Multiple Return Separation for a Full-Field Ranger Via Continuous Waveform Modelling

John P. Godbaz, Michael J. Cree, Adrian A. Dorrington

Department of Engineering, University of Waikato, Hamilton, New Zealand

ABSTRACT

We present two novel Poisson noise Maximum Likelihood based methods for identifying the individual returns within mixed pixels for Amplitude Modulated Continuous Wave rangers. These methods use the convolutional relationship between signal returns and the recorded data to determine the number, range and intensity of returns within a pixel. One method relies on a continuous piecewise truncated-triangle model for the beat waveform and the other on linear interpolation between translated versions of a sampled waveform. In the single return case both methods provide an improvement in ranging precision over standard Fourier transform based methods and a decrease in overall error in almost every case. We find that it is possible to discriminate between two light sources within a pixel, but local minima and scattered light have a significant impact on ranging precision. Discrimination of two returns requires the ability to take samples at less than 90 degree phase shifts.

Keywords: Lidar, Range Imaging, AMCW, Mixed Pixels, Maximum Likelihood, Sparse Spike Train Deconvolution

1. INTRODUCTION

Amplitude Modulated Continuous Wave (AMCW) rangers work on the time-of-flight principle, measuring the range dependent phase shift introduced in an amplitude modulated light beam as it traverses from the ranger to the scene and back. Full-field range-imagers simultaneously acquire range for each pixel across a 2D image. Generally both the light and sensor are modulated with a cyclic rectangular function. Homodyne systems use the same modulation frequency for both light and sensor and take measurements while varying the phase of the sensor modulation. The measurements are samples of the correlated modulation waveforms and can be Fourier transformed in order to calculate the phase shift due to the time-of-flight, hence range to the scene. Heterodyne systems, in contrast, modulate the sensor at a slightly different frequency to the light source, thus heterodyning the two signals to produce a low frequency beat waveform with the same phase as the light return. This allows the waveform to be sampled at a low frequency and analysed using the same Fourier method. Heterodyne methods have been found to improve phase linearity over homodyne methods due to a reduction in aliased harmonics.[1]

Mixed pixels are caused by the integration of light from objects at different distances by a single pixel. Fourier analysis of mixed pixels results in erroneous ranges that can manifest as intermediate values between received signals or as completely anomalous values that wrap around due to the modulo 2π effect of phase offsets. AMCW full-field rangers are particularly susceptible to the mixed pixel problem due to depth-of-focus limitations. Previous work on mixed pixels has been largely limited to identification and removal of these points from images.[2–5] As more full-field AMCW rangers such as SwissRanger[6] and CanestaVision[7] become available, mixed pixels are seen as more problematic. Pulsed systems handle multiple returns well, but lose ranging precision versus an AMCW system because they integrate so much less light. In a previous paper[8] we proposed that a single pixel over time can be regarded as a convolution of an impulse response with the signal returns. We used an off-the-shelf algorithm, the Levy-Fullagar algorithm,[9] to deconvolve the returns – however, it was necessary to average the data over many beat cycles in order to achieve good results. Here we develop methods that require only a single beat cycle. Other systems have used non-sinusoidal modulation but have not utilised the harmonic content of the beat waveform to find multiple returns.[10–12]

Recovering the individual signal returns can be considered a sparse spike train deconvolution problem, commonly encountered in Geophysics when processing seismic traces. Previous methods do not directly produce

---

Contact Details: J.P.G.: E-mail: jpg7@waikato.ac.nz

continuous phase values, limiting range precision. In this paper we propose two methods that produce continuous phase values, thus continuous range values. The first uses a parametric truncated-triangle model for the impulse response and the second uses linear interpolation between translated versions of a sampled waveform. The models are then fit to the data in order to determine the intensity and phase of the signal returns. Advantages of these new methods include handling multiple returns but retaining the ranging precision of an AMCW system in the single return case – in fact, using these models improves ranging precision over the standard Fourier transform based method in the single return case. The primary requirement for implementing a system that uses these or similar methods is harmonic content in the beat waveform (must be a truncated-triangle for the first method) and the ability to take samples at less than 90 degree phase shifts.

We have ignored background light levels for simplicity here. It should be possible to extend the methods here by adding an extra parameter to the model, but this has not yet been attempted.

2. THEORY

We explain the standard sparse-spike train deconvolution model, the two models used for the beat waveform and then the Poisson noise Maximum Likelihood optimisation problem. In the following, square brackets, [], are used for discrete valued parameters and parentheses, (), for continuously valued parameters.

2.1 Spare Spike Train Deconvolution

The problem of determining the intensity and phase of returns within a pixel can be considered a type of Sparse Spike Train (SST) deconvolution problem. This means that it is a deconvolution problem in which the vector describing the data mostly contains zero values. SST deconvolution methods use a 1 dimensional model along the lines of

$$k = H \cdot s + \epsilon \tag{1}$$

Where k is the recorded data, H is a Toeplitz matrix, s is a sparse vector and ϵ is noise. Inversion methods include regularisation and $L1$ norm minimisation[9] – other methods[13–15] explicitly detect spikes and then determine their amplitude using a search algorithm, often using an assumption of a Bernoulli-Gaussian distribution. Some of these methods tend to spread a signal across two adjacent phase bins if the underlying phase is intermediate, however they do not identify that these are coming from the same return – they are limited to a discrete model. Our first method models the data as a continuously variable piecewise function, while the second method implicitly solves eqn. 1 by assuming that each individual return is spread across two adjacent phase bins and then using this to determine the continuously valued phase.

Rather than detecting the spike location via a global or local search algorithm, which can be slow, the algorithms are structured as direct optimisation problems making an assumption of a certain number of returns. The best results are explicitly found for both the single and dual return cases, and it is determined later which is the most likely. Two different models are used for the beat waveform/impulse response, each of which takes intensity and phase of each return as an argument. The intensity and phase values are then determined via a Poisson maximum-likelihood optimisation.

