R. Z. Whyte, A. D. Payne, A. A. Dorrington, and M. J. Cree, "Multiple range imaging camera operation with minimal performance impact," Proc. SPIE-IS&T Electronic Imaging, edited by D. Fofi, K. S. Niel, SPIE vol. 7538, pp. 75380I, (2010).

http://dx.doi.org/10.1117/12.838786

Multiple range imaging camera operation with minimal performance impact

Refael Z. Whyte, Andrew D. Payne, Adrian A. Dorrington* and Michael J. Cree

Department of Engineering, The University of Waikato, Private Bag 3105,

Hamilton 3240, New Zealand

ABSTRACT

Time-of-flight range imaging cameras operate by illuminating a scene with amplitude modulated light and measuring the phase shift of the modulation envelope between the emitted and reflected light. Object distance can then be calculated from this phase measurement. This approach does not work in multiple camera environments as the measured phase is corrupted by the illumination from other cameras. To minimize inaccuracies in multiple camera environments, replacing the traditional cyclic modulation with pseudo-noise amplitude modulation has been previously demonstrated. However, this technique effectively reduced the modulation frequency, therefore decreasing the distance measurement precision (which has a proportional relationship with the modulation frequency). A new modulation scheme using maximum length pseudo-random sequences binary phase encoded onto the existing cyclic amplitude modulation, is presented. The effective modulation frequency therefore remains unchanged, providing range measurements with high precision. The effectiveness of the new modulation scheme was verified using a custom time-of-flight camera based on the PMD19-K2 range imaging sensor. The new pseudo-noise modulation has no significant performance decrease in a single camera environment. In a two camera environment, the precision is only reduced by the increased photon shot noise from the second illumination source.

Keywords: Time of flight, Range imaging, Pseudo-Noise (PN) coding

1. INTRODUCTION

Time of Flight (TOF) range imaging cameras measure the distance to each pixel in the field of view. This is achieved by illuminating the scene with intensity modulated light and measuring the phase shift between the outgoing and incoming amplitude modulation envelope. The distance is then calculated from this phase shift. The problem with current camera designs is that when multiple cameras illuminate the same scene the measured phase is the sum of both camera's phases, corrupting the distance information. Ideally multiple cameras should be capable of operating in a common scene without knowledge or synchronization with each other. Currently multiple cameras can operate with different amplitude modulation frequencies.[1] Due to limited frequencies available, knowledge of all TOF cameras operating close by is needed, making this technique non-ideal.

Figure 1 illustrates an image taken with both a TOF camera and a standard digital camera. The pixel intensity in figure 1b contains the depth information, where darker pixels are closer to the camera. The shape and distances between the coffee cup, garden gnome, drink bottle and Buddha statue can be measured. In comparison the digital photo contains no direct depth information. The white pixels in the range image on the coffee cup and the gnome's beard are caused by too much reflected light returning and saturating the image sensor for this particular capture. TOF cameras also suffer from other measurement errors and inaccuracies, a detailed description of which can be found in May et. al. and Guomundsson et. al.[2,3]

*adrian@waikato.ac.nz; phone +64-7-858-5062; fax +64-7-838-4835

Image Processing: Machine Vision Applications III, edited by David Fofi, Kurt S. Niel,
Proc. of SPIE-IS&T Electronic Imaging, SPIE Vol. 7538, 75380I · © 2010 SPIE-IS&T
CCC code: 0277-786X/10/$18 · doi: 10.1117/12.838271

(a) (b)

Figure 1. Images of a test scene with different objects placed at various distances apart. Image (a) is a digital photo, (b) is a range image taken with a time-of-flight camera. Pixel intensity in (b) contains distance information, darker pixels are closer to the camera, as indicated by the scale on the right.

1.1 Time-of-flight principles

TOF cameras typically use an indirect measurement of the time of flight of light to determine the distance to objects in a scene. The scene is illuminated by a light source that is amplitude modulated. The distance travelled by the modulated light to the object and back produces a phase shift in the modulation envelope. This phase is measured with an image sensor that samples the incoming light at the same frequency as the intensity modulation (homodyne demodulation). The demodulated output $m(\tau)$ is the cross-correlation between the optical signal $s(t)$ and shutter signal $g(t)$,[4] i.e,

$$m(\tau) = s(t) \otimes g(t) = \lim_{T \to \infty} \frac{1}{T} \int_{-\frac{T}{2}}^{+\frac{T}{2}} s(t) \cdot g(t+\tau) dt \tag{1}$$

