited into a higher energy state, when the electron returns to the ground state a photon is emitted. The lifetime of a fluorophore is the average time between the excitation of an electron and its return to the ground state. Fluorescence lifetime imaging (FLI) is an imaging technique for measuring the lifetime of a fluorophore signal. FLI has applications in biotechnology, flow cytometry, medical diagnostics, and DNA sequencing to name a few. Currently two techniques are widespread in resolving the fluorescence lifetime, time domain and frequency domain methods. In time domain measurements the sample is excited by a narrow pulse and the sample intensity with time is measured. The lifetime, τ, is calculated from the gradient of the decay (Lakowicz, 2006). In frequency domain measurements the sample is excited by amplitude modulated light, the lifetime is encoded as a phase shift in the amplitude modulation envelope (Lakowicz, 2006). The lock-in pixels in ToF range cameras are designed to measure phase shifts in amplitude modulation, making them ideal for FLI. In this section a ToF camera is modified to measure the fluorescence life time of a sample. The issues of obtaining an accurate lifetime measurement are explored.

A.3.1 Theory

In order for fluorophore to fluoresce it has to be excited, in our case we are using an amplitude modulated light source. The use of amplitude modulated light for frequency domain lifetime measurements is a well known technique and described by Lakowicz (2006). The amplitude modulation of the excitement is given by the ratio of the average intensity a and the peak to peak height b. The modulation of the emission is defined by the same ratio of the average emission intensity and peak to peak emission height. In fluorescence the modulation, m, is defined as

$$m = \frac{aB}{Ab}. \tag{A.20}$$

The phase, ϕ of the emitted light encodes the lifetime τ_ϕ by

$$\tau_\phi = \frac{\tan(\phi)}{\omega}. \tag{A.21}$$

The modulation also encodes the fluorescence lifetime by

$$\tau_\phi = \frac{1}{\omega}\sqrt{\frac{1}{m^2} - 1}. \tag{A.22}$$

The phase shift in a ToF range camera is the combination of the phase shift from the lifetime and phase shift induced by the distance travelled

$$\phi(\omega) = \frac{2\omega d}{c} + \tan^{-1}(\omega\tau), \tag{A.23}$$

where c is the speed of light. Both the distance and fluorescence lifetime are unknown in equation A.23. The two unknown variables are estimated by taking multiple measurements at different modulation frequencies and minimizing the objective function

$$J(d, \tau) = \hat{\phi}(\omega) - \frac{2\omega d}{c} - \tan^{-1}(\omega\tau), \tag{A.24}$$

where $\hat{\phi}(\omega)$ is the phase measured by the ToF camera with each modulation frequency. Measurements at multiple modulation frequencies are taken to allow the objective function in equation A.24 to converge on the correct distance, d, and lifetime, τ.

A.3.2 Experimental Setup

The sample is a glass slide which contains a dissolved CdS quantum dot sample, which is dissolved in hexane and PMMA. The slide is tilted to stop specular reflections off the glass slide and placed in front of a black non-fluorescing cardboard sheet. The scene is illuminated by an amplitude modulated 405 nm laser diode. A dielectric interference filter (cut-off wavelength of 450 nm) is placed between the sample and ToF camera, the emitted photons from the fluorophore have a longer wavelength than the incident photons. The dielectric filter removes photons that come directly from the light source and contain no information on the fluorescence lifetime. therefore the incident signal at the camera is phase shifted due the lifetime and distance travelled. The difference between the non-fluorescing background and quantum dot encodes the lifetime. At 405 nm many objects fluoresce, including many optical elements. Special optical elements are used in FLI imaging systems, but were not available for this setup. We sweep the modulation frequency of the ToF camera from 1 MHz up to 50 MHz in increments of 1 MHz to collect 50 samples to use in equation

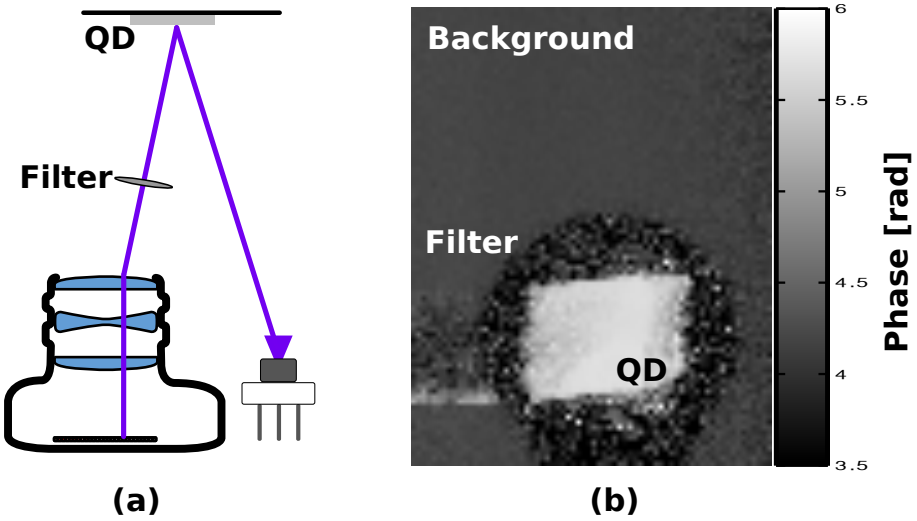


Figure A.11: (a) Experimental setup. A violet laser diode illuminates the scene which excites a fluorescent quantum dot. (b) Sample phase measurement at 40 MHz with the quantum dot and background.

A.24. The ground truth is directly measured by a femtosecond laser. The experimental setup is shown in figure A.11.

A.3.3 Results and Discussion

The results are plotted in figure A.12. The sample was measured with a direct time of flight system to obtain the ground truth lifetime of 32 ns. The ToF range camera measures the lifetime to be 32 ns, which is correct within 2 ns. The recovery time is limited by the bandwidth of modulation frequencies of the ToF camera. The bandwidth of ToF range cameras is expected to continue improving making them more ideal for FLI applications. They have the advantages of being a full field imaging system and are low cost compared to current FLI devices.

