

Mixed Pixel Return Separation For a Full-Field Ranger

J. P. Godbaz¹, M. J. Cree, A. A. Dorrington

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

¹ Email: jpg7@waikato.ac.nz

Abstract

Full-field amplitude modulated continuous wave range imagers commonly suffer from the mixed pixel problem. This problem is caused by the integration of light from multiple sources by a single pixel, particularly around the edges of objects, resulting in erroneous range measurements. In this paper we present a method for identifying the intensity and range of multiple return values within each pixel, using the harmonic content of the heterodyne beat waveform. Systems capable of measurements at less than 90 degree phase shifts can apply these methods. Our paper builds on previous simulation based work and uses real range data. The method involves the application of the Levy-Fullagar algorithm and the use of the cyclic nature of the beat waveform to extract the mean noise power. We show that this method enables the separation of multiple range sources and also decreases overall ranging error by 30% in the single return case. Error in the two return case was found to increase substantially as relative intensity of the return decreased.

Keywords: Lidar, AMCW, Multiple Returns, Mixed Pixels, Limited Depth of Field, Sparse Spike Train Deconvolution

1 Introduction

Range-imaging devices work via the time-of-flight principle, utilising knowledge of the period of time it takes for light to travel to an object and back to the measurement device to determine the range to the object. Amplitude Modulated Continuous Wave (AMCW) lidar systems work by illuminating a scene with amplitude modulated light, often with lasers. Using a sinusoidally modulated waveform one can calculate the range to an object in a scene by measuring the phase offset due to time of flight of the received light, so that an object farther away results in a larger phase offset. Homodyne systems work by gain mixing the received signal with a reference signal at the same frequency as the modulation signal but at several different phase offsets. In practice this is usually a sinusoid at 90 degree phase offsets. From this autocorrelation operation a representative phasor can be calculated via a Fourier transform, $s[0] - s[2] + i(s[1] - s[2])$, where s is the recorded samples for a pixel. From this value, range and intensity can be extracted.

Mixing can take place either before or after detection, such as a modulated image intensifier before detection or a photodiode followed by an electronic mixer device. The University of Waikato Range Imager (1; 2) uses a modulated image intensifier which enables full-field high-resolution range im-

ages to be taken. The ranger uses a heterodyne modulation technique which involves modulating the image intensifier and the lasers at very slightly different frequencies. When these two modulation signals are mixed they produce a beating signal which is at a much lower frequency – low enough to be captured by an off-the-shelf Charge Coupled Device (CCD) camera. The phase offset of this resulting signal is proportional to the phase offset of the laser modulation versus the intensifier modulation and thus range. The ranger is designed for flexibility and is able to provide a large number of samples per beat – a facility which we take advantage of in this paper.

With the availability of full-field range-imaging devices such as SwissRanger (3) and CanestaVision (4), the importance of mixed pixels has increased substantially. The mixed pixel effect is caused by the integration of light from multiple sources at different ranges by a single pixel. This occurs in point scanning systems due to the finite width of the scanning beam, but is much worse in full-field systems because each pixel integrates over a large area. This can be particularly aggravated by limited depth-of-field, which causes defocus. In AMCW systems this results in the formation of pixels with unexpected ranges – either intermediate or values that wrap-around due to the cyclic nature of the measurements. In order for these devices to accurately measure range they must be

able to cope with mixed pixels.

1.1 Previous Work

There is a body of previous work on the mixed pixel problem (5; 6; 7; 8). Discussion has largely been limited to identification and removal of mixed pixel data using either normal-angle methods or restoration using simple methods such as median filters.

Previous ranger systems have used non-sinusoidal modulation (9; 10), which produces a non-sinusoidal output waveform, and papers on pulsed lidar systems have acknowledged the convolutional relationship between signal returns and recorded data (11). However, we are unaware of other work that explicitly considers the AMCW output waveform as a convolution of a Signal Return Model (SRM) with a cyclic waveform.

1.2 Overview

Here we present an end-to-end method for detecting and identifying the returns in mixed pixels extending our previous work (12). There we presented a method for correcting full-field ranger mixed pixel and defocus problems and tested it using simulated data. In this paper we utilise the same approach to identifying multiple returns but we extend it so that it is useful in practical applications.

By modelling the sequence of frames taken by the ranger as a convolution of a SRM with a beat waveform, the Levy-Fullagar (LF) deconvolution algorithm is applied to the data in a order to identify the range and intensity of all light sources within a pixel. We develop methods for extracting the mean noise power for a pixel and continuous range and intensity data.