Square wave amplitude modulation is used in most range imaging cameras due to the simplicity of implementation in a digital circuit. The square wave functions $s(t)$ and $g(t)$ are described by equations 2 and 3 respectively. The optical signal $s(t)$ becomes phased shifted by φ radians due to propagation delay of the return journey.

$$s(t) = [1 + \mathrm{sgn}(\cos(\omega t - \varphi))] \tag{2}$$

$$g(t) = (1 + \mathrm{sgn}(\cos(\omega t)))/2 \tag{3}$$

The returning phase φ can be calculated by taking four samples (A_0 to A_3) of the image sensor output $m(\tau)$ where the shutter signal $g(t)$ is phase shifted by $\pi/2$ radians. Using these four samples the returning light amplitude A, DC offset B, and phase φ are calculated using equations 4 to 6,[5] which are based on the assumption that sine wave modulation is used. When using square wave modulation, odd harmonics are aliased during sampling making these equations only an approximation. Calibration is typically used to compensate for linearity errors introduced by these harmonics.[6] A graphical representation of these three variables is shown in figure 2.

$$A = \frac{\sqrt{(A_0 - A_2)^2 + (A_1 - A_3)^2}}{2} \tag{4}$$

$$B = \frac{A_0 + A_1 + A_2 + A_3}{4} \tag{5}$$

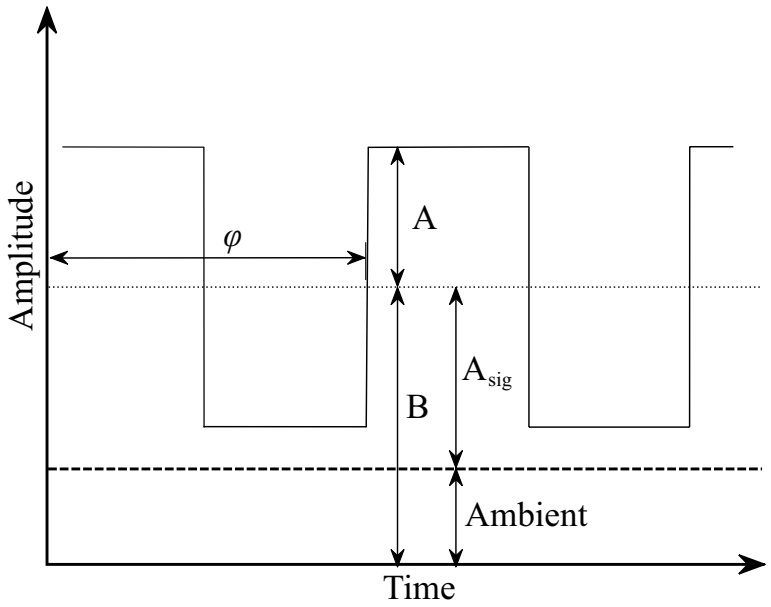


Figure 2. Graphical representation of variables in equations 4 to 6 of an illumination modulation signal. The variables are the amplitude A, offset B and the phase φ. The offset B can be divided into two parts: the part due to ambient light illustrated by the region below the dashed line, and an offset due to the modulation itself, A_{sig}.

$$\varphi = \tan^{-1}\left(\frac{A_0 - A_2}{A_1 - A_3}\right) \tag{6}$$

From the phase measurement φ, the distance d for each pixel is calculated as

$$d = \frac{c\varphi}{4\pi f_{mod}}, \tag{7}$$

where c is the speed of light and f_{mod} is the modulation frequency. The maximum distance the camera can measure before the returning illumination phase shifts by 2π is called the ambiguity distance d_{max}, which is calculated with

$$d_{max} = \frac{c}{2f_{mod}}. \tag{8}$$

The standard deviation σ_R of the measured distance is calculated with equation 9,[7] where c_{demod} is the demodulation contrast, and A_{sig} is the mean signal component resulting from the illumination source. This equation shows that the range measurement standard deviation σ_R is inversely proportional to the modulation frequency f_{mod}.

$$\sigma_R = \frac{c}{4\pi f_{mod}\sqrt{2}} \cdot \frac{\sqrt{B}}{c_{demod} \cdot A_{sig}} \tag{9}$$

1.2 Pseudo-noise modulation

There has been much literature on spread spectrum communications and their applications.[8] In Pseudo-Noise (PN) modulation the data $d(t)$ of bit length T_d is multiplied by a spreading code $c(t)$ with bit length T_c, where commonly $T_c < T_d$. The spectrum of the signal $d(t) \times c(t)$ is spread beyond the bandwidth of $d(t)$. The receiver then replicates the spreading code to demodulate the received signal to extract the original data $d(t)$.

