Characterisation of the Victoria University Range Imaging System

Ben Drayton
Victoria University of Wellington
ben.drayton@gmail.com

Dale A. Carnegie
Victoria University of Wellington

Adrian A. Dorrington
University of Waikato

Abstract
Indirect time of flight cameras are becoming commonplace for real time full field of view range imaging. This paper will characterise the response of the Victoria University Range Imaging System, an indirect time of flight camera. This characterisation focused on the precision, accuracy and a number of phenomena that influence these characteristics. Non-ideal properties of the sensor are explored, in particular non-linearity with respect to intensity and spatial non-uniformity. Particular attention is paid to the effect of the modulation frequency on the sensor and the effect of harmonics in the correlation waveform. The modulation frequency is not normally adjustable on commercial cameras and is therefore less well researched. Measurements are compared to theory and investigations are performed where deviations from the theory occur.

1 Introduction
Indirect time of flight cameras are increasingly being used for real time full field of view range imaging and have significant promise for applications in the field of mobile robotics. Instead of measuring the time of flight for an emitted signal directly, these cameras encode the time of flight into a phase shift between two modulated signals. A light source is modulated at high frequency, generally between 10 and 100 MHz, and an intensity based image sensor is modulated at the same frequency. Previously the sensor modulation has been provided by an image intensifier acting as a shutter on a standard CCD camera [Cree et al., 2006], however these are bulky, expensive and require high voltage signals to operate. Modern indirect time of flight cameras implement custom CMOS based sensors allowing low voltage electronic modulation. For the sensor used in the Victoria University Range Imaging System, each pixel contains two readout gates and two modulation gates. The modulation signal acts to guide charge carriers stimulated by incident light to one or the other readout gate. After integration over a large number of modulation cycles the output of the two readout gates are subtracted, giving the measured intensity value.

\[ d = \frac{ct}{2} \]

Indirect time of flight cameras operate by encoding the time taken for the emitted signal to return to the camera into a phase shift between two modulated signals. A light source is modulated at high frequency, generally between 10 and 100 MHz, and an intensity based image sensor is modulated at the same frequency. Previously the sensor modulation has been provided by an image intensifier acting as a shutter on a standard CCD camera [Cree et al., 2006], however these are bulky, expensive and require high voltage signals to operate. Modern indirect time of flight cameras implement custom CMOS based sensors allowing low voltage electronic modulation. For the sensor used in the Victoria University Range Imaging System, each pixel contains two readout gates and two modulation gates. The modulation signal acts to guide charge carriers stimulated by incident light to one or the other readout gate. After integration over a large number of modulation cycles the output of the two readout gates are subtracted, giving the measured intensity value.
The time taken for the illumination signal to return from the object being imaged introduces a phase shift between the modulation of the illumination signal and the modulation of the sensor. The time taken is related to the phase by the equation

$$ t = \frac{\varphi}{2\pi f_{\text{mod}}} $$

(2)

where \( \varphi \) is the phase between the two signals and \( f_{\text{mod}} \) is the modulation frequency. Substituting (2) into (1), the distance is therefore related to the phase by the equation

$$ d = \frac{c \varphi}{4\pi f_{\text{mod}}} . $$

(3)

The intensity observed by the sensor is integrated over many periods of the modulation frequency to provide an intensity value related to the phase between the two modulation signals. Assuming that the modulation signals are sinusoids, the intensity observed by the sensor is related to the phase shift between the modulation signals by the equation [Jongenelen, 2010]

$$ I = A \cos(\varphi) + B, $$

(4)

where \( A \) is a gain factor including the sensitivity of the sensor, the amplitude of the modulated light, the reflectivity of the object being measured and the inverse square decrease in illumination with distance caused by the spreading of the electromagnetic waves from the illumination source. \( B \) is an offset value caused by background illumination, DC offset in the ADC and, in some sensor architectures, asymmetry in the two pixel readout gates.