2 Theory

We revise the definition of a SRM and frame the problem as a Sparse Spike Train (SST) deconvolution problem. We follow this with an explanation of the nature of the ranger noise and how we can determine the mean noise power, which is required by the LF deconvolution algorithm.

2.1 Signal Return Models

In (12) we explained that the sequence of captures taken by the heterodyne ranger can be regarded as a convolution of a SRM with the heterodyne beat waveform. Our SRM is a radial based voxel model, with three dimensions x , y and r , where x and y correspond to CCD pixel location and r is the radial distance from the camera. We assume that

the majority of pixels contain data from a small finite number of sources, modelled as Dirac delta functions, allowing us to apply SST techniques to the problem. In actuality, the signal returns within a pixel are not necessarily Dirac deltas as this is only true if all objects were segments of surfaces of spheres centred on the ranger.

2.2 Beat Waveform Formation

In the heterodyne ranger system, modulated laser light returned from a scene is further modulated by an intensifier response waveform. We represent the laser waveform as $l(\theta)$ with a frequency f_1 and the intensifier response waveform as $r(\theta)$ with a frequency f_2 . The photoelectrons reaching the phosphor screen of the intensifier thus have a resultant intensity waveform of $p(\theta_1, \theta_2)$, given by

$$p(\theta_1, \theta_2) = l(\theta_1)r(\theta_2). \quad (1)$$

This is sampled using a CCD camera, which integrates light over a period of time (and blurred by the phosphors). In the homodyne case ($f_1 = f_2$) the value sampled is dependent upon the difference in phase between θ_1 and θ_2 and a scale factor ρ , that varies according to integration time. We use $m(\theta)$ to represent the mean light intensity for a phase difference of θ ,

$$m(\theta_2 - \theta_1) = m(\theta) = \rho \int_{\theta - \pi}^{\theta + \pi} l(\phi)r(\phi + \theta)d\phi. \quad (2)$$

Payne (2) identified that only frequencies present in both the laser response waveform $l(\theta_1)$ and the intensifier response waveform $r(\theta_2)$ are present in the beat waveform $m(\theta)$. $m(\theta)$ can be identified as being a correlation of the laser waveform with the intensifier response. We notate the unsampled waveform over time for a single pixel as b , where n is the number of returns and θ_i and a_i are the phase and intensity of each return. Then

$$b(\theta) = \sum_{i=0}^{n-1} a_i m(\theta - \theta_i), \quad (3)$$

which is the convolution of a SST by $m(\theta)$. This is extended to the heterodyne case ($f_1 \neq f_2$) by sampling with a rect function. Here $B[t]$ is the sampled value at discrete timestep t and ω is the phase step per sample, giving us

$$B[t] = \int_{-\infty}^{\infty} b(\phi) \text{rect}(\phi/\omega - t)d\phi. \quad (4)$$

Dorrington (13) showed that since the Fourier transform of a rect function is a sinc function, the heterodyne case attenuates higher frequencies which tend to either contain noise or be the

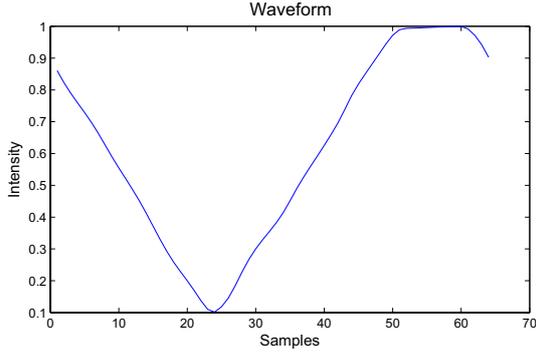


Figure 1: The range-imager beat waveform at 12.0MHz with a laser duty cycle of approximately 55%.

result of aliasing of harmonics above the Nyquist frequency. We consider B_{xy} , the raw data, to be a convolution of a discrete SRM s_{xy} by the beat waveform, for a particular pixel $(x, y) \in \mathbb{Z}^2$, thus

$$B_{xy}[t] = s_{xy} * b. \quad (5)$$

The Waikato Range Imager uses rectangular wave modulation for both the intensifier and the lasers, which has the added benefit of a stronger fundamental than would be possible with a sine wave for a given maximum intensity constraint (2). Importantly, the use of rectangular wave modulation results in harmonic content in the beat waveform which can be used to solve the mixed pixel problem. In theory the resultant beat waveform is a truncated triangle wave – however, in actuality the response of the intensifier and lasers is quite complex and the output shape is not perfect. Figure 1 shows the shape of the ranger system beat waveform at 12.0 MHz. Asymmetry in the beat waveform is beneficial to restoration quality – it results from varying the duty cycle of the modulation waveforms.