The use of PN modulation with $T_c < T_d$ has been demonstrated in TOF cameras,[9,10] however this method has serious drawbacks. The precision is proportional to the modulation frequency of the light source, as described in

section 1.1. In current TOF cameras the light source is modulated as quickly as possible to maximize precision. Under PN modulation the light source's modulation bandwidth needs to handle frequencies up to $1/T_c$, however the average modulation frequency is lower than this, thus reducing the measurement precision.

A new technique where the code bit length is greater than the data length ($T_c > T_d$) is proposed. The technique using $T_c > T_d$ is used to encode the Global Positioning System (GPS).[11] The PN code is binary phase encoded onto the data by phase shifting the data by π radians. If the PN sequence is common to both the illumination and the sensor, the detected signal will not change. However, if the sequences are different, the PN sequence will cause relative changes of π radians in the modulation signal multiple times during an integration period, causing a second camera's light to appear as effectively DC (equivalent to background light). That is, the phase will change at seemingly random times, thus light will collect equally on the integration bins in the pixel structure explained in section 2 below. This encoding changes the $g(t)$ and $s(t)$ functions to

$$s(t) = [1 + \text{sgn}(\cos(\omega t - \varphi - \pi c(t)))], \tag{10}$$

$$g(t) = (1 + \text{sgn}(\cos(\omega t - \pi c(t))))/2. \tag{11}$$

There are a number of different PN sequences, and each sequence has multiple unique codes of variable length. Each PN sequence is generated differently and defined by its autocorrelation function (ACF), a measure of how a code interferes with itself, and the cross-correlation function (CCF), a measure of how different codes interfere with each other. The ACF and CCF should be as small as possible. The different sequences include maximum length sequences (m-sequences for short), gold sequences, Walsh-Hadamard sequences, Barker sequences, etc.[8]

M-sequences have similar properties to a purely random sequence of ones and zeros,[8] with a low ACF which is desired, but a high enough CCF to be undesirable. They are also simple to produce in digital circuits. Gold sequences are generated using special cases of m-sequences, and have opposite properties of m-sequences, undesirable ACFs and good CCFs. Walsh-Hadamard sequences are used in second and third generation cellular phone networks. The codes are generated from binary Hadamard Matrices, where each row is orthogonal to all the others when phase aligned. If the codes are misaligned they have high ACF and CCF values, which is undesirable. Cellular phone networks use the control channels to phase align the codes, however in a multiple range camera environment there is no method of synchronization between cameras thus Walsh sequences are not suitable.

M-sequences were selected as the most suitable PN code for this application. The ACF and CCF for m-sequences have been calculated in the literature.[12] The degree r of the polynomial determines the number of stages in the linear feedback register (LFR), and the code length m is determined by

$$m = 2^r - 1. \tag{12}$$

This shows m-sequences always have an odd length, containing one more '1' than '0'. Primitive polynomials can be calculated with software and are available in tables.[8]

2. EQUIPMENT SETUP

A diagram of the range camera is shown in figure 3. The light source is a bank of eight IR lasers, the image sensor is a PMD19-K2 (PMD Technologies, Germany)[13] and an Altera Stratix III field programmable gate array (FPGA)[14] is used to generate the modulation signals for the image sensor and light source, and to process the range data. A second bank of red lasers is used to simulate the light source from a second range camera.

The PMD19-K2 sensor acts as a high speed shutter demodulating the incoming light in accordance with a signal from the FPGA. To demodulate the light with no moving parts a pixel structure as shown in figure 4 is used.