A.4 Conclusion

This chapter has explored three unique applications of ToF range camera technology. These are measuring radial velocity of objects, multiple light sources for dynamic relighting and colour range imaging, and fluorescence lifetime imaging (FLI). Imaging with multiple light sources at different wavelengths is equivalent to hyperspectral imaging. An application where the phase and amplitude of multiple wavelengths contains useful information is FLI, as different materials fluoresce at different wavelengths. As the bandwidth of ToF cameras

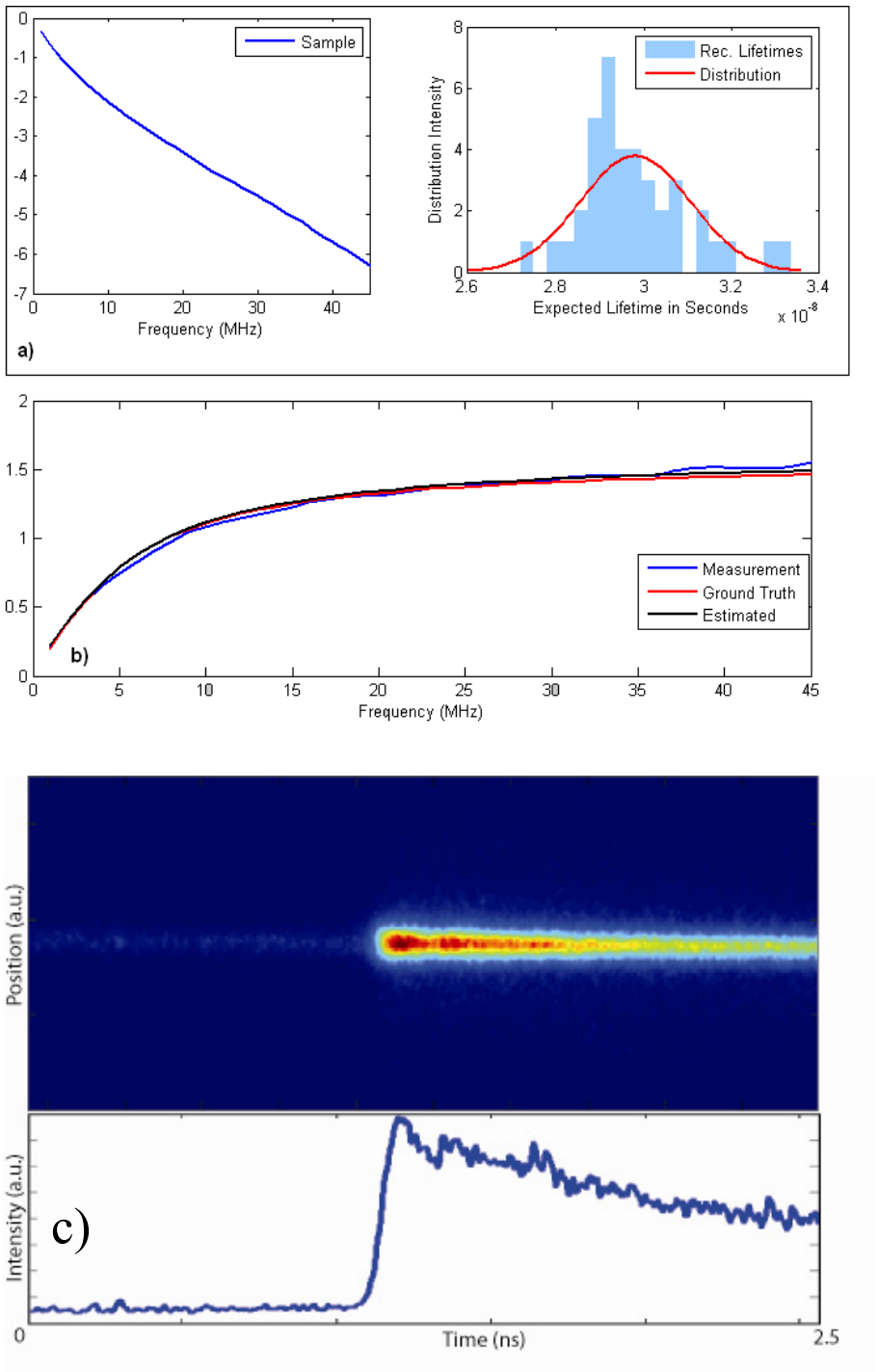


Figure A.12: Lifetime recovery without a reference signal. (a) Numerical fitting results compared to measured phase response of quantum dot pixel; a 30 ns lifetime is recovered. (b) Difference between fluorescent pixel and background with an inverse tangent ($\tau = 32$ ns). (c) Time-resolved measurement of decay provides independent verification of 32 ns lifetime.

increases then the lifetime estimation and velocity measurements will both improve. The use of multiple light sources has demonstrated basic applications and further research could yield more interesting results. This chapter has demonstrated that ToF camera and pixel technology has applications outside of range imaging.

Appendix B

PMD19k Camera User Guide

The camera is designed around a pixel array sensor created by PMD technologies, a time-of-flight camera manufacturer based in Germany. The sensor is the PMD19k3 with a array size of 160 by 120 pixels. The camera design is based off a reference design by Jongenelen Jongenelen (2011). The camera composes of four printed circuit boards the names and functionally are described in table B.1. An overview of the camera design is shown in figure B.1.

A limited UDP stack is implemented on the FPGA for communication to external devices. A Matlab API is developed for configuring and controlling the camera. JTAG programming of the FPGA is provided through the USB blaster port on the FPGA development board. The JTAG also privides limited debugging from the Nios II processor.

Table B.1: Description of the PMD19k3 camera components.

Name	**Functionally**
Illumination	Provides a light source to the camera. A variety of light sources are available. The light source is connected to the camera by a SATA cable carried two LVDS pairs.
RedEye	Interface PCB between the sensor, illumination, and FPGA PCBs. Interfaces the modulation and readout to the sensor, and reading of ADC values. LVDS from the FPGA is converted to single ended.
FPGA	The DE2-115 Cyclone IV development kit. The FPGA controls the camera and communicates data though the ethernet port.
DLP	The DMD based projector from Texas Instruments.

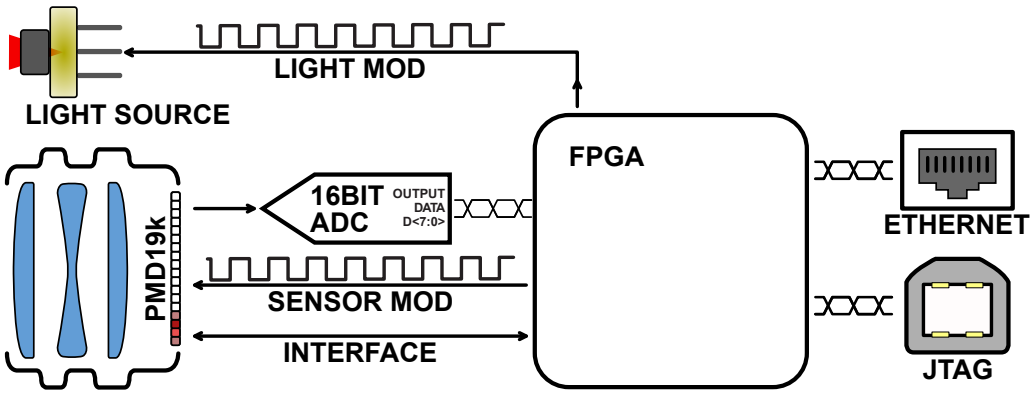


Figure B.1: Main components of the PMD19k3 time-of-flight range camera.

Board	Power Supply
Illumination	6V
RedEye	6V
FPGA	Power Pack supplied
DLP	Power Pack supplied

Table B.2: PMD19k3 time-of-flight range camera power supply requirements.

Table B.3: Short description of major FPGA components.