2.3 Noise Calculation and Deconvolution

We apply the Levy-Fullagar (LF) algorithm on a per pixel basis to the raw ranger data, which inverts equation 5 and produces a SRM which identifies the different returns within each pixel. The LF deconvolution algorithm is fully described in (14) – only a summary is presented here. It works by using knowledge of the impulse response and noise power to set constraints in the Fourier domain. A linearly constrained system is formed, which is then solved for the minimum possible L_1 norm using linear programming methods – this ensures sparsity.

The Waikato Range Imager has several noise sources, primarily shot noise and readout/thermal noise. Shot noise is Poisson distributed and the noise power is proportional to the intensity of

the light. Readout/thermal noise is Gaussian distributed and is consistent across multiple samples on a per pixel basis. We make the assumption that the resultant combined noise for a particular pixel versus time can be approximated as Gaussian white in the Fourier domain. The assumption of whiteness has been previously used for the application of linear techniques such as Wiener filtering to shot noise limited data (15).

The total noise power required for the LF algorithm is estimated by exploiting the repetitive nature of the beat waveform, by calculating the sample variance of the noise for each frame in the beat cycle. For example, if there are 64 samples per beat, then every 64th sample is a sample from the same distribution. Taking $B[t]$ as the raw ranger samples for a single pixel over time, consisting of N complete beats ($N > 2$) and n frames per beat ($n > 4$), we have Nn samples. Then m_i , the mean intensity of sample i in the beat cycle, is calculated as

$$m_i = \frac{\sum_{j=0}^{N-1} B[nj + i]}{N}, \quad (6)$$

and the sample variance for sample i in the beat cycle, σ_i^2 , by

$$\sigma_i^2 = \frac{\sum_{j=0}^{N-1} (B[nj + i] - m_i)^2}{N - 1}, \quad (7)$$

then the mean noise power $\overline{\sigma^2}$ for the pixel is estimated by

$$\overline{\sigma^2} = \frac{\sum_{j=0}^{n-1} \sigma_j^2}{n}. \quad (8)$$

An interesting extension of this method would be to determine the spectral power of the noise at each frequency and utilise it for the deconvolution.

2.4 Identification of Discrete Returns

We now deal with the extraction of continuous range data from the discrete SRM (see equation 3). Once the SRM has been deconvolved using the LF algorithm, for each pixel we have a series of range bins. Each range bin represents a discrete phase of the beat waveform and thus a discrete range. If the actual range of a return lies between two bins, then the intensity of that return is distributed across the two adjacent bins according to the underlying continuous range value.

First we identify each return. We create a mask $m_{xyw}[a] \in \{0, 1\}$ for the region belonging to a particular return w at pixel location $(x, y) \in \mathbb{Z}^2$, where $a \in \mathbb{N}_0$ is a discrete range value by identifying contiguous range bins which have an intensity above 5% of the total range bin intensity for that pixel. The region above the threshold and one bin to either side is assigned the mask value 1,

all other regions the value 0. We can then take the first bin of a Fourier transform for each set of masked range bins and produce a phase offset, thus range measurement for each return. $p_{xyw} \in \mathbb{C}$ is a phasor representative of the intensity and phase of the return at pixel location from the SRM data $s_{xy}[a]$, thus

$$p_{xyw} = \sum_{a=0}^{N-1} e^{2i\pi(a-1)/N} s_{xy}[a] m_{xyw}[a]. \quad (9)$$

The problem of identifying continuous ranges from the deconvolved SRM can be compared to the problem of producing continuous range data from ranging systems. Centroiding, thresholding, simple maxima and second order methods have been applied to this problem (11). We use eqn. 9 because it produces results similar to centroiding but takes into account points that partially wrap around.

3 Methodology

The experiment was carried out at 12 MHz with a beat frequency of 0.5 Hz at a frame rate of 32 fps. The f/number was 4.5 and the scene was focussed such that the object at the front was in focus, leaving the background quite blurred due to limited depth of field – this provides a large number of mixed pixels to work with, however it limits our ability to make a valid comparison to restoration by median filter. In order to minimise the signal to noise ratio sequences of 4096 frames, consisting of 64 beat cycles of 64 frames each, were used. 16 beats or less could have been used, but with reduced quality.