2.2 Fourier Bin Method

For comparison purposes we use the common Fourier bin method in addition to our own methods. This calculates the range to an object by taking the Fourier transform of the beat waveform fundamental frequency. Since the phase of the waveform is proportional to range to the object, this enables ranges to be measured with as few as three samples per beat.

2.3 Truncated-Triangle Wave Model

Our first new method is very different from standard SST deconvolution methods because it uses a continuously variable function to model the impulse response, which allows the underlying structure of the problem to be
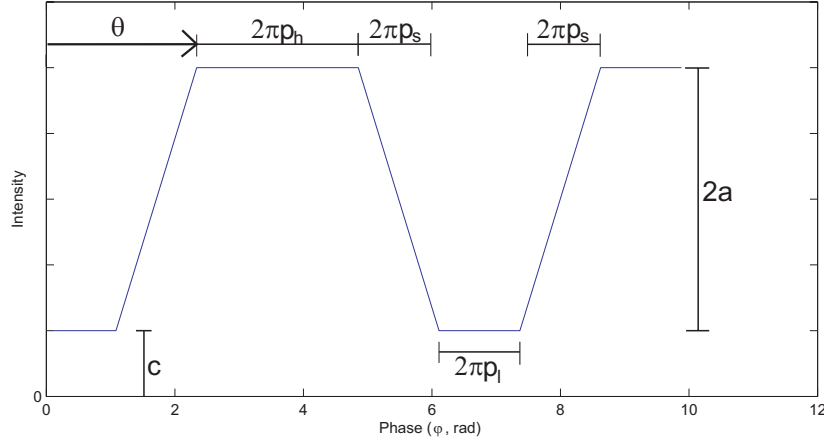
Figure 1. The parameters of a truncated triangle waveform.

better represented. We assume that both the laser waveform, $l(\theta_l)$ and the intensifier response waveform, $r(\theta_r)$ are cyclic rectangular waveforms described by the equations

$$l(\theta_l) = \begin{cases} \rho, & \frac{\theta_l}{2\pi} \bmod 1 < h_l \\ 0, & \frac{\theta_l}{2\pi} \bmod 1 \geq h_l \end{cases} \tag{2}$$

$$r(\theta_r) = \begin{cases} 1, & \frac{\theta_r}{2\pi} \bmod 1 < h_r \\ 0, & \frac{\theta_r}{2\pi} \bmod 1 \geq h_r \end{cases} \tag{3}$$

Where h_l and h_r are the duty cycles of the laser and intensifier waveforms respectively. The resulting waveform in the single return case is a truncated triangle wave as shown in fig. 1 formed via a correlation operation

$$m_t(\theta, \phi) = \int_0^{2\pi} l(2\pi - \theta + \psi) r(\phi + \psi) d\psi \tag{4}$$

where θ and ρ are the phase offset and intensity of the return and ϕ is the phase we are measuring at. This waveform shape, ignoring phase and intensity, can be uniquely described by the tuple $(p_h, p_l, \frac{c}{a})$ which can be calculated directly from the duty cycles of the laser and intensifier waveforms or fit to emperical data. p_h and p_l represent the proportion of time spent at peak and trough intensity respectively, $\frac{c}{a}$ represents the relative size of any constant offset to the amplitude of the waveform (see fig. 1).

Using knowledge of the duty cycle of the illumination and the intensifier shutter, we can now calculate the parameters that describe the shape of the beat waveform, ultimately modelling the impulse response with a continuous piecewise function. For simplicity, take the function $f(\theta, \phi)$, which produces normalised relative phase offsets in the range $[0, 1)$ for $0 \leq \theta < 2\pi$ and $0 \leq \phi < 2\pi$

$$f(\theta, \phi) = \left(1 + \frac{\phi - \theta}{2\pi}\right) \bmod 1 \tag{5}$$

ϕ represents the phase at which the measurement is being taken and θ represents the phase offset of the signal return and thus range to an object. Note the sign of θ – objects at greater distances have signals that are more delayed in time. For simplicity we assume that $\rho = 1$. Firstly we calculate the waveform peak intensity, b, and the period of peak intensity, p_h. The waveform peak intensity is the maximal value of $m_t(\theta, \phi)$ for a particular pixel over time, where $b = 2a + c$.

$$b = 2\pi \min h_l, h_r \tag{6}$$

$$p_h = |h_l - h_r| \tag{7}$$

The phase shift taken for the waveform to descend from an intensity of b to zero is defined as b because the slope is always $\rho \ (= 1)$. While the waveform does not necessarily reach zero, knowledge of the phase shift required

in order to achieve this allows us to determine if there is any constant offset c. r_h is the combined phase shift required for full ascent from zero intensity to a, peak intensity, and descent to zero again. This defines a special case, because if $r_h > 2\pi$ then the impulse response never takes a value of zero.

$$r_h = 2\pi p_h + 2b \tag{8}$$

We can now calculate the proportional size of the flat lower region of the truncated triangle wave, if any

$$p_l = \begin{cases} 0, & r_h > 2\pi \\ 1 - \frac{r_h}{2\pi}, & r_h \leq 2\pi \end{cases} \tag{9}$$

If the value of r_h is too high, then there is a constant offset c to the waveform

$$c = \begin{cases} \frac{r_h - 2\pi}{2}, & r_h > 2\pi \\ 0, & r_h \leq 2\pi. \end{cases} \tag{10}$$

With knowledge of p_l we can now calculate p_s, the actual proportion of sloped ascent or descent region

$$p_s = \frac{1 - p_h - p_l}{2} \tag{11}$$

$$a = \frac{b - c}{2} \tag{12}$$

a represents the amplitude of the waveform itself, and enables us to calculate the last of the defining parameters $\frac{c}{a}$, which provides the relative size of any constant offset.

