Extracting the MESA SR4000 Calibrations

Sean A. Charleston, Adrian A. Dorrington, Lee Streeter, Michael J. Cree

University of Waikato, Private Bag 3105, Hamilton 3240, New Zealand

ABSTRACT

Time-of-flight range imaging cameras are capable of acquiring depth images of a scene. Some algorithms require these cameras to be run in 'raw mode', where any calibrations from the off-the-shelf manufacturers are lost. The calibration of the MESA SR4000 is herein investigated, with an attempt to reconstruct the full calibration. Possession of the factory calibration enables calibrated data to be acquired and manipulated even in 'raw mode'. This work is motivated by the problem of motion correction, in which the calibration must be separated into component parts to be applied at different stages in the algorithm. There are also other applications, in which multiple frequencies are required, such as multipath interference correction. The other frequencies can be calibrated in a similar way, using the factory calibration as a base. A novel technique for capturing the calibration data is described; a retro-reflector is used on a moving platform, which acts as a point source at a distance, resulting in planar waves on the sensor. A number of calibrations are retrieved from the camera, and are then modelled and compared to the factory calibration. When comparing the factory calibration to both the 'raw mode' data, and the calibration described herein, a root mean squared error improvement of 51.3 mm was seen, with a standard deviation improvement of 34.9 mm

Keywords: Time of Flight, Range Imaging, Calibration, MESA, SR4000

1. INTRODUCTION

Time-of-flight range imaging cameras operate on the principle that the precise knowledge of the speed of light can be used to determine the distance to an object. In practice, this is often inferred indirectly, using a technique called *Amplitude Modulated Continuous Wave* (AMCW), due to the high electronic complexity and cost required to directly measure the flight time.[1] In AMCW systems, the amplitude of the transmitted light is modulated, and the phase change of the received signal is measured to infer distance.[2] These cameras require at least three captures (known as phase steps), in order to produce a single phase/amplitude image combination.

The SR4000 has a number of built in calibrations, which are unavailable when using the camera in 'raw mode' (acquiring only the raw phase steps). A number of applications require direct access to the 'raw mode' data, for example the mixed pixel restoration algorithm[3] captures raw data at two frequencies to resolve multipath interference. Other applications include an upcoming motion correction algorithm, which is part of the inspiration for this work.

As with any optical system, calibration is a key step to produce precise and repeatable results. Time-of-flight cameras introduce a variety of additional sources of noise and error compared to standard 2D cameras, due to a number of complications, including the transmission and reception of modulated light. The MESA SR4000 is an industrial grade time-of-flight camera, and is regarded to have a relatively high quality calibration.[4, 5]

In this paper, we present a novel technique for acquiring data for calibration, using a translation stage and a retro-reflector. This technique attempts to produce planar waves on the sensor, providing a robust calibration dataset. Two sets of images are captured, one in 'raw mode' and one with the factory calibrations applied. Methods are applied to retrieve each of the factory calibrations, based on the 'raw mode' data. Finally, the calibrations obtained are analysed, and the applications and limitations of the technique are discussed.

Further author information:
S.A.C.: E-mail: sac60@students.waikato.ac.nz
A.A.D.: E-mail: adrian@waikato.ac.nz

2. BACKGROUND

2.1 Time-of-flight principles

Time-of-flight cameras using the AMCW technique determine the phase of a returning modulation waveform by correlation with the original modulation waveform. At least three phase steps (τ_n for the measurement at the nth phase step), with additional phase offsets (θ_n) added, are required to resolve the phase (ϕ), amplitude (α), and background level (B) of the returning modulation waveform. The modulation waveform is assumed to be sinusoidal, as shown in Figure 1. Four phase steps ($\tau_1, ..., \tau_4$) are typically used, as it simplifies the calculations for hardware implementation.[6] The recovery of these signals is performed by a process known as demodulation, which is as follows,[7]

