

Simultaneous Phase and Frequency Stepping in Time-of-Flight Range Imaging

Carl A. Lickfold, Lee Streeter, Michael J. Cree, and Jonathan B. Scott

School of Engineering, The University of Waikato, Private Bag 3105, Hamilton 3240, New Zealand

Email: cal20@students.waikato.ac.nz

Abstract—Time-of-flight (ToF) range imaging is an established full-field depth acquisition technology. ToF is prone to several forms of error, notably multi-path interference (MPI). Frequency stepping ToF mitigates these errors, but efficient computational methods suffer from bandwidth limitations. We introduce a new acquisition method based on simultaneous phase and frequency stepping, which is capable of running in real time and reduces error compared to phase stepping and frequency stepping alone.

Index Terms—Time-of-flight, range imaging, phase stepping, frequency stepping, multi-path interference

I. INTRODUCTION

Time-of-flight (ToF) range imaging is an efficient depth acquisition solution with a wide range of machine vision applications. Amplitude modulated continuous wave (AMCW) ToF encodes scene depth as the phase shift between the light modulation envelope and the sensor modulation at every pixel [1]. This acquisition scheme has been favoured due to its simplicity in depth computation from as few as three raw frames.

AMCW suffers from several forms of error [2]–[5]. For example, it is possible for light to return by multiple propagation paths to each pixel, causing significant error known as multi-path interference (MPI) [2]. Another major issue is phase wrapping. Higher modulation frequencies are desired because they offer improved range precision over low frequencies, but the unambiguous range is inversely proportional to modulation frequency [6]. Recent advances in ToF sensors, like the Microsoft Azure Kinect, have resulted in maximum modulation frequencies of several hundred MHz [7]. While methods exist to disambiguate high frequency range measurements they require full acquisitions at multiple modulation frequencies, and are confounded by motion artefacts [6]. Recent work [8], [9] has highlighted the benefits of radar acquisition methods applied to ToF hardware. As the bandwidth continues to increase, these methods are becoming more advantageous.

Previous works have investigated an alternative acquisition scheme to AMCW known as stepped frequency continuous wave (SFCW) [8], [9]. This method has several advantages over phase stepping. Multiple propagation paths are separable up to the bandwidth limit of the acquisition and the unambiguous range can be extended by increasing the number of frequency steps [8], [9]. Previous implementations of SFCW use the so-called MUSIC algorithm [10], which has high computational complexity, limiting potential uptake, and previous

work has used calibrated data sets to evaluate its performance in mitigating MPI. A solution that works on uncalibrated data is desirable.

In this work we present an alternative approach by simultaneously stepping phase and frequency. We show that imperfections in the sensor and other background offsets can be separated from range measurements with this method. Our work presents a step towards real-time implementation of radar techniques on current generation ToF sensors. We evaluate the performance on raw data, which could be acquired at video frame rates, and show reduced error when compared to AMCW and SFCW alone.

II. TIME-OF-FLIGHT CAMERA OPERATION

A. AMCW and SFCW

AMCW operates at a single modulation frequency, ω_k where k is constant, and the correlation between transmitted light and reflected light is sampled by stepping the modulation phase θ_j between the light source and sensor by $\Delta\theta$ with each capture. The sensor and light source are modulated at the same frequency (i.e. homodyne modulation) for the duration of the integration period T and the resulting intensity measurement after the j th acquisition is given by (in a general form to encompass both AMCW and SFCW)

$$I_{jk}(\omega_k) = \frac{1}{2}a \cos(\phi_k - \theta_j) + b, \quad (1)$$

where a is the brightness of the light return, b is the background offset, and

$$\phi_k = \frac{2\omega_k d}{c} \quad (2)$$

is the phase shift due to the distance that the light travels to the scene and back, c is the speed of light [11].