Ignoring linear coefficients, the truncated-triangle wave impulse response can be modelled in the homodyne case as

$$m_t(\theta, \phi) = \frac{c}{b} + \begin{cases} 1, & f(\theta, \phi) < p_h \\ 1 - \frac{f(\theta,\phi) - p_h}{p_s}, & p_h \leq f(\theta, \phi) < p_h + p_s \\ 0, & p_h + p_s \leq f(\theta, \phi) < (1 - p_s) \\ \frac{f(\theta,\phi) - (1 - p_s)}{p_s}, & (1 - p_s) \leq f(\theta, \phi) < 1 \end{cases} \tag{13}$$

which equates to eqn. 4. This can be extended to the heterodyne case by an integration of the form

$$m_h[u](\theta) = \int_{u/N}^{u/N+1} m_t(\theta, \tau) d\tau \tag{14}$$

where $\{u \in \mathbb{Z}; 0 \leq u < N\}$ is one of N discrete samples.

2.4 Base-Ratio Waveform Model

The second model of the impulse response is based on an empirical sampled waveform, $\psi_w[v]$, of length N and uses linear interpolation between the waveform and a version shifted by a single sample to generate intermediate phase values

$$g[v](\alpha) = (1 - \alpha)\psi_w[v] + \alpha\psi_w[(v + 1) \bmod N] \tag{15}$$

where $g[v](\alpha)$ is the linear interpolation function where $\{v \in \mathbb{Z}; 0 \leq v < N\}$ is the coarse location of the return and α is the interpolation weighting or fine location. This method could be considered to be similar to the standard SST deconvolution formulation but assuming that there is one spike split across two cells, and not explicitly using a matrix formulation. We attempt to directly optimise for a combined θ:

$$m_w[u](\theta) = g[(N + u - \text{floor}(\frac{N\theta}{2\pi})) \bmod N](1 - (\frac{N\theta}{2\pi} \bmod 1)) \tag{16}$$

where $m_w[u](\theta)$ is the optimisation model, $\{u \in \mathbb{Z}; 0 < u < N\}$ is the current sample and θ is the phase offset of the return. An alternate representation could use a piecewise function with N segments.

2.5 Modelling Multiple Returns/Poisson Maximum Likelihood

In order to model a particular pixel, the models for each return within that pixel are summed together. A pixel sampled at u at image location $(x, y) \in \mathbb{R}^2$, with N_r returns, of phase offsets Θ_i and intensity I_i takes the value

$$m_{x,y}[u](\Theta, I) = \sum_{i=0}^{N_r-1} I_i \cdot m_x[u](\Theta_i) \tag{17}$$

for a single return model $m_x[u](\theta)$ of some sort. Θ and I are vectors of length N_r which are optimised numerically for a likelihood function. In the case of the University of Waikato Range-Imager, the primary noise source is shot noise so we use a Poisson log-likelihood function rather than a Gaussian/least squares approach. There are other noise sources such as thermal noise, readout noise and jitter noise, however modelling them would substantially increase computational complexity.

In the case of Poisson distributed noise, the log likelihood is described by

$$LL(\lambda) = -n\lambda + \log \lambda \sum_{i=0}^{n-1} k_i - \sum_{i=0}^{n-1} \log k_i!, \tag{18}$$

since the recorded data are not free parameters, the final term can be ignored. We maximise eqn. 18 via eqn. 17 using Matlab 7.3's Sequential Quadratic Programming method for solving non-linear equations using numerically calculated gradients.

3. EXPERIMENTAL METHODOLOGY

These experiments were carried out on the University of Waikato full-field heterodyne ranger described in[16] using a laser modulation frequency of 12MHz. The system uses modulated laser illumination and a modulated image intensifier as a high-speed shutter on the camera. The camera was run at 32 frames-per-second – thus at 4 frames-per-beat, each beat was 1/8th of a second long, at 8 frames-per-beat, 1/4 second etc.

The ranger system suffers from zeroed frames, which are frames with very low, near zero intensity values as if not exposed to any light caused by a camera firmware issue. This non-deterministic problem was dealt with by detecting the zeroed frames and dropping in a replacement from the same point in the beat cycle from an extra series of backup frames captured at the end of each experimental capture. It is possible that laser intensity drift over time may have negatively affected results when this replacement occurred.

3.1 Waveform Fitting

In order to use the truncated-triangle waveform model for deconvolution/phase identification it is necessary to determine the three shape parameters (p_h, p_l, $\frac{c}{a}$). In order to do this, a small region was averaged over a number of beat cycles to give the general shape of the waveform. Candidate pixels were required to be sufficiently bright to limit noise and the influence of scattered light, fairly close to the optical axis and not near the edge of an object. The parameters were then found via optimisation using numerical methods – this was repeated on an individual image basis for each experiment. The waveform ψ_w was produced in a similar manner. Additional analysis was also performed on the sampled waveform in order to extract the intensifier shutter response – this helps to explain deviations from the ideal truncated-triangle model.

Measurements produced by the truncated-triangle and base-ratio methods need to be able to be converted to the same form as Fourier bin data – this can be achieved by multiplying by a calibrated complex number. A synthesised truncated-triangle waveform with an intensity of 1 and a phase of 0 was Fourier transformed and the value of the fundamental frequency bin was recorded. This was repeated for the base-ratio method in order to produce conversion factors for both methods.
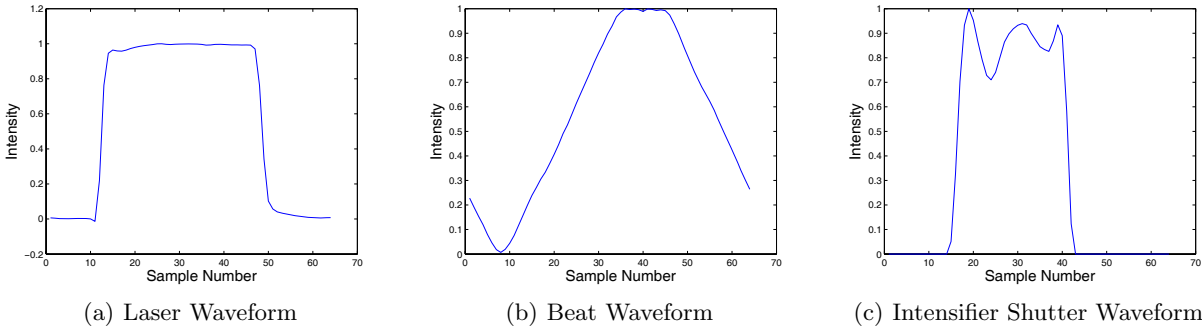
Figure 2. Using knowledge of the laser and beat waveforms, the intensifier shutter response was recovered using 30 iterations of the Richardson-Lucy deconvolution algorithm. It shows ringing effects due to the high modulation frequency.