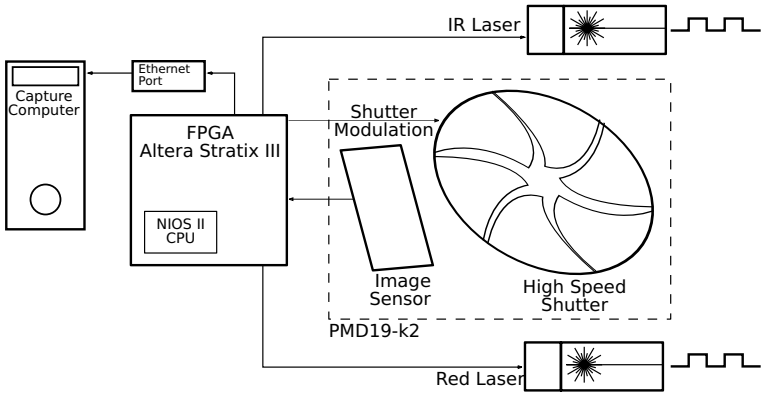


Figure 3. Time-of-flight range camera construction. A field programmable gate array (FPGA) is used to generate modulation signals for a PMD19-K2 sensor and illumination source, an amplitude modulated IR laser bank. The FPGA controls a second bank of red lasers to simulate the light source of another camera. Both laser banks have similar optical power outputs. The FPGA reads and processes the sampled frames from the image sensor and sends them to the host computer via Ethernet.

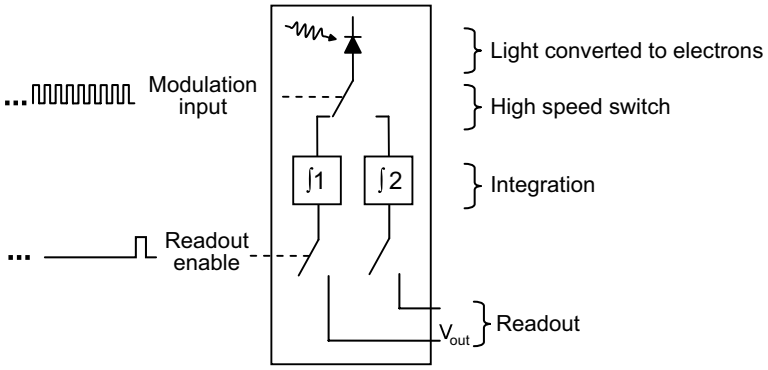


Figure 4. Internal pixel structure of PMD19-K2 sensor. Photons are converted to electrons and collected in two bins over time. By subtracting the output values of the two bins from each other, a sample A_0 to A_3 in equation 6 is calculated.

This particular sensor contains an array of 160 by 120 pixels.

An internal phase locked loop (PLL) in the FPGA generates the modulation signals for the light sources and the shutter. The PLL allows for dynamic reconfiguration of the frequency and relative phases of its output signals, providing a simple method to alter the frequency and phase of the modulation signals. This allows the simulation of another camera's light source by changing the modulation phase over time. Both banks of lasers have similar optical power output.

A series of experiments were conducted to measure the performance of the PN modulation scheme. The shutter and light source signals have a base modulation frequency of 30 MHz. An m-sequence of degree $r = 8$ was selected as it is short enough to quickly run numerical simulations and long enough to be effective as explained in section 3 below. The code is generated at a frequency of 100 kHz, hence the code $c(t)$ will change once every 300 modulation periods. Different modulation techniques for the shutter and light source can be compared by imaging a flat target at a set distance. This is achieved by calculating the mean of the measured phase for a region of 10 by 10 pixels (to reduce noise).

3. RESULTS AND DISCUSSION

To compare the performance of PN and traditional square wave modulation, the error of the measured phase is plotted against time in figure 5. The normal modulation (NM) curve is a single camera operating with a square wave modulated light source. This provides a base line reference to compare against. The normal modulation with interference (NMIF) curve illustrates the effect of the second light source, simulating a second interfering camera using normal modulation. The second camera's phase is swept from 0 to 2π radians in the eight second period resulting in varying error values. The point where the error decreases significantly corresponds to when the two cameras' light sources are phase aligned. The PN modulated with interference (PMIF) curve is the measured error with the camera using PN modulation (using an m-sequence of degree eight), while the second camera's operation is unchanged, continuing to use normal modulation. By applying PN modulation, the interfering light is rejected, allowing distance measurements in a multiple camera environment that previously was impractical due to the resultant errors. However there is an increase in the random measurement error.

Table 1 shows the mean and standard deviation for the measured phase of four different conditions, the three discussed above, and PN modulation (PM) with only one camera operating. The standard deviation between NM and PM is almost identical, indicating that the measurement precision is not significantly affected by PN modulation. However the mean values are different, indicating that PN modulation introduces a constant phase offset. This is surprising as none of the variables in equations 6 or 9 have changed. As this offset is constant it should be able to be calibrated out. A two sample t-test on the measured phase over 148 continuous images between NM and PM tests rejected the null-hypothesis (at the 5% significance level), indicating a statistically significant difference between the measured phase for the modulation schemes. This change remains unexplained.