A minimum of three raw frames are required to recover ϕ , a , and b at each pixel [11]. Taking phase steps evenly spanning the range 0 to 2π , the discrete Fourier transform (DFT) recovers the phase ϕ and amplitude a from the fundamental DFT bin. The background offset b is recovered from the DC bin. Depth is calculated by rearranging Eqn. 2 for d . The recovered phase is always within the range of 0 to 2π , so points in the scene which exceed distance

$$d_{\text{ambg}} = \frac{c\pi}{\omega_k} \quad (3)$$

will “phase wrap” and appear closer than what they actually are [6].

In SFCW acquisition we set the phase step $\Delta\theta = 0$, thus the modulation phase is held constant, e.g. $\theta_j = 0 \forall j$, and the frequency is instead evenly stepped at values ω_k in increments $\Delta\omega$ [8]. Various methods, such as MUSIC and the DFT, are used to obtain a spectral estimate from the raw frames [8], [9]. The distance is calculated from the spectral estimate ω_{est} according to

$$d = \frac{2c\omega_{est}}{B}. \quad (4)$$

The frequency range of the capture sequence determines the bandwidth B , which in turn determines the minimum resolvable range, ΔR , when using the DFT frequency estimation, and is given by

$$\Delta R = \frac{c}{2B}. \quad (5)$$

The ambiguity distance is only limited by aliasing which is determined by the frequency spacing

$$d_{ambg} = \frac{c\pi}{\Delta\omega}. \quad (6)$$

Given appropriate hardware, $\Delta\omega$ may be arbitrarily reduced, increasing d_{ambg} [8].

Under SFCW, the bandwidth and the distance determine the number of observed cycles at each pixel. The limited sensor modulation bandwidth causes limited range resolution. Furthermore, the background offset, which is near DC, interferes with the spectral estimation and is not separable for near returns closer than ΔR .

B. Stepped Phase and Frequency Continuous Wave (SPFCW)

We propose a hybrid acquisition method by simultaneously stepping phase θ_j and frequency ω_k in linear increments, so $j = k$. According to Eqn. 1, SPFCW causes observation of many cycles, despite limited bandwidth. Therefore, the introduced change in phase between frequency steps, $\Delta\theta/\Delta\omega$ increases the Fourier frequency of each distance. The computed distance is simply adjusted back to its correct position after calibration. The tradeoffs for this advantage are the ambiguity distance is reduced, and part of the Fourier spectrum (computed range) is lost by shifting it behind the camera.

III. METHODS

All data are captured using a TIDA-01173 Camera Development Kit (CDK) (Texas Instruments, Dallas, Texas). The CDK comprises an OPT8241 QVGA ToF sensor, an OPT9221 ToF controller and a laser light source. The system, based on the design developed by [12], externally modulates the sensor and light source with signals provided by an AD9959 direct digital synthesizer (DDS) evaluation board (Analog Devices, Norwood, Massachusetts). This system is designed for bandwidths of up to 100 MHz, too low to be ideal for SFCW, however with our proposed data capture and processing method, we demonstrate useful measurements when operating in SFCW modes up to 150 MHz.

The data are captured as sets of 16 phase steps, and 30 replicate sets are acquired at each frequency. The frequency is stepped from 10 MHz to 150 MHz and sampled at 5 MHz

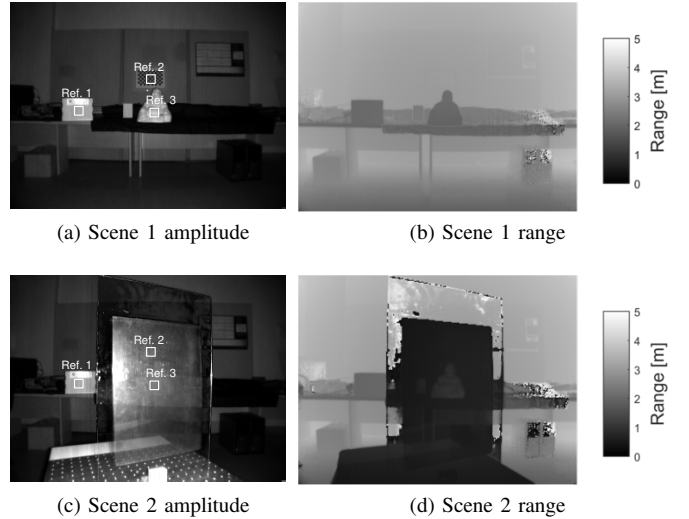


Fig. 1. AMCW measurements at 30 MHz are shown for a visual comparison. The reference locations, used for error analysis, are white squares of 10 by 10 pixels. The regions are shown for scene 1 (a) and scene 2 (c).