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

$$\alpha = \frac{\sqrt{(\tau_1 - \tau_3)^2 + (\tau_2 - \tau_4)^2}}{2}, \tag{2}$$

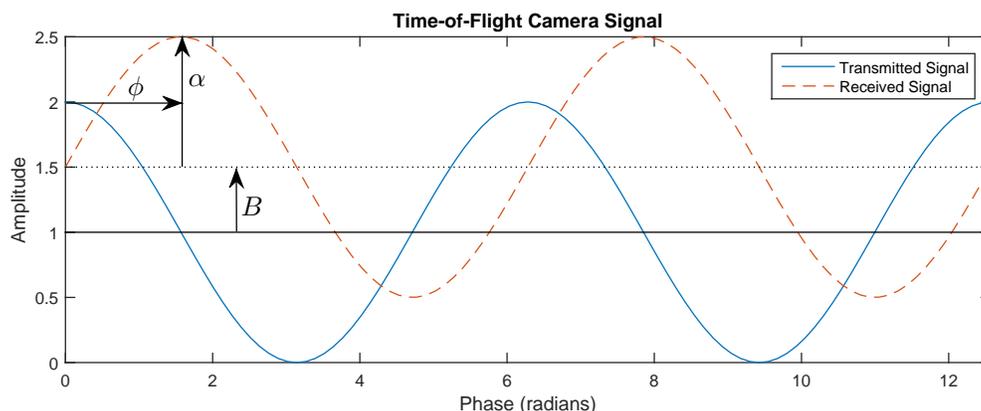$$B = \frac{\tau_1 + \tau_2 + \tau_3 + \tau_4}{4}. \tag{3}$$

Figure 1: Time-of-flight camera signals can be illustrated as a transmitted and received cosine wave. The received signal gives the phase change, amplitude, and background light level.

The standard demodulation technique computes the first bin of a four point discrete Fourier transform.[8] This complex phasor form can be generalised to any number of phase steps.

$$\begin{aligned}
P_{jk} &= \sum_{n=1}^{4} \tau_{jkn} \cdot e^{-i\theta_n} \\
&= 2A_{jk}e^{i\phi_{jk}} \\
&= \alpha_{jk}e^{i\phi_{jk}},
\end{aligned} \tag{4}$$

where P is the phasor for a particular pixel in row j and column k of the pixel array, $i = \sqrt{-1}$, and A is each phase step's amplitude.

This phasor is constructed by demodulating the received signal with the transmitted signal. Each of the phase steps can be represented as,

$$\tau_{jkn} = A_{jk}\cos(\phi_{jk} + \theta_n) + B_{jk}. \tag{5}$$

Time-of-flight camera signals are modulated at a particular modulation frequency (f_m). This modulated light must travel from the camera, to the object, and back to the camera's lens. If this total distance is greater than

the wavelength of the modulated signal, the signal will wrap around and the camera will report an incorrect distance. This maximum distance is known as the ambiguity distance (d_a),

$$d_a = \frac{c}{2f_m},$$ (6)

where c is the speed of light.

The distance to each pixel can then be calculated, based on the phase change and ambiguity range,

$$d_{jk} = d_a \cdot \frac{\phi_{jk}}{2\pi}.$$ (7)

2.2 Error sources in time-of-flight cameras

To build a calibration model we first enumerate offset error sources relevant to time-of-flight imaging. The first error source, fixed in time and independent of range, is a global phase offset ($0\,\text{rad} \neq 0\,\text{mm}$). This constant phase offset is due to fixed propagation delays in the electronics, and is corrected by applying a constant phase correction to all pixels. Secondly, there is a gradual phase offset across the sensor, caused by clocking propagation delays.[9] This gradual phase offset is usually grouped with the global phase offset. Finally, there is a fixed pattern phase offset, due to additional fixed per-pixel delays, including different material properties in the CMOS gates.[10] Example data points are illustrated in Figure 2, where the three sources of phase error arise as offsets to the target plane. These three phase offsets can be combined into one phase correction, δ_{jk} per pixel.
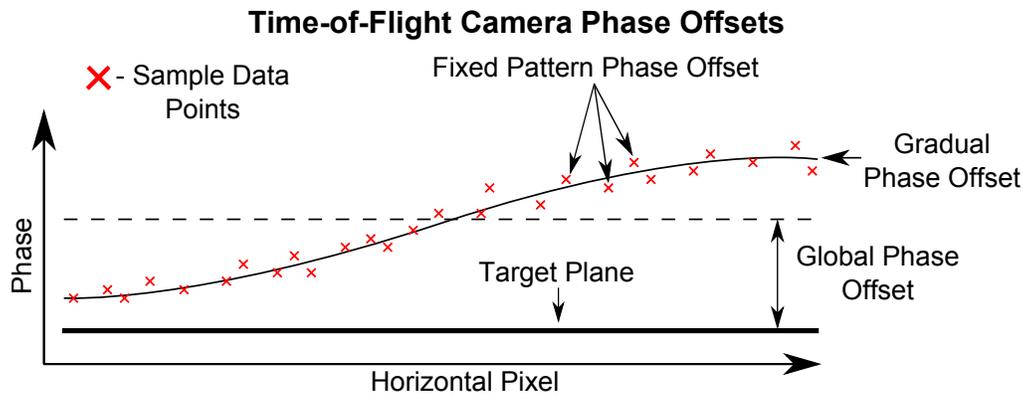