3.2 Single Return Analysis

Single return precision was calculated by placing a flat board in front of the ranger and taking a series of range images. It was assumed for analysis purposes that each pixel contained only a single return. Each individual beat was considered to be a separate measurement. The methods under test were applied to the data over 10 beat cycles and a 50x50 pixel region of the board and the interbeat precision was estimated using the sample standard deviation. An estimate of the ideal range data was created by averaging the Fourier bin method data over time and then blurring it using a Gaussian blur of radius 5 pixels – the overall ranging error for a dataset was then estimated by measuring the Root-Mean-Square error against the ideal model. This allows error to be estimated without complex calibration for effects such as irising. In the single return case, the Fourier bin values were used as seed values for the optimisation.

3.3 Dual Return Analysis

A scene was created where two boards were placed in front of each other so that they overlapped, the camera was then defocussed to produce a blurred region between the two boards and a large number of mixed pixels to work with. 10 beat cycles were analysed in a similar manner to the single return case, except that for analysis purposes it was assumed that every pixel within this region contained two returns. In order to calculate the two return precision the source of each return was determined by a brute force search designed to minimise error – it was assumed that one source within each pixel belonged to the board in front, and one to the board in back. In order to identify the nature of any systematic error, the mean distance between returns was measured and compared to the actual mean distance between the two objects.

4. RESULTS AND DISCUSSION

4.1 Intensifier Response Waveform

Payne[17] sampled the intensifier shutter response waveform of the ranger using a 266 pS pulsed laser diode synchronised to the ranger waveform synthesiser. Instead, by sampling the laser waveform using a photodiode we can use the correlation relationship expressed in eqn. 4 to allow us to calculate the response without such specialised hardware. The laser waveform at a 56.8% duty cycle is shown in fig. 2(a), and the resultant beat waveform is shown in 2(b). The Richardson-Lucy algorithm[18, 19] is a type of Expectation-Maximisation deconvolution algorithm designed for deconvolution with positivity constraints in the case of Poisson distributed noise. We applied 30 iterations of the algorithm to estimate the shutter response shown in fig. 2(c), more iterations resulted in excessive sensitivity to noise while less were unable to fully resolve the shape. The waveform shows similar ringing effects to those reported by Payne, but with reduced magnitude due to use of a different modulation frequency and mitigation efforts.
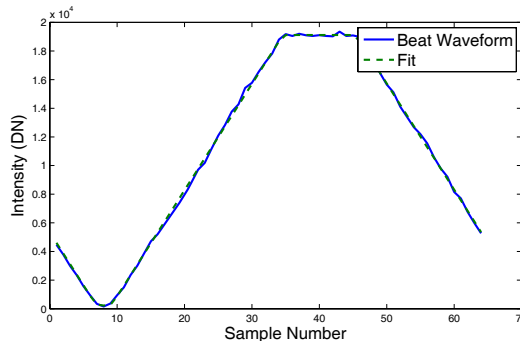
The intensifier shutter response waveform explains why the beat waveform is not a perfect truncated-triangle wave. The ringing in the intensifier waveform results in small oscillations in the beat waveform. However, because the waveform is close enough to a cyclic rectangular function we are able to model the output waveform as a

Table 1. Example fitted truncated-triangle model parameters. These are the parameters of the models shown in fig. 3

	Shape			Scaling/Translation	
Frames-Per-Beat	p_h	p_l	$\frac{c}{a}$	θ (rad)	a (DN)
4	0.1459	0.02842	0	0.7626	16210
64	0.1692	0.03477	0.02398	3.3485	9445
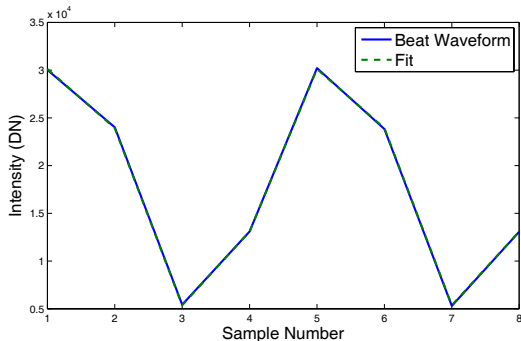
(a) 4 Frames-Per-Beat

(b) 64 Frames-Per-Beat

Figure 3. The results of fitting a truncated triangle model to the beat waveform while varying the number of frames per beat.

truncated-triangle wave. We have found that the intensifier response at certain other frequencies (for example 10MHz) can break the truncated-triangle based model.

4.2 Waveform Models

Table 1 shows the parameters found in the two cases shown in fig. 3. It was not possible to use the base-ratio method at 4 frames-per-beat because the sampled waveforms were not representative enough – typically they contained a significant constant offset. Attempting to apply the base-ratio method resulted in non-continuous range values and discrete levels in the range data – the same errors occur if the truncated-triangle parameters are fit to a poorly chosen region of the image where there is a large constant offset due to multiple returns. When the parameters are properly fit, the range measurements are approximately normally distributed.

4.3 Single Return Phase Measurement

The results from this experiment are presented in table 2 and fig. 4. The precision of the Fourier method decreases after 16 frames-per-beat – this is extremely surprising because it indicates that the extra light being integrated is not actually improving the signal to noise ratio at all, rather it appears to be adding noise. The most probable cause of this is the Direct Digital Synthesiser (DDS) which is used to produce our modulation signals. The DDS runs at 400MHz and uses an accumulator/lookup-table system to produce a roughly sinusoidal waveform which is converted to a rectangular wave via a comparator. Despite the insertion of a low-pass filter into the signal chain, jitter/instability in the rectangular wave output contributes noise.