intervals. The integration time is linearly increased from 300 μ s at 10 MHz to 4.9 ms at 150 MHz to mitigate the effect of amplitude reduction at high frequencies.

Two scenes are used to evaluate the proposed method and all points in the field of view for both scenes are less than 5 m. Scene 1, Fig. 1a, is set up to minimize MPI and is used for calibration. Scene 2, Fig. 1c, is heavily contaminated with MPI and a translucent sheet acts as a foreground scatterer. The translucent sheet is intentionally positioned so that at least one range bin ΔR separates the background scene and the foreground sheet. AMCW measurements at 30 MHz are selected for a comparative analysis for its unambiguous range of approximately 5 m, matching the scene.

IV. RESULTS AND DISCUSSION

SPFCW is evaluated at three different phase step values of $\pi/8$, $\pi/4$ and $3\pi/8$, these are referred to as SPFCW-1, SPFCW-2, and SPFCW-3 respectively. The spectral estimates are obtained by applying a zero-padded DFT of size 2048, see Fig. 2b. The range for SFCW and SPFCW is determined by performing a peak search of the DFT starting from 0 m. Secondary peaks are taken as the second return if they are larger than a set threshold, shown in Fig. 3c and 3d. The DFT is converted to a range using Eqn. 4; all range measurements for both scenes are calibrated to point Ref. 1 in scene 1.

Fig. 2a clearly demonstrates that SPFCW increases the number of observed cycles compared to SFCW. Intrinsic offsets in the sensor cause large DC offsets in the raw data at each frequency, and these offsets change with frequency. For both SFCW and SPFCW, the mean is subtracted to prevent spectral leakage due to the large DC offset. Unlike SFCW, SPFCW increases the frequency of the range measurements, separating them from the DC contamination. Shifting the range measurement back to the correct position, the DC contamination