Figure 2: An example is shown of a row of data points across an image that might be acquired from the MESA SR4000 when imaging a target plane. The red crosses show examples of the 'raw data', along with an illustrative curve representing the gradual phase offset. The fixed pattern phase offset is also shown as phase variation from the gradual phase offset. The global phase offset is represented by a fixed offset for all pixels.

Phase offsets that are dependent on the range to the scene ($\rho_{jk\phi}$) are caused by harmonic distortion, namely harmonics (odd harmonics only if four phase steps are used) of the modulation signal (due to square wave modulation and/or non-linearities) aliased to the fundamental.[11] The error presents itself as a calibratable oscillation on the undistorted phase. Figure 3 is an example of harmonic distortion, showing how the sample data points oscillate about the undistorted phase.

There are also three sources of amplitude error. The first is a fixed pattern gain factor (η_{jk}), due to variation in the quantum efficiency of each pixel, as well as differences in integration storage capacitance (i.e. the same number of photoelectrons produce different voltages at each pixel). The result is a multiplicative error dependent on the signal amplitude.[12] The second source of amplitude error, due to uneven illumination on the sensor (σ_{jk}) is also multiplicative in nature. The spatial output profile of the light source and lens vignetting results in the apparent illumination being strongest in the centre of the image. Finally, the SR4000 includes a distance dependent amplitude correction (γ_ϕ), which looks to correct the inverse square drop in intensity with distance. This is also a multiplier, which is based on the measured phase, and can be calculated after signal demodulation.
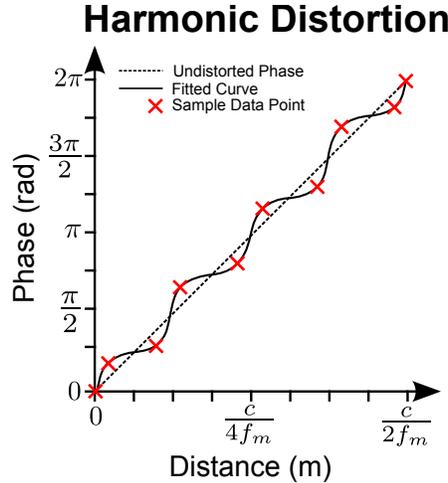
Harmonic Distortion

Figure 3: Harmonic distortion presents itself as a distance based offset, oscillating around the undistorted phase.

The function of this amplitude correction is to make the amplitude image closer to that of a typical greyscale imaging sensor, hence is not always necessary.

Finally, there is a well understood dark current offset per pixel (β_{jk}). This offset is reported by the sensor in the absence of light,[12] and is caused by small currents produced even without any incident photons on the sensor.