Table 1. Statistics of measured phase for: a single camera with normal modulation (NM); PN modulation (PM); two cameras operating with normal modulation with interference (NMIF) and PN modulation on one camera with interference (PMIF). The interfering second camera is phase swept from 0 to 2π radians.

Test	**Mean**	**Standard deviation**
NM	3.0444	0.0026
PM	2.9808	0.0027
NMIF	2.6304	1.1270
PMIF	2.9838	0.0242

PMIF will have a larger standard deviation than NM because the photon shot noise increases due to the effective increase in ambient light from the second light source. The relative increase in measurement standard deviation for the PMIF case can be calculated using numerical simulations to calculate the new A_0 to A_3 values to use in equations 4, 5, and 9. Assuming the demodulation contrast c_{demod} does not change between NM and PMIF, the calculated relative change can be normalized to the measured NM value of 0.0026. The PMIF measurement standard deviation is calculated to be 0.0045. This is much less than the measured 0.0242 in table 1. However the measured phase is a function of the interfering camera's phase. With the systematic offset removed, a new standard deviation of 0.0046 is measured, almost the same as the calculated standard deviation.

Figure 6 shows the phase the PN modulated camera measured against the interfering camera's phase for the PMIF case. The dashed line is the mean measurement from a PM camera, a constant offset from a NM camera, the solid line. The difference between the measured and PN phase is caused by a constant offset as seen in the PM case, and is also dependent on the phase of the interfering camera. The interfering phase causes a significant effect as the difference between the smallest and greatest measured returning phase is greater than three standard deviations of each measurement. Although there is still some dependence on the interfering phase, this decrease in accuracy is small compared to the NMIF.

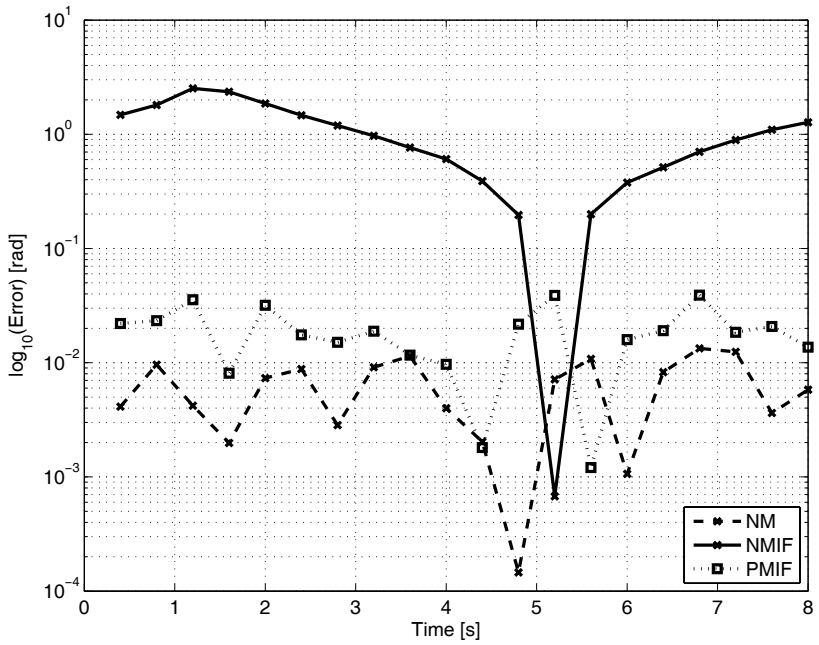


Figure 5. Error in measured phase using a time-of-flight camera comparing PN and square wave modulation showing: one camera using square wave modulation, (normal modulation NM); two time-of-flight cameras both using square wave modulation, (normal modulation with interference NMIF); two cameras with pseudo-noise modulation applied to one camera, (PN modulation with interference PMIF). In the last two cases, interfering camera's phase is swept from 0 to 2π radians over eight seconds.