Table 2. Interbeat precision and rms error against a reference model of phase measurements for different measurement methods in the single return case. The reference model was generated by taking the mean of all Fourier bin phase measurements over time and then applying a Gaussian blur with a blur radius of 5 pixels. The source data was a range image of a flat board and only one optimisation attempt was used. Fpb = frames-per-beat

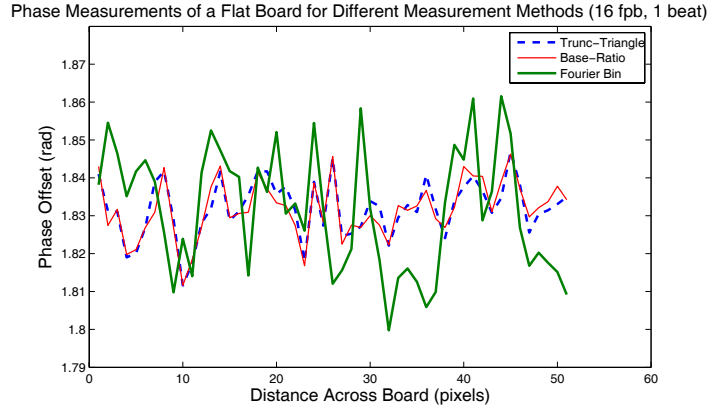|	Fourier-Bin		Trunc-Triangle		Base-Ratio	
Fpb	Precision (1σ)	RMS Error	Precision (1σ)	RMS Error	Precision (1σ)	RMS Error
4	0.0094	0.0193	0.0083	0.0348	N/A	N/A
8	0.0062	0.0163	0.0045	0.0119	0.0051	0.0117
16	0.0062	0.0168	0.0036	0.0086	0.0048	0.0124
32	0.0086	0.0138	0.0046	0.0093	0.0055	0.0086
64	0.0111	0.0123	0.0061	0.0086	0.0062	0.0075

Figure 4. A slice through a range image of a flat board. The base-ratio and trunc-triangle methods appear to substantially decrease random noise relative to the standard Fourier bin based method. The two shape based methods are highly correlated.

Table 3. The interbeat, 1σ precision of phase measurements for different measurement methods in the two returns case. Fpb = frames-per-beat

Fpb	One Optimisation Attempt		Best of Five Optimisation Attempts	
	Trunc-Triangle (rad)	Base-Ratio (rad)	Trunc-Triangle (rad)	Base-Ratio (rad)
8	0.2457	0.3116	0.2137	0.1889
16	0.2715	0.4584	0.1842	0.1657
32	0.2404	0.4248	0.1425	0.0341
64	0.2849	0.3731	0.0225	0.0218

It appears that under most circumstances the non-Fourier algorithms produce phase measurements with better precision and less overall error. An exception to this is the 4 frame-per-beat truncated-triangle case where the precision is quite good, but the overall error is quite high – suggesting that there might be a systematic error that needs further investigation. One possibility is that the calibration to convert to the same scale/phase as the Fourier bin data produced an incorrect phase due to aliasing, one option might be to calibrate using a higher number of frames-per-beat to avoid this. An optimal trade-off between precision and number of frames required appears to be around 8 frames-per-beat with the truncated-triangle method being slightly better than the base-ratio algorithm. One possibility is to average a smaller number of frames-per-beat over several beat cycles – this appears to be the only way to improve results after 16 frames-per-beat. A better base-ratio algorithm could trade-off complexity and precision, perhaps oversampling the waveform so that ψ_w is better representative of the underlying continuously variable waveform shape. One limitation of this analysis is that it does not take into account effects caused by linearity errors – in the future more extensive analysis ought to be completed over a wider range of phase offsets.

4.4 Dual Return Phase Measurement

The results are shown in tables 3 and 4 and graphical examples are shown in fig. 5. We show the optimisation based methods after either one optimisation attempt or five – this shows the large impact that local minima

Table 4. Mean estimated distance between returns in the two return case. Fpb = frames-per-beat

Fpb	Best of Five Optimisation Attempts	
	Trunc-Triangle (rad)	Base-Ratio (rad)
8	1.412	1.132
16	1.436	1.285
32	1.373	1.298
64	1.279	1.298
Actual	1.131	

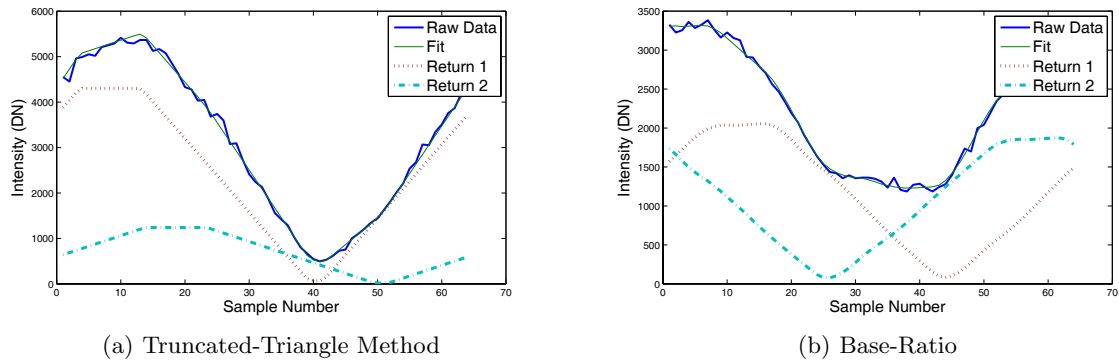(a) Truncated-Triangle Method   (b) Base-Ratio

Figure 5. Results from two different methods at 64 fpb.