Name	Description
PMD19k3 Interface	Controls the readout of data from the camera.
Modulation Controller	Sets the frequency phase and binary sequence of the modulation signals.
UDP Packetizer	Hardware UDP stack reads the ADC values and forms UDP packets to transmit
Triple Speed Ethernet	Hardware MAC the IP is provided by Altera.
Nios II	Softcore microprocessor running software UDP stack.

B.0.1 Hardware

The camera consists of four individual PCBs, each with there own power supply. Table B.2 details the power supply requirements for each PCB.

B.0.2 FPGA Design

The major internal components of the FPGA are shown in Figure B.2. A short description of each component is in table B.3.

The components are connected with the Avalon Interconnect allowing for portability and reusability between design in Altera FPGAs using the Qsys

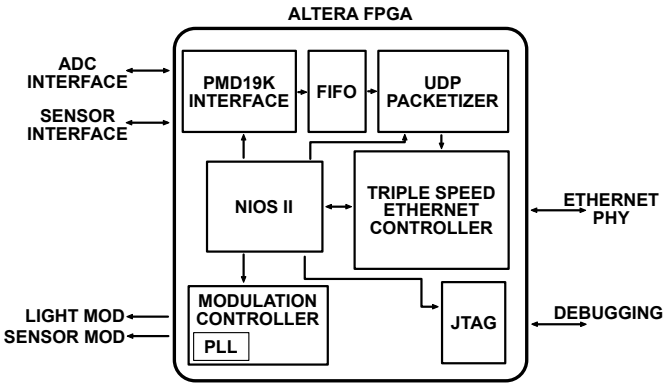


Figure B.2: Internal components of the camera's FPGA design

builder tool suit. The IP developed should be portable for other applications and camera systems. Each IP block was designed in VHDL.

The TSE is designed for up to 1Gbps Ethernet conductivity. The readout speed of the pixel data values is approximately 8MHz this produces 256Mbps of data. Even with the UDP and Ethernet overheads the bottle neck is the readout time of the sensor. The camera can only work on 1Gbps network links due to a bug in the TSE configuration.

Register Map

The camera is configured by a number of register spaces that are interconnected with Avalon Memory map. The Nios II microprocessor is the Avalon Memory Master, the rest of the devices are slaves. The offset address of each space can be located in the Qsys project file. The TSE register map is documented in by Altera (2013b). The ADC register map is a direct interface to the internal registers of the AD9826 and is documented by Analog Devices (2012). The overview of the UDP Packetizer, PMD19k Modulation Control and PMD19k Control is described in tables B.4 to B.6.

Offset	Name	R/W	Description
0x0000	Frame Period	R/W	Sets the frame period in microseconds. The frame period is the integration time plus the readout time.
0x0001	Integration Period Requested	R/W	The integration time in microseconds requested.

0x0002	Integration Period Actual	R	The actual integration period in microseconds.
0x0008	Frames per Output Frame	R/W	Unused
0x0011	FPB	R/W	Number of phase steps in a range frame.
0x0018	Modulation Type	R/W	Controls the binary sequence used to modulate the light source and sensor. Four operating modes and each mode can be static or dynamic. In dynamic modes the PLL clocks phase offset is updated each frame by the value in the `Step Size` register. The operating modes are: 0: AMCW, 1: AMCW Static, 2: PN, 3: PN Static, 4: BPSK, 5: BPSK Static, 6: Shift Register, 7: Shift Register Static
0x0019	MOD Reset	R/W	Unused.
0x001D-23	C0-C6 Start Step	R/W	The starting phase step of PLL clocks.
0x0024-2A	C0-C6 Step Size	R/W	In non-static modulation modes these registers set the increment in phase steps of each PLL clock between frames.
0x002B-31	C0-C6 Step	R/W	In static modulation modes these registers set the phase step of each PLL clock.
0x0032	Modulation Duty Cycle	R/W	Bits 0-6 are used as the `rselodd` bit in the corresponded C0-C6 PLL reconfiguration stream. Setting to '1' forces 50% duty cycle from PLL.

0x0033	Frequency Reconfig Enable	R/W	Sets the `bypass` bit of the PLL reconfiguration stream. The 7 LSB bits correspond to the `bypass` bit in the C0-C6 PLL clocks. Setting the bit to 1 disables reconfiguration of the selected clock.
0x0034	PLL Reconfig	W	Writing '1' to this register reconfigures the modulation PLL. This bit is cleared in hardware.
0x0035-42	C0-C6 Hi and Low PLL dividers	R/W	The hi and low PLL counter set the output frequency of each clock by dividing the PLL's VCO frequency.
0x0043-44	M Hi and Low	R/W	The M counter value used in setting the PLL VCO frequency.
0x0045-46	N Hi and N Low	R/W	Not yet implemented.
0x0047	M-sequence Polynomial	R/W	The primitive polynomial used to generate the maximum length sequences.
0x0048	Code Length	R/W	The length of the maximum length sequence as defined by 0x0047.
0x0049	Shift Length	R/W	The length of data in the modulation shift register.
0x0100-13F	Light Register	R/W	The initial values of the light shift register.
0x0140-17F	Sensor Register	R/W	The initial values of the sensor shift register.
0x0180-01BF	Spare Register	R/W	The initial values of the spare shift register.

Table B.5: Overview of PMD19k Modulation Controller Register Space.

Offset	**Name**	**R/W**	**Description**

0x0	Reset	R/W	Bit 0: Reserved. Bit 1: Reset UDP connection. Bit 2: Reserved. Bit 3: Reset PMD19k3 interface.
0x1	Pause	R/W	Writing '1' to this register should pause the camera operation. NEVER TESTED.
0x2	FPGA Version	R	Register value is FPGA build version.
0x3	Nios Version	R	Register value is the Nios Software build version.
0x4	Light Enable	R/W	Select which light outputs are active.

Table B.6: Overview of PMD19k Control Register Space

B.0.3 Nios II UDP Stack

The Nios II provides an interface between the Ethernet and the registers. UDP packets can read or write registers. The format of the UDP packet is defined as:

R/W Register Space Offset Address Data

All the parts of the packet are represented by 32bit long unsigned integers. The data part can be an array. The R/W word has three values, 0 is defines a read, 1 is a write and 2 is a returning packet from the camera. The values that the register space can have are defined in table B.7. The offset address is the registers offset address as defined in the tables B.6 and B.4. The data is a single 32bit value or an array of 32bit values. The UDP Stack allows for changing global variables in the Nios II. The variables are accessed in the same method as the FPGA registers space the equivalent register space is in table B.8. It is possible to expand this to be able to call routines.

B.0.4 Known Issues

Below are a list of known bugs and issues with the camera and how to avoid and correct them.

- During camera boot if the RedEye PCB is powered after the FPGA is programmed the RedEye PCB draws a large amount of current until the

Table B.4: Overview of UDP Packetizer Register Space.