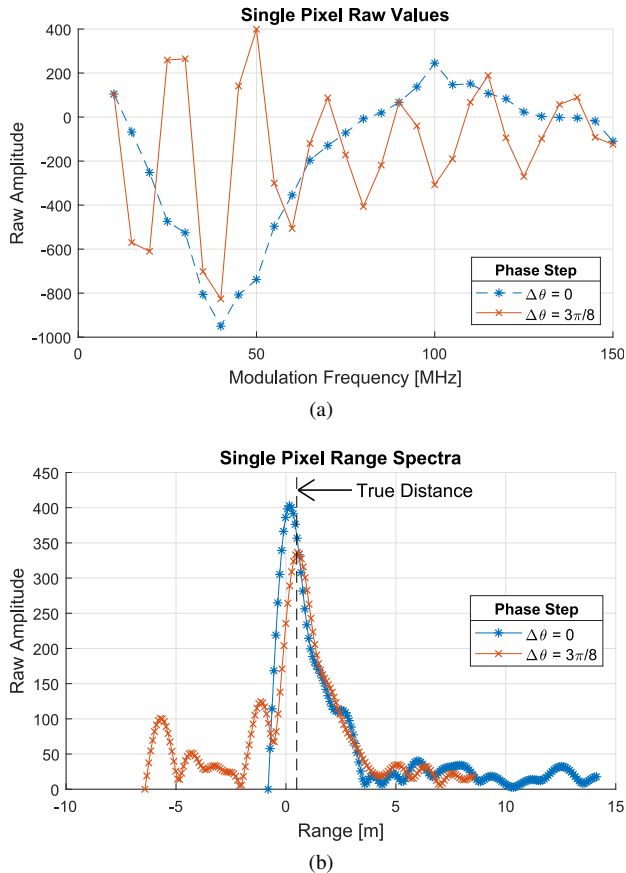


Fig. 2. (a) shows raw values for SFCW and SPFCW for a single pixel at location (175, 230) with a known distance of 0.49 m; (b) is the corresponding range spectra. Both range spectra show values at negative distances due to calibration. SPFCW increases the number of observed periods and the intrinsic sensor offsets are shifted to a further negative range value.

appears at easily identified non-physical negative distances. Fig. 2b shows the calibrated single-sided range spectrum for SFCW and SPFCW. The peak using SPFCW is much closer to the true distance than SFCW, and several peaks appear at negative distances due to the intrinsic offsets in the sensor. The reduced ambiguity distance with SPFCW is also apparent.

The root-mean-square error (RMSE) at three 10 by 10 pixel reference locations are compared between methods. The reference locations, which are manually measured using a Leica DISTO™ plus (Leica Geosystems AG, Heerbrugg, Switzerland), are shown in Fig. 1. The error in these regions is summarised in tables I and II and the lowest error for each region is shown in bold. With the exception of AMCW at Ref. 1 in scene 1, our calibration region, SPFCW outperforms AMCW and SFCW in all other regions. SPFCW is capable of recovering multiple returns at each pixel and the secondary return for SPFCW-3, shown in Fig. 3d, has an RMSE of 0.162 m at Ref. 3.

AMCW is not capable of distinguishing multiple propagation paths, therefore the range image in Fig. 1d shows significant error due to MPI in the region of the statue. The primary return for SPFCW-3 in Fig. 3b shows almost no MPI

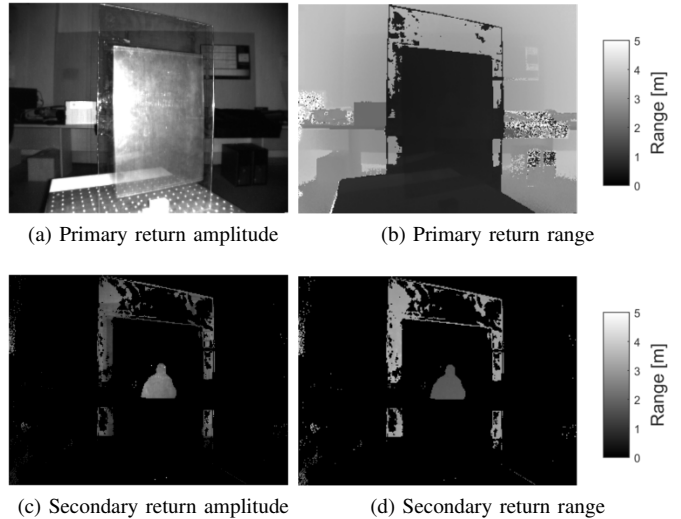


Fig. 3. The amplitude and range is recovered for scene 2 using SPFCW-3. The primary (a), (b) and secondary (c), (d) returns are extracted from the range spectra and show reduced error compared to AMCW and SFCW.

TABLE I
SCENE 1 RMSE [m]

Location	AMCW	SFCW	SPFCW-1	SPFCW-2	SPFCW-3
Ref. 1	0.010	0.016	0.018	0.015	0.016
Ref. 2	0.164	1.552	0.225	0.038	0.180
Ref. 3	0.099	0.097	0.053	0.075	0.133

TABLE II
SCENE 2 RMSE [m]

Location	AMCW	SFCW	SPFCW-1	SPFCW-2	SPFCW-3
Ref. 1	0.071	0.019	0.029	0.014	0.024
Ref. 2	0.105	0.361	0.341	0.133	0.090
Ref. 3	0.457	0.396	0.318	0.153	0.081

in the region of the statue. However, there are multiple returns which are not separable in the lower region of the sheet due to the limited bandwidth.