Due to the dependence of the coefficients \( A \) and \( B \) on a number of parameters exterior to the camera, a single intensity measurement is not sufficient to calculate the phase shift. Instead \( N \) measurements are taken with a phase step \( \delta \) introduced between the illumination and sensor modulation signals for each measurement. The intensity \( I_n \) for frame \( n \) \( (n = 1 \ldots N) \) is therefore

$$ I_n = A \cos(\varphi(n-1) \delta - \varphi) + B. $$

(5)

The phase steps are normally at regular intervals of \( N/2\pi \) radians. The sign of the phase has been changed to match convention.

This set of \( N \) intensity frames forms what is commonly referred to as the correlation waveform, as it represents the correlation between the modulation signals of the illumination source and the sensor. The phase of the correlation waveform is the phase of the first intensity measurement \( (n = 1) \) and therefore the phase representing the distance to the object. Using a Discrete Fourier Transform the phase can be determined as

$$ \varphi = \tan^{-1} \left( \frac{\sum_{n=1}^{N} I_n \sin(2\pi(n-1)\delta)}{\sum_{n=1}^{N} I_n \cos(2\pi(n-1)\delta)} \right). $$

(6)

It is common to use four frames per phase measurement as this simplifies (6) to

$$ \varphi = \tan^{-1} \left( \frac{I_2 - I_4}{I_1 - I_3} \right). $$

(7)

The amplitude of the correlation waveform is

$$ A = \frac{1}{2} \sqrt{(I_1 - I_3)^2 + (I_2 - I_4)^2}. $$

(8)

and the offset is

$$ B = \frac{I_1 + I_2 + I_3 + I_4}{4}. $$

(9)

Due to the cyclic nature of the phase there is a maximum distance at which unambiguous measurements can be recorded. Beyond this distance the phase will wrap around. Equation (3) can therefore more accurately be represented as

$$ d = \frac{c}{2f_{\text{mod}}} \left( \frac{\varphi}{2\pi} + k \right) $$

(10)

where \( k \) is an integer. Increasing the modulation frequency decreases the maximum unambiguous measurement distance of the camera.

3 Camera Hardware

The camera characterised in this paper is the Victoria University Range Imaging System, a custom built camera designed for research into mobile robotic applications. It uses a PMD19K-2 electronically modulated image sensor (PMDTechnologies GmbH, Siegen, Germany) with a resolution of 160 \( \times \) 120 pixels. A Cyclone III EP3C120 F780C7 FPGA (Altera, San Jose, CA, USA) is used to provide processing of the intensity images and control of the modulation signals. It has four Phase Locked Loops (PLLs) that are used to provide the phase stepped modulation signals required for indirect time of flight range imaging and approximately 3.8 mega-bits of internal RAM for storing data from the sensor for processing and processed data for outputting to either a VGA monitor or a computer for long term storage. A photograph of the camera is shown in Figure 1.

![Figure 1 Photograph of the Victoria University Range Imaging System [McClymont et al., 2011]](image-url)
range measurement entirely.

To measure the precision of the Victoria University Range Imaging System, a flat, stationary object was imaged over 100 measurements for a number of different measurement times. The object used for these measurements was a white piece of card. The standard deviation of the measurements was calculated for each pixel and averaged over a 60 × 10 region of pixels. The results are shown in Figure 2. As expected, an inverse relationship is observed until the sensor reaches saturation. When saturated the four intensity frames have essentially the same value and therefore the output is largely random, leading to a very high standard deviation.

Both the amplitude $A$ and the offset $B$ are expected to increase linearly with the measurement time [Jongenelen, 2010]. Therefore, since the precision is proportional to $\sqrt{B}/A$, the precision of the phase measurements should have an inverse square root relationship with measurement time. However, improving the precision by increasing the measurement time has a negative impact on the frame rate of the camera. It also has a limit imposed by the sensor architecture. If the measurement time is too long the returning light will saturate the sensor, destroying the

$$
\sigma_{\phi} = \frac{c_{d}}{2\pi f_{mod}}
$$

meaning that the distance precision can be improved by increasing the modulation frequency. However, as discussed in Section 2 this has the trade off of decreasing the unambiguous measurement distance.