Offset	Name	Description
0x00	Configure	Bit 0 enables streaming of ADC values. Setting bit 1 starts the computation of the IP header checksum. This must be set before valid packets can bet streamed or if any field of the IP header changes. Bit 2 enables UDP Packetizer debugging mode. Bits 0 and 2 are never to both be set at the same time the behavior is undefined.
0x01	MAC Dst Hi	32 MSB of the destination MAC address.
0x02	MAC Dst Low	16 LSB of the destination MAC address.
0x03	MAC Src Hi	32 MSB of the source MAC address.
0x04	MAC Src Low	16 LSB of the source MAC address.
0x05	IP Src Addr	Camera IP address in network byte order.
0x06	IP Dst Addr	Destination IP address in network byte order.
0x07	UDP Dst Port	UDP Destination port.
0x08	UDP Src Port	UDP Source port.
0x09	Data Length	Number of 32bit words of ADC data to transmit per packet. Recommended to use 0x13F.
0x0A	Current State	Debugging register to read current state of the UDP Packetizer.
0x0B	Ready Counter	Debugging register.
0x0C	ADC Level	Debugging register, contains the current level of the data FIFO.
0x0D	Dropped Counter	Debugging register, records number of dropped frames.
0x0E	Dump level	Debugging register.
0x0F	Status	Debugging register.
0x10	Abort level	Debugging register.
0x011	End of Packet Count	Debugging Register.
0x012	Stop of Packet Count	Debugging Register.
0x013	Frame Count	Debugging register, allows read access to the frame count register.
0x014	Row Count	Debugging register, allows read access to row count register.

Table B.7: Register space base address from Nios II UDP interface.

Register Space Address	Register Space Name
0	Modulation Control
1	UDP Packetizer
2	Camera Control
3	ADC
4	Nios Variables
5	TSE

Offset	Name	R/W	Description
0	Nios Source IP Address	R/W	The IP address used by the Nios II UDP Stack.
1	Nios Destination IP Address	R/W	
2	Nios UDP Source Port	R/W	
3	Nios UDP Destination Port	R/W	
4	Nios Source MAC Address	R/W	Needs to be the same as the TSE source MAC address for correct operation.
5	Nios Destination MAC Address	R/W	
6	Nios Reset Camera	W	

Table B.8: Nios II Global variables register space.

Nios II is programmed. To avoid the only power the RedEye PCB when the FPGA is fully programmed.

- Sometimes every 3rd column of the readout is invalid. The fix is reprogramming the FPGA.

- Sometimes the columns are mixed during readout. The fix is to reset the camera from Matlab.

- Sometimes the Ethernet fails on the camera. This has been observed to correlate with higher temperatures. Cool down the FPGA.

Function	Description
PMD19kConfigADC(camera)	Configure the ADC.
PMD19kConfigInterface(camera)	Configure the PMD19k3 interface.
PMD19kConfigModulation(camera)	Configure the modulation controller.
PMD19kConfigPLL(camera)	Configure the modulation PLL.
PMD19kConfigUDP(camera)	Configure the UDP Packetizer.
PMD19kResetCamera(camera)	Reset the camera.
PMD19kSetLights(camera)	Configure the lights
PMD19kConfigInterface_ADCTesting (camera)	Configure the ADC timing. Must call during configuration of camera.
PMD19kResetInterface(camera)	Reset the PMD19k3 interface.
pmd19k_send_packet (camera.connection,pkt)	Send a command packet to the Nios II processor.
pmd19k_capture_stream (camera.stream_con,camera)	Capture a stream of frames to disk. Saves data contained in UDP packet to disk.
pmd19k3_process_capture (inFile,outFile,maxFrames)	Process the captured in coming packets on the UDP stream into raw frames. Writes the raw frames to a binary file.
PMD19k3ProcessedToRawMean (captureName)	Process the binary file containing raw frames to matlab files.

Table B.9: Matlab top level API functions.

- The Nios II Ethernet and IP address are fixed. In theory they can be reconfigured however this has bugs. The addresses have to be hard coded at compile time.

B.0.5 Matlab API

The Matlab API interfaces to Mex function that use socket programming to read and write UDP packets. All configuration functions are passed a structure that contains the relevant information. This section over views the camera configuration structure and functions.

The following table describes the variables in the `camera` structure used to configure the camera.

Variable	sub	Description

modType		Modulation type. 0 for square wave and 6 for binary sequence. Sets register 0x0018 in Modulation Control registers.
integrationTime		Integration time of raw frame in microseconds.
frameTime		Total frame time including readout and integration time in microseconds. If same as integrationTime than integrationTime is automatically reduced.
pll	dutyCycle	Should set duty cycle bit of PLL. Does not work.
	freqReconfig Enable	Sets the reconfigEnable of the corresponding PLL counter. Only needed when mhigh or nnigh are set to 0.
	mhigh	
	mlow	
	nlow	
	nhigh	
	counters	Vector of 14 elements. Contains [hi low hi low] of the PLL down counters.
shiftRegisters	sensorCode	Binary sequence to be sent to the light source.
	lightCode	Binary sequence to be sent to the sensor.
	spareCode	
	spareCode2	
	spareCode3	
startStep		Vector of 7 elements. Each element the initial phase offset of each clock output of the PLL.
staticStep		Unused.

nSteps		Number of phase steps in each depth frame.
stepSize		Vector of 7 elements. Each element is the phase step size between raw frames.
adc	gainLeft	
	gainRight	
	gainMid	
	offsetLeft	
	offsetRight	
	offsetMid	
lights		The light sources to enable.
light_default		Set the default state of each of the light outputs. Currently set to 0.
light_pllsel		Set to 0.
connection	camera_ip	String contains IP address of camera. Must be set to '192.168.202.5' unless changed in Nios II code.
	src_port	The source port of the UDP connection to the Nios II. Must be set to 2223 unless changed in the Nios II code.
	dst_port	The destination port of the UDP connection to the Nios II. Must be set to 54155 unless changed in the Nios II code.
stream_con	camera_ip	The IP address of the camera. Must be set to '192.168.202.5'.
	src_port	The source port of the UDP stream containing camera images. Must be set to 2225.

	`dst_port`	The destination port of the UDP stream containing camera images. Must be set to 54156.
`capture`	`nFrames`	Total number of raw frames to capture.
	`dumpFrames`	Number of frames to dump before capturing.
	`pktDataSize`	Must set to `hex2dec('13F')`.

Table B.10: Internal variables in `camera` structure.

B.0.6 DMD API

The following functions are provided to configure the DLP projector from Matlab. The function from Matlab is `dmd_displayStaticImage(imgFile)`. Where `imgFile` is a string containing the file and location of the image wanting to be displayed.

Warning: The DMD does not throw an error message to Matlab if it fails to write the image to the projector. Make sure the correct image has been displayed. Power cycle the projector if image writing fails.

References

Altera. *Stratix III Development Kit Reference Manual*. Altera, San Jose, CA, USA (2008a). Rev 1.2.

Altera. *Using PLLs in Cyclone Devices*. Altera, San Jose, CA, USA (2008b).