have on the resulting phase precision. The truncated-triangle and base-ratio methods both seem to perform reasonably well with multiple optimisation attempts and with a high number of frames-per-beat, however in general the precision is not very good. More detailed analysis of the raw data indicates that the majority of the measurements have a precision of around 0.02 radians, however there is a significant population of outliers which distort the metric. A good example is shown in fig. 7(d) and 7(h) where multiple optimisation attempts have a huge impact on local minima. One possible solution to this may be to change from a 1D model to a 3D model and use spatial information to reduce the impact of the outliers. One major limitation with this experiment is that the precision is being calculated over a large area with a wide variety of intensity ratios (near object intensity/far object intensity) on which we found that the range error depends.[8]

Like other systems[20] the ranger suffers from issues with scattered light, particularly in dark regions – this ends up being a major source of subtly mixed pixels. This is contributed to by the use of a relay lens to couple the camera to the image intensifier. For example, in the case of fig. 6 the mean light intensity across the entire image after calibration is 5.8×10^3 DN and the peak intensity is 4.5×10^4 DN. In theory, the regions outside the image intensifier should be completely black, but because of stray light the mean intensity in the corner regions of the image is 1.5×10^2 DN. If scattered light is integrated by a pixel it tends to result in an extra constant offset. Table 4 shows that there is a systematic error resulting in returns appearing farther apart than they actually are – it is probable that this is contributed to by stray light. One method for testing the ranger system in the two return case without scattering effects may be direct illumination of the image intensifier with several diffuse light sources at different ranges rather than using a physical scene. Another alternative might be a simulation based approach to determine if there are any other possible contributing factors to a systematic error. Future work on characterisation of the scattering might allow the effect to be predicted and subtracted/deconvolved from the range data.

4.5 Estimating the Number of Returns

An example of the data output by the section 2 phase determination methods is in fig. 6. The challenge is to now identify the actual number of returns at each pixel. Note that the primary return/first return is the brightest return.

In an ideal situation there would be a simple local prior distribution describing the probability of a specific number of returns occurring within any pixel which could be incorporated into the optimisation process. For example, many sparse spike train deconvolution methods use a Bernoulli-Gaussian distribution as a prior – however this does not necessarily have any relevance to an actual physical scene. Rather than use any specific prior, we instead use a simple method for global optimisation by prioritising returns which allows the user or possibly an advanced computer programme to determine the value of a single controlling parameter C_l.

The log-likelihood calculated as a part of the optimisation can then be used to determine the quality of a transition – a change in state produced by incrementing the number of returns assumed to be within a pixel. The quality of the transition is defined by ΔLL_{2-1}, which is the change in log-likelihood caused by incrementing the assumed number of returns from one to two. The parameter, C_l, is the number of iterations that a greedy
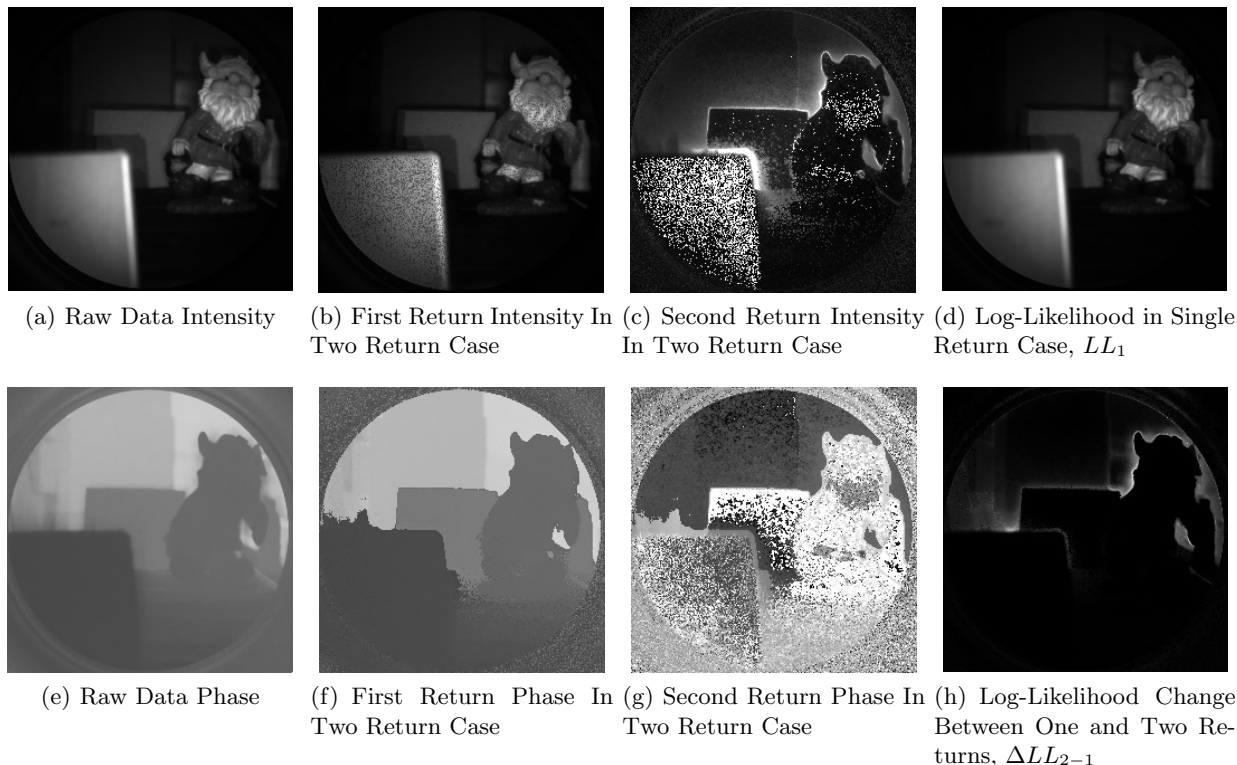
(a) Raw Data Intensity   (b) First Return Intensity In Two Return Case   (c) Second Return Intensity In Two Return Case   (d) Log-Likelihood in Single Return Case, LL_1

(e) Raw Data Phase   (f) First Return Phase In Two Return Case   (g) Second Return Phase In Two Return Case   (h) Log-Likelihood Change Between One and Two Returns, ΔLL_{2-1}