The phase step received at a particular pixel including all error sources considered, is given by

$$\tau_{jkn} = \eta_{jk}\sigma_{jk}\gamma_\phi A_{jk}\cos(\phi_{jk} + \theta_n + \delta_{jk} + \rho_{jk\phi}) + (B_{jk} + \beta_{jk}). \tag{8}$$

3. METHOD

The camera used is the SR4000, from MESA Imaging.[13] The SR4000 operates at a modulation frequency of 30 MHz with an illumination wavelength of 850 nm. The camera has a pixel array size of 144×176, with a field of view of $43.6° \times 34.6°$. In the camera's regular operation, factory calibrations and filters are applied by default, correcting errors and noise sources in the image while outputting only the phase change and amplitude of the scene. The camera is able to be set to 'raw mode', which allows the user to acquire the four raw phase steps with no calibrations or filters applied (raw data).

Typical calibration techniques involve imaging a planar target,[14] or a checkerboard pattern[15] to attempt to correct the image to some ground truth. One of the downsides of imaging a planar target is that time-of-flight cameras measure the radial distance to the scene, hence a transformation is required if the calibration is to be derived from a fitted plane.

The key novelty in retrieving the SR4000 calibrations is in acquiring the data. The SR4000 was set up on a precision linear translation stage (Macron Dynamics, Inc., Croydon, PA, USA). A retro-reflector of size 20 mm by 29 mm was set on the stage, and positioned by the translation stage from 0.5 m up to 3.5 m from the camera. The purpose of the retro-reflector is that it approximates a point source, and generates a near planar wave at the sensor. Because the lens cannot be removed from the camera (as it will spoil the factory calibration), a diffuser was placed directly over the camera lens to distribute the light evenly across the sensor, and prevent saturation due to the use of the retro-reflector. The experimental setup of the retro-reflector on the translation stage is illustrated in Figure 4. Images were acquired from the SR4000, both in the camera's normal calibrated operation and in 'raw mode'.

Using this setup, images were taken in increments of 0.05 m, over the 3 m, giving measurements at 60 unique distances. A total of 100 images were captured at each distance, and the mean image calculated.

The first error source found is harmonic distortion. This is found by taking the mean value at each distance, and plotting the phase versus the distance. If the translation stage were to cover the entire ambiguity range
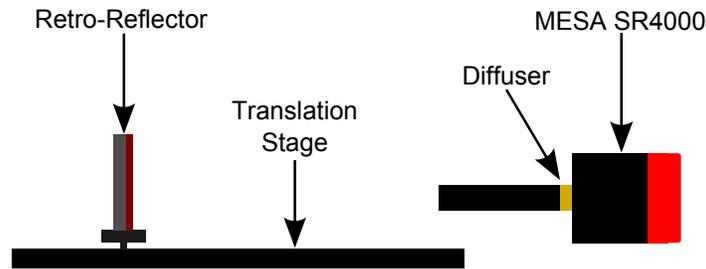
Figure 4: The setup used to capture the calibration data is illustrated in a side elevation view. The camera images a retro-reflector on a moving translation stage. A diffuser was placed in front of the sensor as close to the lens as possible, along with a small tube to reduce scattered light entering the camera.

of the SR4000 (5 m), the harmonics could be found and compensated for, but in our case images could only captured in the 0.5 m to 3.5 m range. The harmonic distortion is corrected using an interpolated look-up table, and any distances recorded outside the measured range are not able to be corrected.

The next source of error to be corrected is the global phase offset. This was found in a similar way to the harmonic distortion, simply by plotting the phase values against the distance values. This produces a plotted curve with some slope, and the offset is found as the y-intercept, as illustrated in Figure 5.
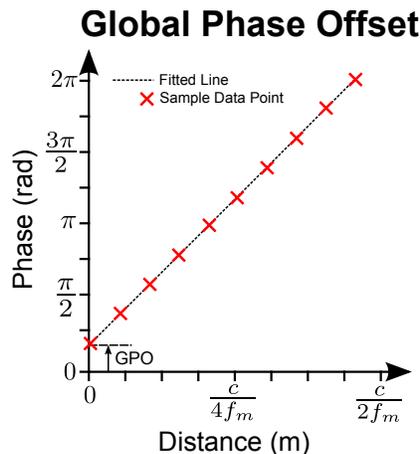

Figure 5: The global phase offset is found by comparing the phase output by the camera, to the corresponding distance. The phase axis intercept gives the offset.