Altera. *Stratix III Device Handbook, Volume 1*. Altera, San Jose, CA, USA (2011). Rev 2.2.

Altera. *Cyclone IV Device Handbook, Volue 1*. Altera, San Jose, CA, USA (2013a).

Altera. *Triple-Speed Ethernet MegaCord Function User Guide*. Altera, San Jose, CA, USA (2013b).

Analog Devices. *Complete 16-Bit Imaging Signal Processor*. Analog Devices, Norwood, MA, USA (2012). Rev B.

Bajwa, W. U., J. Haupt, A. M. Sayeed, and R. Nowak. Compressed channel sensing: A new approach to estimating sparse multipath channels. *Proc. IEEE*, **98(6)**, pp. 1058–1076 (2010).

Bajwa, W. U., J. D. Haupt, G. M. Raz, S. J. Wright, and R. D. Nowak. Toeplitz-structured compressed sensing matrices. In: *14th Workshop on Statistical Signal Processing (SSP)*, pp. 294–298 (2007).

Bamji, C. Method and system to avoid inter-system interference for phase-based time-of-flight systems. US Patent 7405812 (2008).

Bamji, C., P. O'Connor, T. Elkhatib, S. Mehta, B. Thompson, L. Prather, D. Snow, O. Akkaya, A. Daniel, A. Payne, T. Perry, M. Fenton, and V.-H. Chan. A 0.13 μm CMOS system-on-chip for a 512 $\times$ 424 time-of-flight image sensor with multi-frequency photo-demodulation up to 130 MHz and 2 GS/s ADC. *IEEE J. Solid-State Circuits*, **50(1)**, pp. 303–319 (2015).

Benedetti, A., T. Perry, M. Fenton, and V. Mogallapu. Methods and systems for geometric phase unwrapping in time of flight systems. US Patent 20140049767 (2014).

Bhandari, A., A. Kadambi, R. Whyte, C. Barsi, M. Feigin, A. Dorrington, and R. Raskar. Resolving multi-path interference in time-of-flight imaging via modulation frequency diversity and sparse regularization. *Optics Letters*, **39(6)**, pp. 1705–1708 (2014).

Böhme, M., M. Haker, T. Martinetz, and E. Barth. Shading constraint improves accuracy of time-of-flight measurements. *Computer vision and image understanding*, **114(12)**, pp. 1329–1335 (2010).

Burg, J. P. Maximum entropy spectral analysis. In: *37th Annual International Meeting*. Society of Exploration Geophysics (1967).

Büttgen, B., 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, Proc. SPIE*, volume 5302, pp. 9–20 (2004).

Büttgen, B., 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: *Proc. of 1st range imaging research day*, pp. 21–32 (2005).

Büttgen, B. and P. Seitz. Robust optical time-of-flight range imaging based on smart pixel structures. *IEEE Trans. Circuits Syst. I*, **55(6)**, pp. 1512–1525 (2008).

Charvat, G. L. *Small and Short-Range Radar Systems*. CRC Press (2014).

Choi, O., S. Lee, and H. Lim. Interframe consistent multifrequency phase unwrapping for time-of-flight cameras. *Optical Engineering*, **52(5)**, pp. 057005–057005 (2013).

Crabb, R. and R. Manduchi. Probabilistic phase unwrapping for single-frequency time-of-flight range cameras. In: *Proc. International Conference on 3D Vision (3DV 14)* (2014).

Davis, P. J. *Circulant matrices*. American Mathematical Soc. (1979).

Donoho, D. Compressed sensing. *IEEE Trans. Inf. Theory*, **52(4)**, pp. 1289–1306 (2006).

Dorrington, A. A., M. J. Cree, D. A. Carnegie, A. D. Payne, R. M. Conroy, J. P. Godbaz, and A. P. Jongenelen. Video-rate or high-precision: a flexible range imaging camera. In: *Electronic Imaging 2008*, volume 6813, pp. 07–12 (2008).

Dorrington, A. A., M. J. Cree, A. D. Payne, R. M. Conroy, and D. A. Carnegie. Achieving sub-millimetre precision with a solid-state full-field heterodyning range imaging camera. *Measurement Science and Technology*, **18(9)**, p. 2809 (2007).

Dorrington, A. A., J. P. Godbaz, M. J. Cree, A. D. Payne, and L. V. Streeter. Separating true range measurements from multi-path and scattering interference in commercial range cameras. In: *Three-Dimensional Imaging, Interaction, and Measurement, Proc. SPIE*, volume 7864, p. 786404 (2011).

Droeschel, D., D. Holz, and S. Behnke. Probabilistic phase unwrapping for time-of-flight cameras. In: *41st International Symposium on Robotics (ISR) and 6th German Conference on Robotics (ROBOTIK)*, pp. 1–7 (2010).

Duarte, M. F., M. A. Davenport, D. Takhar, J. N. Laska, T. Sun, K. E. Kelly, R. G. Baraniuk, *et al.* Single-pixel imaging via compressive sampling. *IEEE Signal Processing Magazine*, **25(2)**, p. 83 (2008).

Falie, D. and V. Buzuloiu. Distance errors correction for the time of flight (ToF) cameras. In: *IEEE International Workshop on Imaging Systems and Techniques (IST)*, pp. 123–126 (2008).

Frank, M., M. Plaue, H. Rapp, U. Koethe, B. Jähne, and F. A. Hamprecht. Theoretical and experimental error analysis of continuous-wave time-of-flight range cameras. *Optical Engineering*, **48(1)**, pp. 013602–013602 (2009).

Freedman, D., E. Krupka, Y. Smolin, I. Leichter, and M. Schmidt. SRA: fast removal of general multipath for ToF sensors. In: *Computer Vision - ECCV 2014*. Springer International Publishing (2014).

Fuchs, S. Multipath interference compensation in time-of-flight camera images. In: *20th International Conference on Pattern Recognition (ICPR)*, pp. 3583–3586 (2010).

Fuchs, S., M. Suppa, and O. Hellwich. Compensation for multipath in tof camera measurements supported by photometric calibration and environment integration. In: *Computer Vision Systems*, volume 7963, pp. 31–41. Springer Berlin Heidelberg (2013).

Gao, L., J. Liang, C. Li, and L. V. Wang. Single-shot compressed ultrafast photography at one hundred billion frames per second. *Nature*, **516(7529)**, pp. 74–77 (2014).

Gariepy, G., N. Krstajić, R. Henderson, C. Li, R. R. Thomson, G. S. Buller, B. Heshmat, R. Raskar, J. Leach, and D. Faccio. Single-photon sensitive light-in-flight imaging. *Nature communications*, **6** (2015).

Godbaz, J. *Ameliorating Systematic Errors in Full-Field AMCW Lidar*. Ph.D. thesis, University of Waikato, New Zealand (2012).

Godbaz, J., M. Cree, and A. Dorrington. Mixed pixel return separation for a full-field ranger. In: *23rd International Conference on Image and Vision Computing New Zealand (IVCNZ)*, pp. 1–6 (2008).