Both the amplitude $A$ and the offset $B$ are expected to increase linearly with the measurement time [Jongenelen, 2010]. Therefore, since the precision is proportional to $\sqrt{B}/A$, the precision of the phase measurements should have an inverse square root relationship with measurement time. However, improving the precision by increasing the measurement time has a negative impact on the frame rate of the camera. It also has a limit imposed by the sensor architecture. If the measurement time is too long the returning light will saturate the sensor, destroying the

$$
\sigma_{d} = \frac{c_{d}}{2\pi f_{mod}}
$$

The inverse square root relationship predicted between the precision and the measurement time is reliant on both the amplitude and the offset of the correlation waveform being proportional to the measurement time. To investigate if this is actually the case for our sensor, the intensity of the four intensity frames was recorded for different measurement times using the same setup as for measuring the precision. The results are shown in Figure 4. While the intensity is expected to increase linearly with measurement time the result from our sensor is highly non-linear. It should be noted that as described in Section 2, the intensity value is the subtraction of two readout gates. The sign of the intensity values therefore only indicates which of the gates collected a larger number of charge carriers.

12 V gel battery.

- The FPGA Board – This board provides processing power and control for the camera. The FPGA also produces the phase stepped modulation signals required for indirect time of flight imaging using onboard PLLs. Additional external memory is included on this board to future proof the system against advances in sensor technology leading to higher resolution sensors.

- The Image Capture Board – This board provides high current modulation drivers to modulate the sensor, which is essentially a large capacitive load. It also contains an ADC to convert the analogue video output stream from the sensor into digital frames that can be processed using the FPGA.

- The Illumination Board – This board provides the modulated illumination signal required for indirect time of flight range imaging. A bank of 16 red laser diodes arranged in a concentric circle around the lens is used to provide illumination, with a total optical power of 800 mW.

The stacked structure of the system also means it is simple to replace both the sensor being used and the lens assembly. The lens gap is adjustable to allow a large variety of lenses to be used.

4 Measurement Precision

According to theory, the precision of indirect time of flight measurements is described by the equation [Buttgen et al., 2005]

$$
\sigma_{\phi} = \frac{\sqrt{B}}{2\pi c_{d} b_{sig}}
$$

where $\sigma_{\phi}$ is the standard deviation of the phase measurements, $c_{d}$ is the demodulation contrast and $b_{sig}$ is the offset of the correlation waveform due to the illumination signal. There is also some offset due to ambient light conditions $b_{amb}$ such that $B = b_{sig} + b_{amb}$. The demodulation contrast is equal to $A / b_{sig}$, and therefore (11) can be rewritten as

$$
\sigma_{\phi} = \frac{\sqrt{B}}{2A}
$$

The precision of the distance measurements is related to the precision of the phase measurements by the equation

where $\sigma_{d}$ is the standard deviation of the distance measurements. For shorter measurement times, the precision of the distance measurements is related to the standard deviation of the phase measurements by the equation

$$
\sigma_{d} = \frac{c_{d}}{4\pi f_{mod}}
$$

The inverse square root relationship expected.

$$
\sigma_{d} = \frac{c_{d}}{4\pi f_{mod}}
$$

measuring the exact relationship of the data in Figure 2 a log-log plot, excluding the data from saturation, is shown in Figure 3. A linear fit of this data gives a slope of -1.2, significantly different from the inverse square root relationship expected.

Figure 2 Standard deviation of phase measurements versus measurement time

Figure 3 Log-log plot of standard deviation of phase measurements versus measurement time excluding saturation

The inverse square root relationship predicted between the precision and the measurement time is reliant on both the amplitude and the offset of the correlation waveform being proportional to the measurement time. To investigate if this is actually the case for our sensor, the intensity of the four intensity frames was recorded for different measurement times using the same setup as for measuring the precision. The results are shown in Figure 4. While the intensity is expected to increase linearly with measurement time the result from our sensor is highly non-linear. It should be noted that as described in Section 2, the intensity value is the subtraction of two readout gates. The sign of the intensity values therefore only indicates which of the gates collected a larger number of charge carriers.
Using the data from Figure 4, the amplitude and offset of the correlation waveform can be measured and is shown in Figure 5. While the amplitude appears to have an approximately linear relationship with the measurement time, until saturation is reached, the offset shows non-linear behaviour significantly before saturation. It should be noted that as the measured intensity values are digital representations of the measured intensity that have been scaled, the amplitude and offset are in arbitrary units.