Typically the ranger uses coaxial illumination by positioning four laser sources closely around the main ranger optics, however this results in multiple returns for off-axis regions. To ensure that this was not an issue, only one of the laser sources was used. A high number of frames was required in this particular case because of the low light levels and comparatively low modulation frequency. Submillimetre precision is achievable with higher frequencies and greater light (1).

3.1 Zeroed Frames

The Pantera 1F TM60 camera used in the Waikato Range Imager has a buggy firmware version that results in occasional zeroed frames. These show normal pixel non-uniformity, but have values as if not exposed to any light. In order to restore zeroed frames, several cycles of extra frames at the end of each capture are taken and used to replace zeroed frames with a new sample from the same position in the beat cycle. Zeroed frames were detected by comparing the mean frame intensity against a threshold.

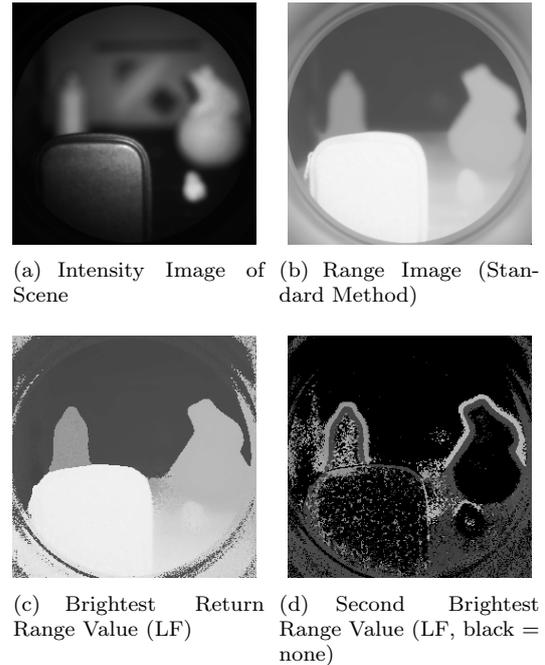


Figure 2: Range image data before (2(a-b)) and after processing (2(c-d)). White represents an object close to the camera, dark grey farther away. Note the light scattering outside the circular image intensifier. The original data are subject to limited depth of field which accentuates the mixed pixel effect.

3.2 Calibration

In order to accurately represent the intensity of the returns, the system was calibrated for pixel variation (constant offsets and flat-fielding). There are several motivations for this: firstly there is pixel-to-pixel additive and multiplicative variation on the CCD and secondly there are intensity variations introduced by the image intensifier.

The waveform was calibrated for by taking a range image of a planar board and averaging the waveform over 64 beat cycles over a small region of several hundred pixels in the centre of the board. This mean waveform was then used as a model for the impulse response by the LF algorithm.

3.3 Measurement of Ranging Error

We compared the standard method of range measurement, which calculates the phase of the Fourier transform bin corresponding to the fundamental frequency of the beat waveform, to the LF based method. The standard method is known to give reliable results when aliasing effects are minimised. We used 64 frames per beat in order to achieve this. To analyse the overall single return error, an image was taken of a flat board. A model of the form