Figure 6. Scene 1. For 6(h) black represents values below zero, white the greatest positive ΔLL_{2-1}. Created using the base-ratio method and 64 frames-per-beat. None of the intensity images are to the same scale and 6(c) is deliberately saturated so as to bring out fine detail in the background. Note that the noisy regions in 6(c) mostly have a very low ΔLL_{2-1} in 6(h). We have not provided the intensity and phase results for the assumed single return case because they are almost exactly the same as the raw images. Sharpening of the edges around objects is visible in 6(f) – the soft edges were due to mixed pixels. A close-up example of the recovered image is provided in 7 and the full recovered image in 8.

algorithm is allowed to execute. The algorithm starts by assuming that every pixel has a single return, and then each iteration finds the transition for which ΔLL_{2-1} is maximal and executes it. In this particular case the maximum number of returns is limited to two, however there is no intrinsic reason why the algorithm could not use a much larger number, or even start assuming no returns. A high ΔLL_{2-1} indicates that two returns fits the data better than one return. Given adequate fitting time, two returns should always be able to fit data at least as well as one return – however, we have found that a high ΔLL_{2-1} tends to indicate that LL_1 is a particularly poor fit and contained multiple returns.

An example of the effect of C_l on a subregion of a range image is given in 7, the figure also demonstrates the effect of local minima. Notice that the C_l limiting was applied as a global optimisation, and this figure is just a subregion. A recovered version of 6 is given in fig. 8.

5. CONCLUSION

We presented two novel methods for calculating the phase and intensity of returns within a pixel for an AMCW ranger system. Both methods were found to improve ranging precision versus the standard Fourier bin method in all cases and decrease overall ranging error in most cases. Although the algorithms were able to separate out multiple returns within a pixel, the precision was fairly low (0.2 radians) and the measurements suffered from systematic error, most likely contributed to by scattered light. Future improvements to mitigate this might include scattered light subtraction/deconvolution methods and the imposition of spatial smoothness constraints to help limit the effects of local minima. Characterisation of dual return separation performance across a wider range of relative phase offsets might help better understand the mechanisms behind the poor results.
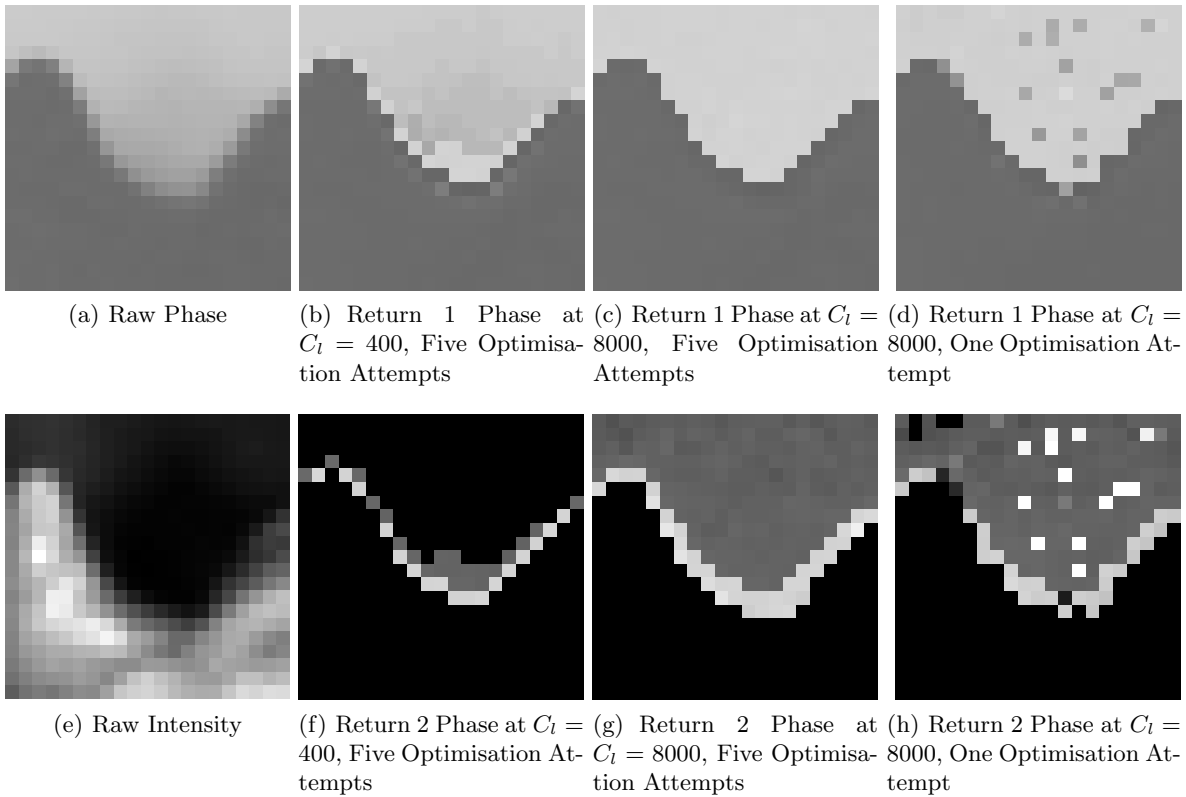
(a) Raw Phase (b) Return 1 Phase at $C_l = 400$, Five Optimisation Attempts (c) Return 1 Phase at $C_l = 8000$, Five Optimisation Attempts (d) Return 1 Phase at $C_l = 8000$, One Optimisation Attempt

(e) Raw Intensity (f) Return 2 Phase at $C_l = 400$, Five Optimisation Attempts (g) Return 2 Phase at $C_l = 8000$, Five Optimisation Attempts (h) Return 2 Phase at $C_l = 8000$, One Optimisation Attempt

Figure 7. An example of the effect that C_l has on an image using a subregion of scene 1 (from fig. 6). Notice that the region within which multiple returns have been identified is very dark and thus much more susceptible to multiple returns from scattered light within the ranger. At the same time, the lack of intensity makes it harder to identify the range to the source of each return. Light from the object in front has been scattered onto the object at the back. The noise that is present in the images produced using only one optimisation attempt is due to local minima. In the $C_l = 400$ case only the clearly mixed pixels around the borders between objects are identified and separated out. As C_l increases, less obviously mixed pixels containing scattered light are identified – eventually, if C_l is set too high, incorrect data is included in the recovered image.