Godbaz, J. P., M. J. Cree, and A. A. Dorrington. Closed-form inverses for the mixed pixel/multipath interference problem in AMCW Lidar. In: *Computational Imaging X, Proc. SPIE*, volume 8296, p. 829618 (2012).

Gokturk, S., H. Yalcin, and C. Bamji. A time-of-flight depth sensor - system description, issues and solutions. In: *Computer Vision and Pattern Recognition Workshop, 2004. CVPRW '04. Conference on*, pp. 35–35 (2004).

Gold, R. Optimal binary sequences for spread spectrum multiplexing. *IEEE Trans. Inf. Theory*, **13(4)**, pp. 619–621 (1967).

Golomb, S. and R. Scholtz. Generalized barker sequences. *IEEE Transactions on Information Theory*, **11(4)**, pp. 533–537 (1965).

Golomb, S. W. and G. Gong. *Signal design for good correlation: for wireless communication, cryptography, and radar*. Cambridge University Press (2005).

Gottesman, S. R. and E. E. Fenimore. New family of binary arrays for coded aperture imaging. *Appl. Opt.*, **28(20)**, pp. 4344–4352 (1989).

Gu, J., T. Kobayashi, M. Gupta, and S. Nayar. Multiplexed illumination for scene recovery in the presence of global illumination. In: *IEEE International Conference on Computer Vision (ICCV)*, pp. 691–698 (2011).

Gulden, P., M. Vossiek, P. Heide, and R. Schwarte. Novel opportunities for optical level gauging and 3D-imaging with the photoelectronic mixing device. *Instrumentation and Measurement, IEEE Transactions on*, **51(4)**, pp. 679–684 (2002).

Gupta, M., A. Agrawal, A. Veeraraghavan, and S. G. Narasimhan. A practical approach to 3D scanning in the presence of interreflections, subsurface scattering and defocus. *International journal of computer vision*, **102(1-3)**, pp. 33–55 (2013).

Hathi, N., M. Rodrigues, I. Darwazeh, and J. O'Reilly. Analysis of the influence of walsh-hadamard code allocation strategies on the performance of multi-carrier CDMA systems in the presence of HPA non-linearities. In: *The 13th IEEE International Symposium on Personal, Indoor and Mobile Radio Communications*, volume 3, pp. 1305–1309 vol.3 (2002).

Haykin, S. *Communication systems*. John Wiley & Sons (2008).

Heide, F., W. Heidrich, M. Hullin, and G. Wetzstein. Doppler time-of-flight imaging. *ACM Transactions on Graphics (TOG)*, **34(4)**, p. 36 (2015).

Heide, F., M. B. Hullin, J. Gregson, and W. Heidrich. Low-budget transient imaging using photonic mixer devices. *ACM Transactions on Graphics (TOG)*, **32(4)**, p. 45 (2013).

Heide, F., L. Xiao, W. Heidrich, and M. B. Hullin. Diffuse mirrors: 3D reconstruction from diffuse indirect illumination using inexpensive time-of-flight sensors. In: *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 3222–3229 (2014a).

Heide, F., L. Xiao, A. Kolb, M. B. Hullin, and W. Heidrich. Imaging in scattering media using correlation image sensors and sparse convolutional coding. *Optics express*, **22(21)**, pp. 26338–26350 (2014b).

Hofbauer, M., J. Seiter, M. Davidovic, and H. Zimmermann. A processing approach for a correlating time-of-flight range sensor based on a least squares method. In: *Sensors Applications Symposium (SAS)*, pp. 355–359 (2014).

Horn, B. K. and B. G. Schunck. Determining optical flow. In: *1981 Technical Symposium East*, pp. 319–331 (1981).

Huhle, B., P. Jenke, and W. Straßer. On-the-fly scene acquisition with a handy multi-sensor system. *International Journal of Intelligent Systems Technologies and Applications*, **5(3)**, pp. 255–263 (2008).

iC Haus. *iC-HK, iC-HKB, 155 MHz Laser Switch*. iC Haus, Bodenhiem, Germany (2009). Rev E1.

Intersil. *Ultra-High Current Pin Driver*. Intersil, Milpitas, CA, USA (2007).

Jensen, H. W., S. R. Marschner, M. Levoy, and P. Hanrahan. A practical model for subsurface light transport. In: *Proceedings of the 28th annual conference on computer graphics and interactive techniques*, pp. 511–518. ACM (2001).

Jiménez, D., D. Pizarro, M. Mazo, and S. Palazuelos. Modelling and correction of multipath interference in time of flight cameras. In: *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 893–900 (2012).

Jongenelen, A. P. P. *Development of a compact, configurable, real-time range imaging system*. Ph.D. thesis, Victoria University, New Zealand (2011).

Kadambi, A., R. Whyte, A. Bhandari, L. Streeter, C. Barsi, A. Dorrington, and R. Raskar. Coded time of flight cameras: sparse deconvolution to address multipath interference and recover time profiles. *ACM Transactions on Graphics (TOG)*, **32(6)**, p. 167 (2013).

Kadambi, A., H. Zhao, B. Shi, and R. Raskar. Occluded imaging with time of flight sensors. *ACM Transactions on Graphics (TOG)*, p. To Appear (2015).

Kahlmann, T. *Range Imaging Metrology: Investigation, Calibration and Development*. Ph.D. thesis, University of Hannover, Germany (2007).

Kahlmann, T. and H. Ingensand. Calibration and development for increased accuracy of 3D range imaging cameras. *Journal of Applied Geodesy*, **2(1)**, pp. 1–11 (2008).

Kahlmann, T., F. Remondino, and H. Ingensand. Calibration for increased accuracy of the range imaging camera swissranger. *Image Engineering and Vision Metrology (IEVM)*, **36(3)**, pp. 136–141 (2006).

Kay, S. M. and S. L. Marple Jr. Spectrum analysis a modern perspective. *Proc. IEEE*, **69(11)**, pp. 1380–1419 (1981).

Kim, D. and K.-J. Yoon. High-quality depth map up-sampling robust to edge noise of range sensors. In: *19th IEEE International Conference on Image Processing (ICIP)*, pp. 553–556 (2012).

Kim, S. K., B. Kang, J. Heo, S.-W. Jung, and O. Choi. Photometric stereo-based single time-of-flight camera. *Optics letters*, **39(1)**, pp. 166–169 (2014).

Kirmani, A., A. Benedetti, and P. A. Chou. SPUMIC: simultaneous phase unwrapping and multipath interference cancellation in time-of-flight cameras using spectral methods. In: *IEEE International Conference on Multimedia and Expo (ICME)*, pp. 1–6 (2013).

Kolb, A., E. Barth, R. Koch, and R. Larsen. Time-of-flight sensors in computer graphics. In: *Proc. Eurographics*, pp. 119–134 (2009).

Lakowicz, J. R. *Principles of Fluorescence Spectroscopy*. Springer Science, 233 Spring Street, New York, NY, USA, third edition (2006).