The next error source is the gradual phase offset, which presents as a gradual change in phase across the image sensor. In the case of the SR4000, the pattern is almost entirely horizontal across the sensor (i.e. there is no gradual phase offset in the vertical axis). This offset is corrected by fitting a surface to the raw image data (with the harmonic distortion and global phase offsets corrected). The difference between this surface, and a surface fitted to the factory calibrated data can then be stored in a look-up table, to be subtracted from the phase data in calibration.

Once the gradual offset is corrected, the fixed pattern phase offset can be corrected. This is an additional phase offset, which is found simply with the difference between each pixel and the MESA data (as all other phase offsets should now be corrected). This offset can also be stored in a look-up table.

Hussmann et al.[12] suggest the use of a flat plane for amplitude calibration, as all pixels should receive uniform illumination. Differences in this illumination can then be calibrated out. Due to the fact that the cameras measure light radially, a flat plane could give reduced amplitude with distance from the centre of illumination, becauses the distance the light travels increases with distance from the centre. The proposed technique with a retro-reflector should result in a uniform illumination, as planar waves imaging on the sensor should be achieved. Other amplitude correction includes investigating the fact that the amplitude is non-linear over distance. One

correction for this comes from modelling the amplitude decay with distance.[16] MESA provide a calibration for both of these amplitude corrections, however they are combined into the same on-camera calibration.[17] These both multiply the amplitude image. The distance based correction from the SR4000 is a multiplier which is proportional to the square of the distance, normalised to the ambiguity distance,

$$\gamma_\phi \propto \frac{\phi^2}{d_a}. \tag{9}$$

Dividing out γ_ϕ from the the amplitude, then taking the ratio of the camera amplitude to the raw amplitude gives the radial amplitude correction caused by uneven illumination,

$$\alpha = \hat{\alpha}\gamma_\phi\sigma, \tag{10}$$

where $\hat{\alpha}$ is the raw amplitude.

The gain fixed pattern correction is then found from a uniformly illuminated sensor. If the sensor is in fact uniformly illuminated, then no other calibrations should be required (such as illumination calibration) to find the gain fixed pattern correction.

The dark current offset is usually found in the absence of light, taking an image of a scene without any illumination. With the SR4000, the authors were unable to capture an image with the light source disabled, resulting in a more difficult calibration. One option is to butt a foam (or other non-scratching) gasket to the edge of the lens, and attach a lens cap. This can be difficult with the way the SR4000 is designed, and there can still be some light leakage through to the sensor. An ideal solution would be to take an image in an infinite dark open space, where light from the SR4000 is unable to return to the sensor. Unfortunately this impossible in practice, so an alternative solution is to take an image of a large open space in darkness (i.e. there is minimal light returning from the camera's light source, and no external light from the environment is entering the sensor).

Because the translation stage being used provides precise distance measurements from the camera, a second calibration was performed. This calibration assumes that the incident light on the sensor is in fact planar, hence the calibrations (C_p) can be modelled to a flat plane using the distances of the translation stage. The same calibration steps as above were used, however instead of using data retrieved from the MESA camera, the data is calibrated from a plane formed at each distance.

The calibrations were then tested on an independent set. A flat wall was imaged from a number of angles and distances. A statistical analysis was then performed, to compare the factory calibrated SR4000 data to both the raw data and the data calibrated using the techniques described above.

4. RESULTS AND DISCUSSION

Using the techniques described above, the phase offsets are first found. The harmonic distortion of the 'raw data' is shown in Figure 6a. The global phase offset can be seen in Figure 6b, where the 'raw data' has been extrapolated to find the y-intercept. These two graphs are the real data equivalents of Figure 3 and Figure 5 respectively.

A single row of captured data is shown in Figure 7. The raw camera data is compared to the data output by the camera in its factory calibrated state, as well as the raw data after being calibrated using the techniques above. The data has been transformed to 3D coordinates, and shows the data from imaging a perpendicular plane. It can be seen that the raw data varies much more than the calibrated data, compared to the camera's factory calibrated output. This image shows only the effect of the phase calibrations, namely the harmonic correction, global phase offset, gradual phase offset, and fixed pattern phase offset.