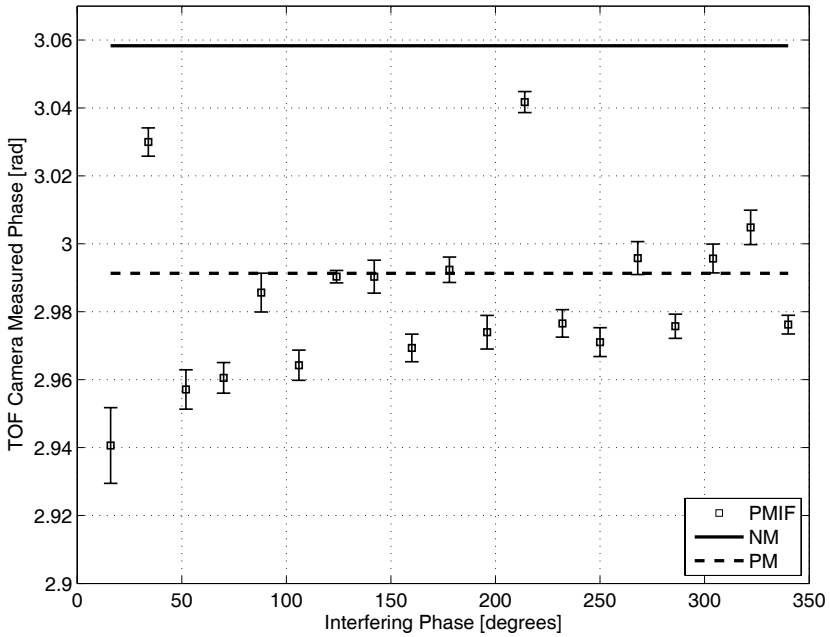
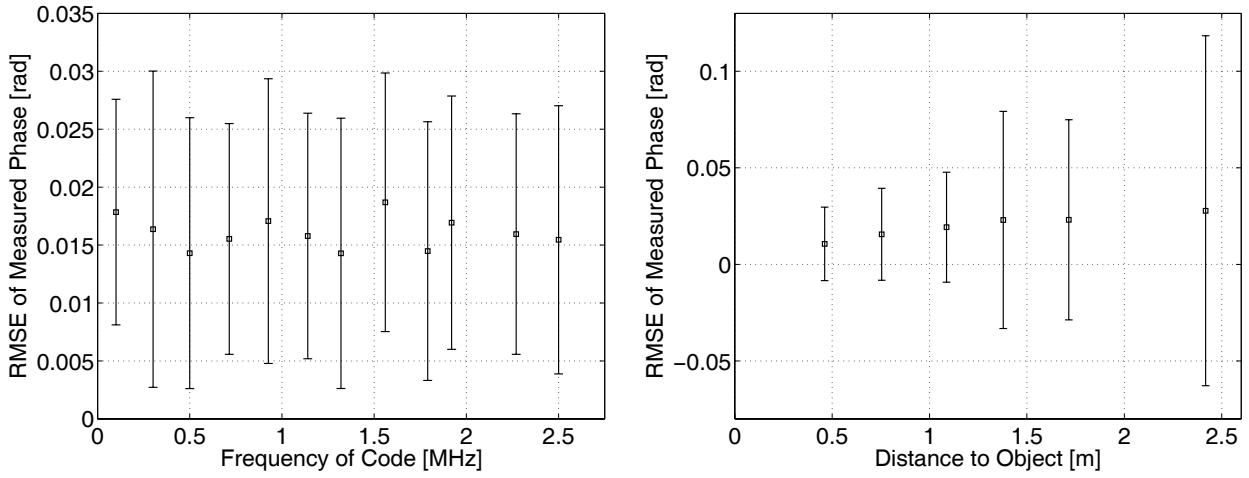


Figure 6. Results from two time-of-flight cameras illuminating the same scene. The measuring camera uses PN modulation and the interfering camera uses square wave modulation and sweeping phase from 0 to 2π radians. The measured returning phase is plotted against the phase of the interfering camera. The solid line is the measured phase from a normal modulated (NM) camera. The dashed line is what a PN modulated (PM) camera measures.



(a) (b)

Figure 7. Root mean square error (RMSE) of phase measurement for two time-of-flight cameras operating simultaneously. The measuring camera uses PN modulation while the interfering camera uses square wave modulation which continuously shifts in phase. Test (a) alters the PN code generation frequency, and (b) changes the distance to the object from the camera. The increase in uncertainty in (b) is due to less optical power returning to the camera because the object is further away. Both variables have no significant effects on RMSE.