Lange, R. *3D Time-of-Flight Distance Measurement with Custom Solid-State Image Sensors in CMOS/CCD-Technology*. Ph.D. thesis, University of Siegen, Germany (2000).

Lange, R. and P. Seitz. Solid-state time-of-flight range camera. *IEEE J. Quantum Electron.*, **37(3)**, pp. 390–397 (2001).

Lanman, D. and G. Taubin. Build your own 3D scanner: 3D photograhy for beginners. In: *SIGGRAPH '09: ACM SIGGRAPH 2009 courses*, pp. 1–87. ACM, New York, NY, USA (2009).

Larkins, R., M. Cree, A. Dorrington, and J. Godbaz. Surface projection for mixed pixel restoration. In: *24th International Conference on Image and Vision Computing New Zealand (IVCNZ)*, pp. 431–436 (2009).

Le, A. V., S.-W. Jung, and C. S. Won. Directional joint bilateral filter for depth images. *Sensors*, **14(7)**, pp. 11362–11378 (2014).

Lee, S., B. Kang, J. D. Kim, and C. Y. Kim. Motion blur-free time-of-flight range sensor. In: *Sensors, Cameras, and Systems for Industrial and Scientific Applications XIII, Proc. SPIE*, volume 8298, p. 82980U (2012).

Lenzen, F., H. Schäfer, and C. Garbe. Denoising time-of-flight data with adaptive total variation. In: *Advances in Visual Computing*, volume 6938, pp. 337–346. Springer Berlin Heidelberg (2011).

Lichti, D. D., J. C. Chow, E. Mitishita, J. A. S. Centeno, F. M. M. d. Silva, R. A. Barrios, and I. Contreras. New models for scattering bias compensation in time-of-flight range camera self-calibration. *Journal of Surveying Engineering*, **140(2)**, p. 04014003 (2014).

Lin, J., Y. Liu, M. B. Hullin, and Q. Dai. Fourier analysis on transient imaging by multifrequency time-of-flight camera. In: *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, p. to appear (2014).

Lindner, M. and A. Kolb. Lateral and depth calibration of pmd-distance sensors. In: *Advances in Visual Computing*, volume 4292, pp. 524–533. Springer Berlin Heidelberg (2006).

Lindner, M. and A. Kolb. Compensation of motion artifacts for time-of-flight cameras. In: *Dynamic 3D Imaging*, volume 5742, pp. 16–27. Springer Berlin Heidelberg (2009).

Lindner, M., M. Lambers, and A. Kolb. Sub-pixel data fusion and edge-enhanced distance refinement for 2D/3D images. *International Journal of Intelligent Systems Technologies and Applications*, **5(3)**, pp. 344–354 (2008).

Lindner, M., I. Schiller, A. Kolb, and R. Koch. Time-of-flight sensor calibration for accurate range sensing. *Computer Vision and Image Understanding*, **114(12)**, pp. 1318–1328 (2010).

Marple, S. L. *Digital Spectral Analysis*. Englewood Cliffs, NJ: Prentice-Hall (1987).

McClure, S. H., M. J. Cree, A. A. Dorrington, and A. D. Payne. Resolving depth-measurement ambiguity with commercially available range imaging cameras. In: *Image Processing: Machine Vision Applications III, Proc. SPIE*, volume 7538, p. 75380K (2010).

Mei, J., A. Kirmani, A. Colaco, and V. K. Goyal. Phase unwrapping and denoising for time-of-flight imaging using generalized approximate message passing. In: *20th IEEE International Conference on Image Processing (ICIP)*, pp. 364–368 (2013).

Mesa Imaging. *SR4000 User Manual*. MESA Imaging, Zurich, Switzerland (2010). Rev 1.4.

Mure-Dubois, J. and H. Hügli. Optimized scattering compensation for time-of-flight camera. In: *Two and Three-Dimensional Methods for Inspection and Metrology V, Proc. SPIE*, volume 6762, p. 67620H (2007).

Muthukrishnan, S. *Data streams: Algorithms and applications*. Now Publishers Inc (2005).

Muttayane, A. *Towards Colour Imaging with the Image Ranger*. Master's thesis, The University of Waikato, New Zealand (2006).

Naik, N., A. Kadambi, C. Rhemann, S. Izadi, R. Raskar, and S. B. Kang. A light transport model for mitigating multipath interference in TOF sensors. *arXiv preprint arXiv:1501.04878* (2015). To be published.

Nayar, S. K., G. Krishnan, M. D. Grossberg, and R. Raskar. Fast separation of direct and global components of a scene using high frequency illumination. *ACM Trans. Graph.*, **25(3)**, pp. 935–944 (2006).

Niclass, C., 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 J. Solid-State Circuits*, **40(9)**, pp. 1847–1854 (2005).

Opnext. *HL6545MG Visible High Power Laser Diode for Recordable-DVD*. opnext, San Jose, CA, USA (2008).

O'Toole, M., F. Heide, L. Xiao, M. B. Hullin, W. Heidrich, and K. N. Kutulakos. Temporal frequency probing for 5D transient analysis of global light transport. *ACM Trans. Graph. (Proc. SIGGRAPH)*, **33(4)** (2014a).

O'Toole, M., J. Mather, and K. N. Kutulakos. 3D shape and indirect appearance by structured light transport. In: *Computer Vision and Pattern Recognition (CVPR), 2014 IEEE Conference on*, pp. 3246–3253. IEEE (2014b).

Pati, Y. C., R. Rezaiifar, and P. Krishnaprasad. Orthogonal matching pursuit: Recursive function approximation with applications to wavelet decomposition. In: *1993 Conference Record of The Twenty-Seventh Asilomar Conference on Signals, Systems and Computers*, pp. 40–44 (1993).

Payne, A. *Development of a Full-Field Time-of-Flight Range Imaging System*. Ph.D. thesis, University of Waikato, New Zealand (2008).

Payne, A. D., A. A. Dorrington, and M. J. Cree. Illumination waveform optimization for time-of-flight range imaging cameras. In: *Videometrics, Range Imaging, and Applications XI, Proc. SPIE*, volume 8085, p. 80850D (2011).

Payne, A. D., A. A. Dorrington, M. J. Cree, and D. A. Carnegie. Improved linearity using harmonic error rejection in a full-field range imaging system. In: *Three-Dimensional Image Capture and Applications, Proc. SPIE*, volume 6805, p. 68050D (2008).

Payne, A. D., A. A. Dorrington, M. J. Cree, and D. A. Carnegie. Characterization of modulated time-of-flight range image sensors. In: *Three-Dimensional Imaging Metrology, Proc. SPIE*, volume 7239, pp. 723904–723904 (2009).

Peters, C., J. Klein, M. B. Hullin, and R. Klein. Solving trigonometric moment problems for fast transient imaging. *ACM Transactions on Graphics (Proc. SIGGRAPH Asia)*, **34**, p. To Appear (2015).

Raskar, R., A. Agrawal, and J. Tumblin. Coded exposure photography: motion deblurring using fluttered shutter. *ACM Transactions on Graphics (TOG)*, **25(3)**, pp. 795–804 (2006).