Quantitatively, the results of the phase calibrations can be shown by taking the root mean squared error (RMSE) and standard deviation (SD), of the calibration techniques to the imaged plane. Table 1 shows the results when comparing the factory calibrated data to the raw data, the camera based calibration data, as well as the planar based calibration (C_p) mentioned above. It can be seen that the camera based calibration performs better than the planar based calibration. This is unsurprising, as the reference data is the camera's factory output, which camera based calibration is derived from.
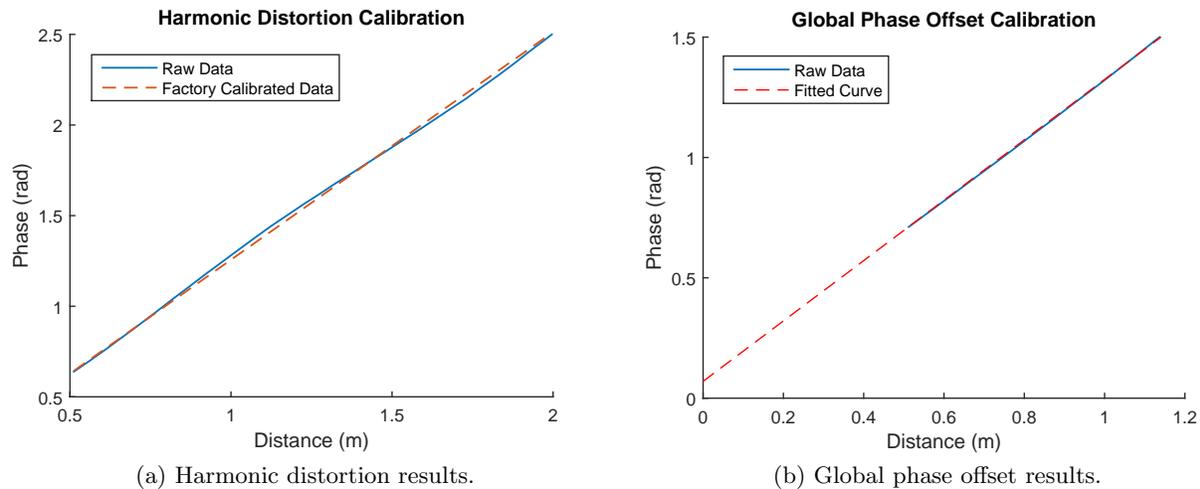
(a) Harmonic distortion results.   (b) Global phase offset results.

Figure 6: The harmonic distortion is seen where the 'raw data' oscillates around the factory calibrated data in 6a. The global offset is shown in 6b, where the 'raw data' is extrapolated, and the y-intercept of this fit is taken.
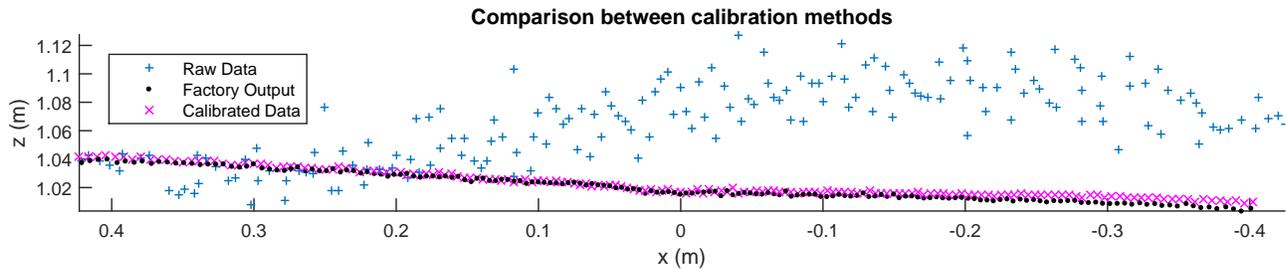

Figure 7: Comparison between raw data, factory calibrated data, and data calibrated using the presented techniques.