$$r(x, y) = a + \sqrt{bx^2 + cy^2 + dxy + ex + fy + g}, \quad (10)$$

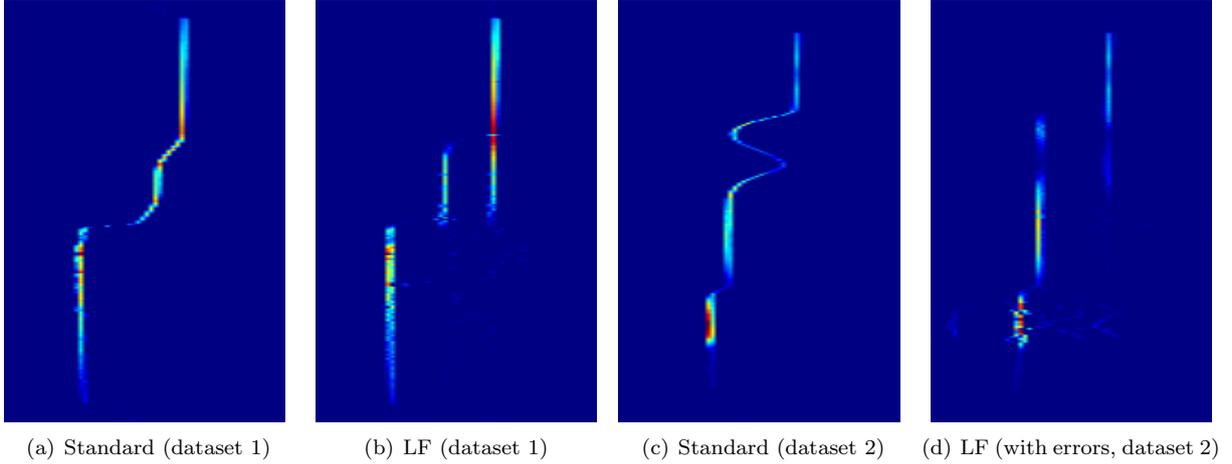


Figure 3: Vertical slices through a SRM created via the standard method or the LF algorithm method. The x axis represents range, the y axis represents vertical slice location – blue represents a low light level and red high. Any line parallel to the x axis which intersects more than one return in the recovered SRM is a mixed pixel in the standard version.

which handles constant offsets, sloping boards and radial curvature, was fit to range data extracted by the standard method. This model was then used as a reference to estimate the overall ranging error for the LF method results.

A second experiment was carried out using two flat boards under defocus. The boards were positioned so that there was a blurred boundary between them. Eqn. 10 was then fit to a selected region from each board, thus allowing the extrapolation of the range data corresponding to each board over the blurred region.

4 Results

The figure 2 example clearly shows the blurred object boundaries of the original range data and the improvement in figure 2c. Figure 2d shows the range of the second brightest return, and regions where objects have blurred onto each other are clearly visible. The figure also shows light scattering, probably due to the relay lens coupling between the image intensifier and CCD, which results in changes to the measured ranges of dark objects.

A different way to look at the problem is by viewing slices through the SRM. Two slices through the figure 2 model are given in figures 3(a) to 3(d). The second example shows a breakdown in the algorithm – this effect is known to occur in the brightest regions of images particularly when high microchannel plate voltages are used, and may be caused by the image intensifier.

Using the method in section 3.3, the LF method RMS error versus the flat board model in the single return case was found to be smaller than the stan-

Table 1: Board Experiment Results

Single Return Results (RMS Error vs. Model)	
Standard Method (64 Beats)	0.0204m \pm 0.0005
Standard Method (1 Beat)	0.0580m \pm 0.0127
LF Method	0.0143m \pm 0.0020
LF Method (Ignr. False Ret.)	0.0105m \pm 0.0002
LF False Multiple Returns	0.61%
Two Returns Results (RMS Error vs. Model)	
Standard Method	N/A
LF Method	See Fig. 4
LF False Single Returns	1.15%

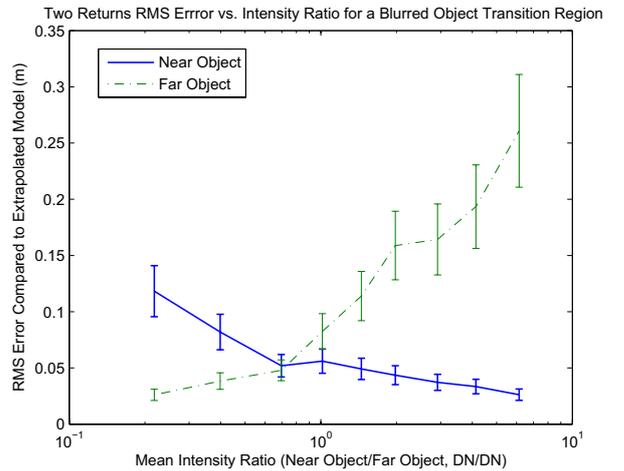


Figure 4: RMS Error of Recovered Returns vs. Intensity Ratio in the Case of Two Returns.

dard method RMS error (table 1) despite the standard method data being the basis for the model. This shows that in the single return case, the LF method can decrease overall error by 30%. In the two returns case the size of the residuals depends upon the relative intensities of the returns. Fig-

ure 4 shows RMS Error versus the intensity ratio. The intensity ratio is the relative intensity of the returns composing each pixel (near object intensity/far object intensity). Error in measuring the brightest return is fairly low, however the darker return appears to be negatively affected by the bounds inherent in the LF method.