The root mean square error (RMSE) of the returning phase in radians versus m-sequence frequency is plotted in figure 7a for a PMIF setup. The modulation frequency f_{mod} is 30 MHz, and the m-sequence frequency f_c is changed between 100 kHz and 2.5 MHz. The output modulation bit stream changes phase by π radians when the code bit changes, this doubles the period for the current modulation bit. When $f_c \rightarrow f_{mod}$ the phase change occurs frequently enough for f_{mod} to be changed, affecting the measurement precision. When $(1/f_c) \cdot m <$ integration period, the full m-sequence has not been transmitted, allowing for possible interference. For $f_c \ll f_{mod}$ and $(1/f_c) \cdot m \leq$ integration period, f_c should have no affect on measurement precision. In figure 7a no change in error is observed, as expected because f_c keeps within the above bounds.

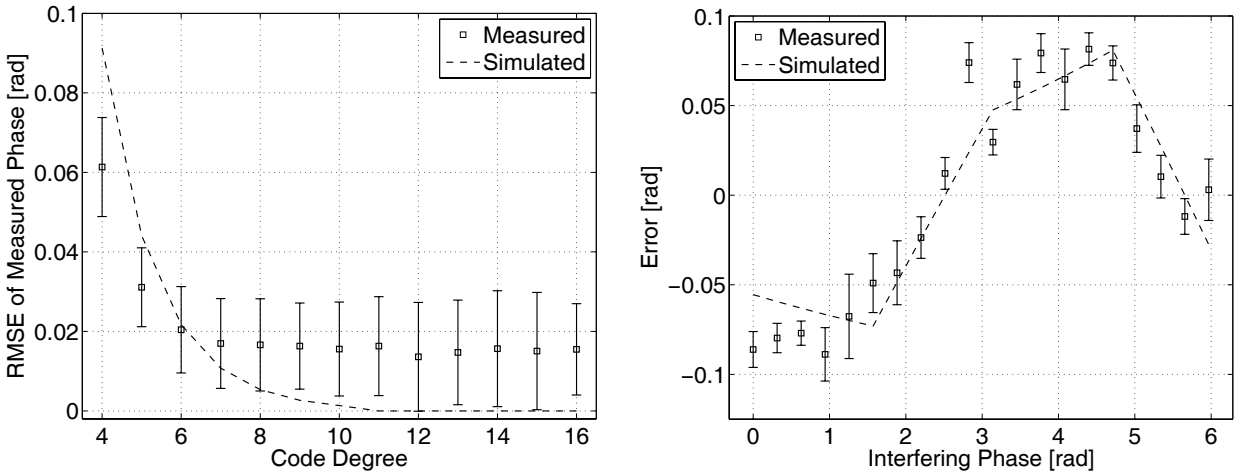
The RMSE of measured returning phase versus distance to the object for a PMIF setup is plotted in figure 7b. The increase in uncertainty as the distance d increases is caused by the decrease in returning optical power by $1/d^2$. The figure indicates that error introduced by PN modulation does not appear to be dependent on distance.

Figure 8a depicts the code length versus RMSE for measured and simulated results for a PMIF setup. The simulation shows that the error should nearly halve every time the degree r is increased by one. This is because the m-sequence length approximately doubles when degree r is increased by one, as shown by equation 12. The error is caused by m-sequences being unbalanced, having one more '1' than '0'. The PN changes the shutter modulation by π radians, thus with a short code one integration bin in the pixel structure is favored. As the code length increases, the effect of it being unbalanced decreases. The RMSE is can be approximated as,

$$\text{RMSE} \propto \frac{1}{2^r - 1}. \tag{13}$$

The measured RMSE doesn't follow the simulated results after degree 7, instead it plateaus. That is, above degree 7, other error sources become dominant.

Figure 8b depicts the simulated result and measured error versus interfering phase for a degree 4 PN modulated camera in a PMIF setup. The simulation agrees with the measured data, where the misalignment could be from a number of factors, e.g. an uneven modulation duty cycle.



(a) (b)

Figure 8. Simulated and measured results for two time-of-flight cameras operating simultaneously. The measuring camera uses pseudo-noise modulation while the interfering camera is square wave modulated with a continuously shifting phase. The two graphs are, (a) the root mean square error (RMSE) versus m-sequence degree r, (b) error versus the interfering camera's phase for a degree 4 m-sequence.

In TOF cameras the demodulation contrast c_{demod} is the difference between the received optical power in the high and low parts of the modulation signal.[5] Ideally this is as large as possible. The maximum value of c_{demod} is limited by sensor saturation and the minimum value by the ambient light. As an increasing number of cameras illuminate a scene, the effective ambient light increases, therefore decreasing the possible modulation contrast and lowering the achievable distance measurement precision. Some range imaging sensor designs reportedly incorporate background light rejection features. Such a sensor could be less susceptible to this decrease in modulation contrast.