The factory amplitude correction in the SR4000 combines both the distance based amplitude correction (γ_ϕ), as well as the radial amplitude correction due to uneven illumination (σ_{jk}). Because the distance based amplitude correction is undesired, the two calibrations are separated. The distance based amplitude correction is undesired, because it is applied after the demodulation. In the applications of these calibrations, the raw data is calibrated before and during the demodulation, meaning a retrospective calibration is unnecessary. The radial amplitude pattern obtained in retrieving this calibration is displayed in Figure 8. It can be seen that the multiplier is approximately 1 at the centre, and increases radially from the centre.

The SR4000 doesn't appear to correct the fixed pattern gain amplitude factor, where there is variation in the amplitude response between pixels. It is very difficult to obtain this factor, as even with the retro-reflector setup, there are still effects caused by uneven illumination. Further work needs to be done to attempt to obtain

Table 1: Root mean squared error and standard deviation of phase calibrations, comparing the raw data, camera based calibrations, and distance based calibrations to the factory calibrations.

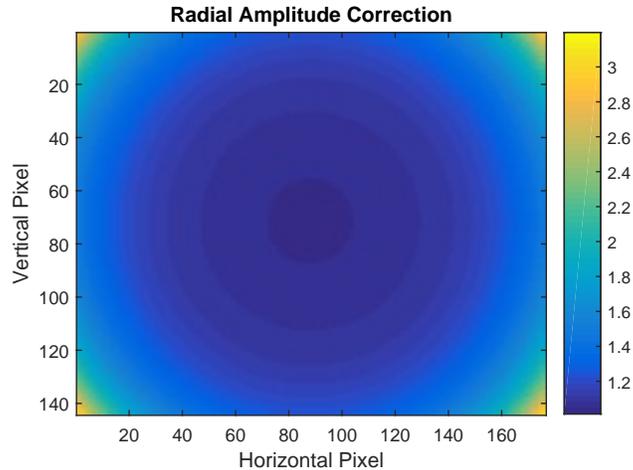| View	Raw Data		Camera Based Calibration		Planar Based Calibration	
	RMSE (m)	SD (m)	RMSE (m)	SD (m)	RMSE (m)	SD (m)
Perpendicular	0.0626	0.0376	0.0082	0.0022	0.0109	0.0029
Left/Right	0.0613	0.0384	0.0049	0.0031	0.0094	0.0033
Up/Down	0.0632	0.0378	0.0051	0.0037	0.0083	0.0031
Average	0.0624	0.0379	0.0061	0.0030	0.0095	0.0031

Figure 8: Extracted radial amplitude correction multiplier from the SR4000

a more uniform illumination on the sensor, however it appears likely to be the case that this gain fixed pattern is negligible compared to the radial amplitude correction.

Because the dark current is inherently corrected in the standard demodulation of the signal, it is unable to be extracted from the calibrated SR4000 data. This correction is only required in the motion correction algorithm, due to the modification of the demodulation technique (where the background and dark current offsets are no longer inherently corrected). This means that it is unable to be compared to any of the MESA calibrations for verification.

In future work we will apply our calibration data to motion correction algorithms. These algorithms require raw data to be captured, such that each phase step can be aligned. The calibrations are then introduced when needed, with compensation for the effects of motion on these calibrations.

The calibrations will also been applied in the Mixed Pixel Restoration Algorithm, where two frequencies are required. The SR4000 is able to operate at multiple frequencies, however only the 30 MHz frequency is calibrated. This calibration technique can be modified to be applied at different frequencies. For example, the SR4000 also operates at 15 MHz. The 30 MHz calibrated phase data can be halved, and the 15 MHz raw phase data can then be calibrated from this. The results of this calibration are not shown, as there is no noticeable difference to the 30 MHz calibration other than the phase values being halved.

5. CONCLUSION

The calibration set extracted from the MESA SR4000 in this paper has shown large improvements over using only the data acquired in 'raw mode', especially in the phase calibrations. Amplitude and dark current calibrations have also successfully been obtained. These calibrations have been separated in such a way that a user is able to apply specific calibrations to the 'raw data' on demand.

There are still other calibrations that would help to improve the extracted calibration set. There is a variation with increasing temperature in the camera, which has only been briefly investigated. It appears that there is a compensation in the SR4000 for this temperature variation. This effect would be relatively simple to extract, by modelling the compensation from the SR4000 with changing temperature. Additionally, there is the consideration of integration time in the camera. The integration time can have an effect on measured distance,[18] hence should be considered unless the integration time remains the same between the calibrations and the acquisition of practical data.