5 Conclusion

We have shown that deconvolution techniques can be applied to AMCW lidar systems capable of sampling at less than 90 degree phase shifts to successfully identify the range of light sources within mixed pixels. In the single return case, use of the LF method was shown to improve overall ranging error by 30%. However, ranging error was found to increase substantially as the relative intensity of a return decreased in the two return case.

6 Acknowledgements

J. Godbaz acknowledges the support of a Tertiary Education Commission Top Achiever Doctoral Scholarship.

References

- [1] A. A. Dorrington, M. J. Cree, A. D. Payne, R. M. Conroy, and D. A. Carnegie, "Achieving sub-millimetre precision with a solid-state full-field heterodyning range imaging camera," *Meas. Sci. and Tech.*, vol. 18, no. 9, pp. 2809–2816, 2007.
- [2] A. D. Payne, "Development of a full-field time-of-flight range imaging system," Ph.D. dissertation, University of Waikato, 2008.
- [3] T. Oggier, M. Lehmann, R. Kaufmann, M. Schweizer, M. Richter, P. Metzler, G. Lang, F. Lustenberger, and N. Blanc, "An all-solid-state optical range camera for 3d real-time imaging with sub-centimeter depth resolution (SwissRanger)," in *Opt. Des. and Eng.*, vol. 5249. SPIE, 2004, pp. 534–545.
- [4] S. Gokturk, H. Yalcin, and C. Bamji, "A time-of-flight depth sensor – system description, issues and solutions," *Computer Vision and Pattern Recognition Workshop, 2004*, pp. 35–35, 2004.
- [5] M. Hebert and E. Krotkov, "3d measurements from imaging laser radars: How good are they?" *IVC*, vol. 10, pp. 170–178, 1992.
- [6] M. D. Adams and P. J. Probert, "The interpretation of phase and intensity data for amcw light detection sensors for reliable ranging," *Int. J. Rob. Res.*, vol. 15, no. 5, pp. 441–458, 1996.
- [7] J. Tuley, N. Vandapel, and M. Hebert, "Analysis and removal of artifacts in 3-d lidar data," in *IEEE Int. Conf. on Robotics and Automation*, 2005, pp. 2203 – 2210.
- [8] P. Tang, D. Huber, and B. Akinici, "A comparative analysis of depth-discontinuity and mixed-pixel detection algorithms," in *Proc. Sixth Int. Conf. on 3D Im. and Mod.*, Los Alamitos, CA, USA, 2007, pp. 29–38.
- [9] R. M. Schwarte, "Breakthrough in multichannel laser-radar technology providing thousands of high-sensitive lidar receivers on a chip," in *Laser Radar Techniques for Atmospheric Sensing. Proc. SPIE, Volume 5575, pp. 126-136 (2004).*, vol. 5575, 2004, pp. 126–136.
- [10] S. Hsu, S. Acharya, A. Rafii, and R. New, "Performance of a time-of-flight range camera for intelligent vehicle safety applications," in *Advanced Microsystems for Automotive Applications 2006*, 2006, pp. 205–219.
- [11] W. Wagner, A. Ullrich, T. Melzer, C. Briese, and K. Kraus, "From single-pulse to full-waveform airborne laser scanners: Potential and practical challenges," in *International Society for Photogrammetry and Remote Sensing XXth Congress*, 2004.
- [12] J. P. Godbaz, M. J. Cree, A. A. Dorrington, and R. Künnemeyer, "Defocus restoration for a full-field heterodyne ranger via multiple return separation," in *Proc. IVCNZ 2007*, Hamilton, New Zealand, 2007, pp. 52–57.
- [13] A. A. Dorrington, M. J. Cree, D. A. Carnegie, A. D. Payne, R. M. Conroy, J. P. Godbaz, and A. P. P. Jongenelen, "Video-rate or high-precision: a flexible range imaging camera," in *Image Processing: Machine Vision Applications. Proc. SPIE, Volume 6813, pp. 681307-681307-12 (2008).*, vol. 6813, 2008.
- [14] S. Levy and P. K. Fullagar, "Reconstruction of a sparse spike train from a portion of its spectrum and application to high-resolution deconvolution," *Geophysics*, vol. 46, no. 9, pp. 1235–1243, 1981.
- [15] M. A. King, P. W. Doherty, and R. B. Schwinger, "A Wiener filter for nuclear medicine images," *Med. Phys.*, vol. 10, no. 6, pp. 876–880, 1983.