V. CONCLUSION

We show that SPFCW has reduced error when compared to AMCW and SFCW. The disadvantage of SPFCW is more frames are required with the implemented DFT spectral estimation compared to AMCW. SPFCW is able to separate intrinsic sensor offsets from range measurements and the primary and secondary returns from SPFCW have significantly less error in an MPI contaminated scene. This work presents a first-step towards running ToF range imaging cameras in real-time using SPFCW. Future work will investigate efficient spectral estimation, and evaluate motion artefacts with dynamic scenes using this new range imaging modality.

ACKNOWLEDGMENT

This work was funded by a Royal Society of New Zealand Marsden Fund Council Fast Start grant (UOW1501).

REFERENCES

- [1] C. S. Bamji, P. O'Connor, T. Elkhatib, S. Mehta, B. Thompson, L. A. Prather, D. Snow, O. C. Akkaya, A. Daniel, A. D. Payne, T. Perry, M. Fenton, and V. Chan, "A 0.13 μm CMOS system-on-chip for a 512×424 time-of-flight image sensor with multi-frequency photo-demodulation up to 130 MHz and 2 GS/s ADC," *IEEE Journal of Solid-State Circuits*, vol. 50, no. 1, pp. 303–319, 2015.
- [2] R. Z. Whyte, "Resolving measurement errors inherent with time-of-flight range imaging cameras," Ph.D. dissertation, University of Waikato, 2015.
- [3] J. P. Godbaz, "Ameliorating systematic errors in full-field AMCW lidar," Ph.D. dissertation, University of Waikato, Hamilton, New Zealand, 2012.
- [4] A. D. Payne, A. A. Dorrington, M. J. Cree, and D. A. Carnegie, "Improved measurement linearity and precision for AMCW time-of-flight range imaging cameras," *Applied Optics*, vol. 49, no. 23, pp. 4392–4403, 2010.
- [5] M. Hansard, S. Lee, O. Choi, and R. P. Horaud, *Time-of-flight cameras: principles, methods and applications*. Springer Science & Business Media, 2012.
- [6] A. P. P. Jongenelen, D. A. Carnegie, A. D. Payne, and A. A. Dorrington, "Maximizing precision over extended unambiguous range for ToF range imaging systems," in *Proceedings of the IEEE International Instrumentation and Measurement Technology Conference*, Austin, Texas, 2010, pp. 1575–1580.
- [7] C. S. Bamji, S. Mehta, B. Thompson, T. Elkhatib, S. Wurster, O. Akkaya, A. Payne, J. Godbaz, M. Fenton, V. Rajasekaran *et al.*, "IMpixel 65nm BSI 320MHz demodulated TOF image sensor with $3\mu\text{m}$ global shutter pixels and analog binning," in *2018 IEEE International Solid-State Circuits Conference-(ISSCC)*. IEEE, 2018, pp. 94–96.
- [8] R. Whyte, L. Streeter, M. J. Cree, and A. A. Dorrington, "Application of lidar techniques to time-of-flight range imaging," *Applied Optics*, vol. 54, no. 33, pp. 9654–9664, Nov 2015.
- [9] A. Kadambi, J. Schiel, and R. Raskar, "Macroscopic interferometry: Rethinking depth estimation with frequency-domain time-of-flight," in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Seattle, WA, USA, 2016, pp. 893–902.
- [10] P. Stoica and R. Moses, "Introduction to spectral analysis. vol. 1 prentice hall," *Upper Saddle River, NJ*, 1997.
- [11] A. D. Payne, "Development of a full-field time-of-flight range imaging system," Ph.D. dissertation, The University of Waikato, 2008.
- [12] S. Shrestha, F. Heide, W. Heidrich, and G. Wetzstein, "Computational imaging with multi-camera time-of-flight systems," *Transactions on Graphics*, vol. 35, no. 4, pp. 33:1–33:11, 2016.