Future work will look to investigate and implement the effect of the temperature calibration and integration time, providing a more robust calibration. This work will also be applied to upcoming algorithms.

ACKNOWLEDGMENTS

REFERENCES

[1] 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," *Proceedings of the 1st Range Imaging Research Day* , 21–32 (2005).

[2] Dorrington, A. A., Kelly, C. B. D., McClure, S. H., Payne, A. D., and Cree, M. J., "Advantages of 3D time-of-flight range imaging cameras in machine vision applications," in [*16th New Zealand Electronics Conference (ENZCon'09)*], 95–99 (2009).

[3] Dorrington, A. A., Godbaz, J. P., Cree, M. J., Payne, A. D., and Streeter, L. V., "Separating true range measurements from multi-path and scattering interference in commercial range cameras," in [*Three-Dimensional Imaging, Interaction, and Measurement*], *Proc. SPIE* **7864**, 786404 (2011).

[4] Chiabrando, F., Piatti, D., and Rinaudo, F., "SR-4000 ToF camera: further experimental tests and first applications to metric surveys," *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences* **38**(5), 149–154 (2010).

[5] Piatti, D. and Rinaudo, F., "SR-4000^{TM}and CamCube3.0 time of flight (ToF) cameras: Tests and comparison," *Remote Sensing* **4**(4), 1069–1089 (2012).

[6] Cree, M. J., Streeter, L. V., Conroy, R. M., and Dorrington, A. A., "Analysis of the softkinetic depthsense for range imaging," in [*Image Analysis and Recognition*], 668–675, Springer (2013).

[7] Lange, R. and Seitz, P., "Solid-state time-of-flight range camera," *IEEE Journal of Quantum Electronics* **37**(3), 390–397 (2001).

[8] Streeter, L. and Dorrington, A. A., "Coded exposure correction of transverse motion in full-field range imaging," *Optical Engineering* **53**(10), 102109 (2014).

[9] Fuchs, S. and Hirzinger, G., "Extrinsic and depth calibration of ToF-cameras," in [*IEEE Conference on Computer Vision and Pattern Recognition, 2008. CVPR 2008.*], 1–6, IEEE (2008).

[10] Foix, S., Alenya, G., and Torras, C., "Lock-in time-of-flight (ToF) cameras: a survey," *IEEE Sensors Journal* **11**(9), 1917–1926 (2011).

[11] Payne, A. D., Dorrington, A. A., Cree, M. J., and Carnegie, D. A., "Improved linearity using harmonic error rejection in a full-field range imaging system," in [*3D Image Capture and Applications VII*], *Proc. SPIE* **6805**, 68050D (2008).

[12] Hussmann, S. and Edeler, T., "Pseudo-four-phase-shift algorithm for performance enhancement of 3D-TOF vision systems," *IEEE Transactions on Instrumentation and Measurement* **59**(5), 1175–1181 (2010).

[13] MESA Imaging AG, [*SR4000 Data Sheet*].

[14] Kahlmann, T., Remondino, F., and Ingensand, H., "Calibration for increased accuracy of the range imaging camera swissranger^{TM}," *Image Engineering and Vision Metrology (IEVM)* **36**(3), 136–141 (2006).

[15] Abdo, N. and Borgeat, A., "3D camera calibration," *Freiburg, Germany: Albert-Ludwigs-University Freiburg* (2010).

[16] Stürmer, M., Penne, J., and Hornegger, J., "Standardization of intensity-values acquired by time-of-flight-cameras," in [*IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops, 2008. CVPRW'08*], 1–6, IEEE (2008).

[17] MESA Imaging AG, [*SR4000/SR4500 User Manual*], vol. 3.

[18] May, S., Werner, B., Surmann, H., and Pervolz, K., "3D time-of-flight cameras for mobile robotics," in [*2006 IEEE/RSJ International Conference on Intelligent Robots and Systems*], 790–795, IEEE (2006).