(a) First Return Intensity (b) First Return Phase (c) Second Return Intensity (d) Second Return Phase
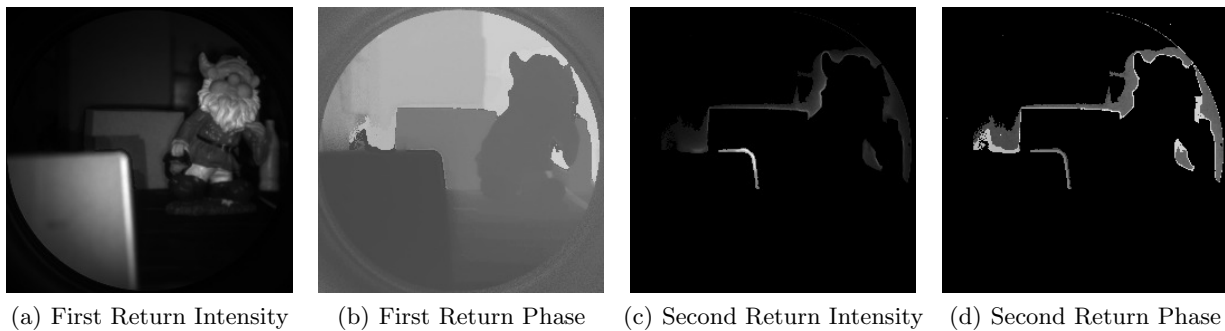
Figure 8. Recovered version of the scene from 6 using $C_l = 4000$. Mixed pixels around the edges of objects and some limited dark regions subject to scattered light have been recovered. Fig. 8(b) shows a noticible transition between the one and two return regions due to the systematic overestimation of the phase difference between two returns. There is also an extremely dark object on the left which was not clearly recovered.

In general the truncated-triangle fitting method was found to be slightly better than the base-ratio method. The maximum likelihood method for determining how many returns were present in each pixel appears to do a reasonable job without the necessity of a predefined prior distribution – however only limited results have been presented here.

ACKNOWLEDGMENTS

REFERENCES

[1] Conroy, R. M., Dorrington, A. A., Künnemeyer, R., and Cree, M. J., "Ranger performance comparison in homodyne and heterodyne operating modes," Proc. SPIE vol. 7239 (2009).

[2] Hebert, M. and Krotkov, E., "3D measurements from imaging laser radars: How good are they?," IVC 10, 170–178 (1992).

[3] Adams, M. D. and Probert, P. J., "The interpretation of phase and intensity data for AMCW light detection sensors for reliable ranging," Int. J. Rob. Res. 15(5), 441–458 (1996).

[4] Tang, P., Huber, D., and Akinci, B., "A comparative analysis of depth-discontinuity and mixed-pixel detection algorithms," Proc. Sixth Int. Conf. on 3D Im. and Mod. , 29–38 (2007).

[5] Tuley, J., Vandapel, N., and Hebert, M., "Analysis and removal of artifacts in 3-D ladar data," IEEE Int. Conf. on Robotics and Automation , 2203 – 2210 (2005).

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

[7] Gokturk, S., Yalcin, H., and Bamji, C., "A time-of-flight depth sensor – system description, issues and solutions," Computer Vision and Pattern Recognition Workshop, 2004 , 35–35 (2004).

[8] Godbaz, J. P., Cree, M. J., and Dorrington, A. A., "Mixed pixel return separation for a full-field ranger," Proc. IVCNZ 2008 (2008).

[9] Levy, S. and Fullagar, P. K., "Reconstruction of a sparse spike train from a portion of its spectrum and application to high-resolution deconvolution," Geophysics 46(9), 1235–1243 (1981).

[10] Schwarte, R. M., "Breakthrough in multichannel laser-radar technology providing thousands of high-sensitive lidar receivers on a chip," Laser Radar Techniques for Atmospheric Sensing. Proc. SPIE, Volume 5575 , 126–136 (2004).

[11] Hsu, S., Acharya, S., Rafii, A., and New, R., "Performance of a time-of-flight range camera for intelligent vehicle safety applications," Advanced Microsystems for Automotive Applications 2006 , 205–219 (2006).

[12] 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," Proc. SPIE vol. 6805 (2008).

[13] Kormylo, J. and Mendel, J., "Maximum likelihood detection and estimation of Bernoulli-Gaussian processes," Information Theory, IEEE Transactions on 28(3), 482–488 (1982).

[14] Kaaresen, K. F., "Deconvolution of sparse spike trains by iterated window maximisation," IEEE Trans. Signal Proc. 45(5), 1173–1183 (1997).

[15] Velis, D. R., "Stochastic sparse-spike deconvolution," Geophysics 73(1), R1–R9 (2008).

[16] Dorrington, A. A., Cree, M. J., Payne, A. D., Conroy, R. M., and Carnegie, D. A., "Achieving sub-millimetre precision with a solid-state full-field heterodyning range imaging camera," Meas. Sci. and Tech. 18(9), 2809–2816 (2007).

[17] Payne, A. D., Dorrington, A. A., Cree, M. J., and Carnegie, D. A., "Characterizing an image intensifier in a full-field range imaging system," Sensors Journal, IEEE 8(11), 1763–1770 (2008).

[18] Richardson, W. H., "Bayesian-based iterative method of image restoration," Journal of the Optical Society of America (1917-1983) 62, 55–59 (1972).

[19] Lucy, L. B., "An iterative technique for the rectification of observed distributions," Astronomical Journal 79, 745 (1974).

[20] Mure-Dubois, J. and Hugli, H., "Optimized scattering compensation for time-of-flight camera," Two- and Three-Dimensional Methods for Inspection and Metrology V 6762, SPIE, Boston, MA, USA (2007).