From (11) we know that the precision should be proportional to $\sqrt{B/A}$. As $\sqrt{B/A}$ is no longer expected to have an inverse square root relationship with measurement time, this proportionality could still be experimentally verified. Using the amplitude and the offset, the ratio $\sqrt{B/A}$ was calculated and is shown in Figure 6 along with the standard deviation, scaled appropriately. The measured standard deviation divided by the ratio $\sqrt{B/A}$ is shown in Figure 7. Since there is some scaling involved in measuring the intensity values, the proportionality constant will not necessarily be $1/\sqrt{Z}$, instead we are simply looking for proportionality. Despite the non-linearity of our sensor, the ratio $\sqrt{B/A}$ is still proportional to the precision of the phase measurements.

As this could be a problem only with our particular model of sensor, these measurements were repeated using a SwissRanger 4000 commercial indirect time of flight camera (Mesa Imaging, Zurich, Switzerland). The results are shown in Figure 8. For these measurements the camera was operated in “raw” mode, which provides uncalibrated intensity values. The intensity values measured using the SwissRanger 4000 camera appear to be linear. To confirm this, the amplitude and offset were calculated from these data and are shown in Figure 9. The amplitude has a linear increase with measurement time however there is still significant non-linearity observed in the offset for short measurement times. The cause for this non-linearity is not currently known.
Figure 9 Amplitude (top) and offset (bottom) versus measurement time for a static object using SR4000 camera

It is also expected that there will be a spatial variation in the precision of the phase measurements [Kahlmann et al., 2006]. Due to the geometry of the illumination source and the optics of the lens the highest precision is expected in the centre of the field of view with a curved surface towards the worst precision in the corners of the sensor. The actual distance to a flat surface will also be slightly greater at the edges than at the centre, however this has been shown to be negligible [Kahlmann and Ingensand, 2005]. Measurements of a flat surface, shown in Figure 10, confirm this prediction.

Figure 10 Spatial variation in phase measurement precision

Both the amplitude and the offset of the correlation waveform can be graphed and are shown in Figure 11. As expected, near the centre of the field of view the amplitude is much higher, causing an improvement in the precision of the phase measurements. The offset is reasonably constant across the sensor.

Figure 11 Spatial variation in Amplitude (top) and offset (bottom) of the correlation waveform

5 Measurement Accuracy

A number of systematic errors are known to affect the accuracy of indirect time of flight measurements. One significant effect is due to the presence of harmonics in the correlation waveform. In Section 2 the phase detection algorithm was derived assuming that the modulation of both the sensor and the illumination source was perfectly sinusoidal. In reality these modulation signals will contain harmonics as they are normally generated using square waves, due to the ease of doing so digitally. The transfer function of the modulation drivers, the sensor and the illumination source will also influence the harmonic content of the correlation signal.

With low numbers of correlation waveform samples, such as four typically used, harmonics can violate the Nyquist sampling criteria and be aliased onto the fundamental, resulting in a sinusoidal error in the phase measurement. Using square wave modulation it is expected that the third harmonic will be the strongest as the correlation waveform will be triangular. With four samples, the third harmonic is expected to cause a sinusoidal error with four cycles within the unambiguous measurement distance of the camera. The non-linearity of our camera, shown in Figure 12, was measured by recording a stationary object and introducing an artificial phase offset between the two modulation signals. A modulation frequency of 30 MHz was used for these measurements. As expected, a four cycle error is observed within the unambiguous measurement distance.
There have been a number of attempts to calibrate this error using b-splines [Fuchs and Hirzinger, 2008; Lindner and Kolb, 2006] and look up tables [Kahlmann et al., 2006]. However, these methods are dependent on the harmonic content of the signal remaining constant. In reality there are a number of factors that can influence the harmonic content of the correlation waveform, and some of these factors may change with time.