Rauhut, H. Circulant and toeplitz matrices in compressed sensing. In: *Signal Processing with Adaptive Sparse Structured Representations* (2009).

Reynolds, M., J. Dobos, L. Peel, T. Weyrich, and G. Brostow. Capturing time-of-flight data with confidence. In: *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 945–952 (2011).

Saha, S. and S. Kay. Maximum likelihood parameter estimation of superimposed chirps using Monte Carlo importance sampling. *Signal Processing, IEEE Transactions on*, **50(2)**, pp. 224–230 (2002).

Schoukens, J. and J. Renneboog. Modeling the noise influence on the Fourier coefficients after a discrete Fourier transform. *IEEE Trans. Instrum. Meas.*, **1001(3)**, pp. 278–286 (1986).

Schuon, S., C. Theobalt, J. Davis, and S. Thrun. High-quality scanning using time-of-flight depth superresolution. In: *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*, pp. 1–7 (2008).

Seiter, J., M. Hofbauer, M. Davidovic, and H. Zimmermann. Investigation of the distance error induced by cycle-to-cycle jitter in a correlating

time-of-flight distance measurement system. *Optical Engineering*, **53(7)**, pp. 073104–073104 (2014).

Seitz, S., Y. Matsushita, and K. Kutulakos. A theory of inverse light transport. In: *10th IEEE International Conference on Computer Vision (ICCV)*, volume 2, pp. 1440–1447 Vol. 2 (2005).

Streeter, L., M. J. Cree, and A. A. Dorrington. A strategy for the correction of effects of jitter in AMCW Lidar images. In: *28th International Conference on Image and Vision Computing New Zealand (IVCNZ), 2013*, pp. 500–505 (2013).

Streeter, L. and A. A. Dorrington. Coded exposure correction of transverse motion in full-field range imaging. *Optical Engineering*, **53(10)**, p. 102109 (2014).

Tadmor, E., I. Bakish, S. Felzenshtein, E. Larry, G. Yahav, and D. Cohen. A fast global shutter image sensor based on the VOD mechanism. In: *2014 IEEE Sensors*, pp. 618–621 (2014).

Tajahuerce, E., V. Durán, P. Clemente, E. Irles, F. Soldevila, P. Andrés, and J. Lancis. Image transmission through dynamic scattering media by single-pixel photodetection. *Optics express*, **22(14)**, pp. 16945–16955 (2014).

Terasic Technologies. *DE2-115 User Manual*. Terasic Technologies, Hsinchu City, Taiwan (2013). Rev 1.

Texas Instruments. *DLP Digital Controller for the DLP3000 DMD*. Texas Instruments, Dallas, TX, USA (2013a).

Texas Instruments. *DLP LightCrafter Evaluation Module (EVM) User's Guide*. Texas Instruments, Dallas, TX, USA (2013b).

Texas Instruments. *Introduction to the Time-of-Flight (ToF) System Design*. Texas Instruments, Dallas, TX, USA (2014).

Thorlabs. *L850P010 Laser Diode*. Thorlabs, Newton, NJ, USA (2010). Rev C.

Tibshirani, R. Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society. Series B (Methodological)*, pp. 267–288 (1996).

Tsai, R. A versatile camera calibration technique for high-accuracy 3D machine vision metrology using off-the-shelf tv cameras and lenses. *IEEE Trans. Robot. Autom.*, **3(4)**, pp. 323–344 (1987).

US Lasers. *D405-120 405nm 120mW laser diode*. US Lasers, Baldwin Park, CA, USA (2013).

Van Nieuwenhove, D., W. Van der Tempel, R. Grootjans, and M. Kuijk. A CAPD based time-of-flight ranging pixel with wide dynamic range. In: *Optical and Digital Image Processing, Proc. SPIE*, volume 7000, p. 70000N (2008).

Velten, A., T. Willwacher, O. Gupta, A. Veeraraghavan, M. G. Bawendi, and R. Raskar. Recovering three-dimensional shape around a corner using ultrafast time-of-flight imaging. *Nature Communications*, **3**, p. 745 (2012).

Velten, A., D. Wu, A. Jarabo, B. Masia, C. Barsi, C. Joshi, E. Lawson, M. Bawendi, D. Gutierrez, and R. Raskar. Femto-photography: Capturing and visualizing the propagation of light. *ACM Transactions on Graphics (TOG)*, **32(4)**, p. 44 (2013).

Wang, Y. Residue-to-binary converters based on new chinese remainder theorems. *IEEE Trans. Circuits Syst. I*, **47(3)**, pp. 197–205 (2000).

Wasza, J., S. Bauer, and J. Hornegger. Real-time preprocessing for dense 3-D range imaging on the GPU: Defect interpolation, bilateral temporal averaging and guided filtering. In: *IEEE International Conference on Computer Vision Workshops (ICCV Workshops)*, pp. 1221–1227 (2011).

Whyte, R., A. Bhandari, L. Streeter, M. J. Cree, and A. A. Dorrington. Time frequency duality of time-of-flight range cameras for resolving multi-path interference. In: *29th International Conference on Image and Vision Computing New Zealand (IVCNZ)*, pp. 247–252 (2014a). ISBN 978-1-4503-3184-5.

Whyte, R., L. Streeter, M. Cree, and A. Dorrington. Review of methods for resolving multi-path interference in time-of-flight range cameras. In: *2014 IEEE Sensors*, pp. 629–632 (2014b).

Whyte, R., L. Streeter, M. Cree, and A. Dorrington. Resolving multiple propagation paths in time of flight range cameras using direct and global separation methods. *Optical Engineering*, **54(11)**, p. 113109 (2015a).

Whyte, R., L. Streeter, M. J. Cree, and A. A. Dorrington. Application of lidar techniques to time-of-flight range imaging. *Appl. Opt.*, **54(33)**, pp. 9654–9664 (2015b).

Whyte, R. Z., A. D. Payne, A. A. Dorrington, and M. J. Cree. Multiple range imaging camera operation with minimal performance impact. In: *Machine Vision Applications III, Proc. SPIE*, volume 7538, p. 75380I (2010).

Wu, D., A. Velten, M. O'Toole, B. Masia, A. Agrawal, Q. Dai, and R. Raskar. Decomposing global light transport using time of flight imaging. *International Journal of Computer Vision*, **107(2)**, pp. 123–138 (2014).

Xu, Z., R. Schwarte, H.-G. Heinol, B. Buxbaum, and T. Ringbeck. Smart pixel : photonic mixer device (PMD) ; new system concept of a 3D-imaging camera-on-a-chip. In: *Proceedings : M2VIP '98; Nanjing, China, 10th - 12th September 1998*, edited by E. H. M. Cheung, pp. 259 – 264 (1998).

Zhang, Z. A flexible new technique for camera calibration. *IEEE Trans. Pattern Anal. Mach. Intell.*, **22(11)**, pp. 1330–1334 (2000).