4. CONCLUSION

Our ability as humans to perceive our environment in three dimensions is one of our most important senses. TOF cameras allow an equivalent sense for three dimensional machine vision applications. Principles of TOF cameras, PN modulation, and custom camera design have been presented. The setup uses a PMD19-K2 image sensor, laser diode illumination and an FPGA to control the camera and generate modulation signals.

We have shown that the distance to objects in a scene can still be accurately measured by a TOF camera using PN modulation while working in a two camera environment, with no communication or synchronization between cameras. When using square wave modulation with two cameras operating at the same modulation frequency, distance measurements become corrupt. A PN modulated camera operating alone does not have a negative impact on the measurement precision. A constant offset was observed, however this offset could easily be removed during camera calibration. In environments with multiple TOF cameras operating there is a decrease in precision caused by the calculable increase in photon shot noise. There is also a reduction in measurement accuracy, due to a systematic error that is a function of the other camera's illumination phase. Measurements have shown that this error can be made negligible by selecting long m-sequences. Only two cameras operating simultaneously were tested, but this technique has potential to operate with many more cameras.

ACKNOWLEDGMENTS

The authors would like to thank all the members of Chronoptics, the range imaging research group at the University of Waikato. Thanks also to Adrian Jongenelen at the University of Victoria for initial design and setup of the range imaging equipment, and funding from the University of Waikato strategic investment fund.

REFERENCES

[1] MESA Imaging, Zurich, Switzerland, *SR4000 User Manual* (2008). Version 0.1.2.2.

[2] May, S., Werner, B., Surmann, H., and Pervölz, K., "3D time-of-flight cameras for mobile robotics," *IEEE/RSJ International Conference on Intelligent Robots and Systems* , 790 – 5 (2006).

[3] Guomundsson, S. A., Aans, H., and Larsen, R., "Environmental effects on measurement uncertainties of time-of-flight cameras," *ISSCS - International Symposium on Signals, Circuits and Systems, Proceedings* **1**, 113 – 116 (2007).

[4] Oggier, T., Lehmann, M., Kaufmann, R., Schweizer, M., Richter, M., Metzler, P., Lustenberger, F., Blanc, N., and Lang, G., "An all-solid-state optical range camera for 3D real-time imaging with sub-centimeter depth resolution (swissranger)," *Proc. SPIE* **5249**, 534 – 545 (2004).

[5] Payne, A., Dorrington, A., Cree, M., and Carnegie, D., "Characterization of modulated time-of-flight range image sensors," *Proc. SPIE* **7239**, 723904 (2009).

[6] Kahlmann, T., Remondino, F., and Ingensand, H., "Calibration for increased accuracy of the range imaging camera swissranger," *ISPRS Commission V Symposium 'Image Engineering and Vision Metrology'* , 136 – 141 (2006).

[7] Büttgen, B., Oggier, T., Lehmann, M., Kaufmann, R., and Lustenberger, F., "CCD/CMOS lock-in pixel for range imaging: Challenges, Limitations and State-of-the-Art," *1st Range Imaging Research Day* , 31–32 (2005).

[8] Peterson, R. L., Ziemer, R. E., and Borth, D. E., [*Introduction to Spread Spectrum Communications*], Prentice Hall, Upper Saddle River, NJ (1995).

[9] Büttgen, B., Mechat, M., Lustenberger, F., and Seitz, P., "Pseudonoise optical modulation for real-time 3-D imaging with minimum interference," *IEEE Trans. Circuits Syst. I* **54**(10), 2109 – 19 (2007).

[10] Büttgen, B. and Seitz, P., "Robust optical time-of-flight range imaging based on smart pixel structures," *IEEE Trans. Circuits Syst. I* **55**, 1512 – 25 (July 2008).

[11] Parkinson, B. W. and Spilker, J. J., [*The global positioning system: theory and applications*], vol. 1, American Institute of Aeronautics and Astronautics, Reston, VA (1996).

[12] Sarwate, D. and Pursley, M., "Crosscorrelation properties of pseudorandom and related sequences," *Proc. IEEE* **68**, 593 – 619 (May 1980).

[13] PMD Technologies, Siegen, Germany, *PMD 19k-2 High Resolution 3D Video Sensor Array 160x120 Pixels Data Sheet* (2008). Version 1.0.

[14] Altera Corporation, San Jose, CA, *Volume 1 - Stratix III Device Handbook* (2009). Version 1.8.