The most obvious influence is the modulation frequency. The limited bandwidth of components used in the camera means that increasing the modulation frequency should decrease the influence of the harmonics as they will become more attenuated. It can be advantageous to have a dynamic modulation frequency as the modulation frequency affects both the distance measurement precision and the maximum unambiguous measurement distance of the sensor.

To quantitatively measure this effect, the relative amplitude of the third harmonic to the fundamental was compared for a number of different modulation frequencies. The third harmonic was chosen as it is expected to be the strongest of the harmonics using square wave modulation. The results are shown in Figure 13. Up to a modulation frequency of 30 MHz the relative amplitude of the third harmonic decreases with increasing frequency as expected. Above 30 MHz the relative amplitude begins to increase again. This is explained by the fact that at these higher frequencies the amplitude of the fundamental is starting to decrease due to bandwidth limitations, as shown in Figure 14.

There is also the potential for the harmonic content of the correlation waveform to have spatial variation across the sensor. As the modulation signals generally enter on one side and propagate across the sensor, the harmonic content of the correlation waveform can change across the sensor, with each pixel acting essentially as a capacitive load. There will also be some variation between each individual pixels response due to the manufacturing process. To measure this phenomenon the harmonic content of an area of pixels was measured and both horizontal and vertical cross-sectional averages are shown in Figure 15. The harmonic content of the correlation waveform was measured by setting \( N \) to 64, sufficiently high that no aliasing should occur of the harmonics, and the correlation waveform was sampled over several phase measurements. A Fourier transform was then performed to acquire the harmonic content of the correlation waveform.

For our sensor the modulation signals enter the sensor along the top edge. There is no noticeable relationship between the horizontal pixels and the relative amplitude of the third harmonic, however there is a significant amount of random variation between the pixels. For the vertical cross-section there is a clear relationship between the relative amplitude of the third harmonic and the position across the sensor.
As well as harmonic content, the propagation of the modulation signal across the sensor introduces a phase shift across the sensor for objects that are at the same actual distance. A surface plot of the data recorded for a flat object is shown in Figure 16. Vertically, as the signal propagates through the sensor a smooth phase shift is introduced. Horizontally the sensor is separated into four modulation columns to reduce the capacitive load on each modulation input, this introduces some differences in response between each column due to internal routing in the sensor [Payne et al., 2009].

The spatial phase error is likely to have a dependence on the modulation frequency. The measurements were repeated using a number of modulation frequencies and the results are shown in Figure 17. This confirms that changing the modulation frequency does have an effect on this error. As the modulation frequency increases the phase shift over the sensor increases.

It is common to combine a single Fixed Pattern Noise calibration and a distance calibration to attempt to calibrate for the various spatial and distance based errors described in this section [Kahlmann et al., 2006], however some systematic error is likely to remain using this approach. Further work is required in this area to find a method of efficiently performing a multi-parameter calibration of indirect time of flight cameras.

6 Conclusions

In this paper the precision and accuracy of the Victoria University Range Imaging System was characterised in terms of a number of parameters, with particular attention to variation in the modulation frequency. It was found that while the intensity recorded by the sensor used in this camera was non-linear the relationship between the measurement precision and the ratio $\sqrt{B/A}$ was maintained. Non-linearity of the PMD19K-2 sensor was found to mainly be expressed in non-linearity of the offset of the correlation waveform.

The harmonic content of the correlation waveform, which is responsible for a linearity error with distance, was shown to have both spatial variation due to the modulation signals propagating from one side of the sensor and dependence on the modulation frequency. Spatial phase errors due to the propagation of were also found to have a dependence on the modulation frequency.

Acknowledgments

The authors would like to thank the New Zealand Foundation for Research Science & Technology (Contract-20759) for funding this work.

References


