J. P. Godbaz, M. J. Cree, A. A. Dorrington, R. Künnemeyer, 'Defocus Restoration for a Full-Field Heterodyne Ranger via Multiple Return Separation', *Proceedings of Image and Vision Computing New Zealand 2007*, pp. 52–57, Hamilton, New Zealand, December 2007.

Defocus restoration for a full-field heterodyne ranger via multiple return separation

J. P. Godbaz[1], M. J. Cree, A. A. Dorrington and R. Künnemeyer

Department of Engineering, University of Waikato.
[1]Email: jpg7@waikato.ac.nz

Abstract

Full-field heterodyne time-of-flight range-imagers allow a large number of range measurements to be taken simultaneously across an entire scene; these range measurements may be corrupted due to limited depth of field. We propose a new method for deblurring heterodyne range images by identifying multiple signal returns within each pixel via deconvolution, thus reducing the spatially variant deblurring problem to a sequence of spatially invariant deconvolutions. We have applied this method to simulated data, showing significant improvement in the restored images.

Keywords: Defocus, heterodyne, range imaging, lidar, spatially variant blur

1 Introduction

Full-field ranging is a developing technology that allows the simultaneous capture of two dimensional images, where each pixel value represents the range to an object in the scene. Range data have many valuable applications such as robot vision and quality control inspection, where three dimensional models can be quickly produced of parts and checked against specifications. We focus here on heterodyne time-of-flight range imaging, in particular the heterodyne range imager described in refs. (1; 2).

The ranger works by illuminating a scene with modulated laser light and then imaging with a gain modulated image intensifier that shutters a charge coupled device (CCD) camera. By modulating the laser at a slightly different frequency to the image intensifier, it is possible to measure how much distance the light has travelled. Image sequences captured with the camera show beating objects, where the phase offset of the beating corresponds to the phase delay of the modulated light, and thus the range to the object.

One issue with high frequency modulation of lasers and image intensifiers is that for some systems the net amount of light transmitted to the CCD can decrease as modulation frequencies increase, due to alterations in duty cycles. In order to maintain as high a signal to noise ratio (SNR) as possible, it can be necessary to use larger apertures. However, there is an inherent trade-off between larger apertures, which suffer from smaller depth of field, and smaller apertures, which allow less light through but have much greater depth of field. This is exacerbated because many applications require a full-field ranger that can operate over a wide range of distances.

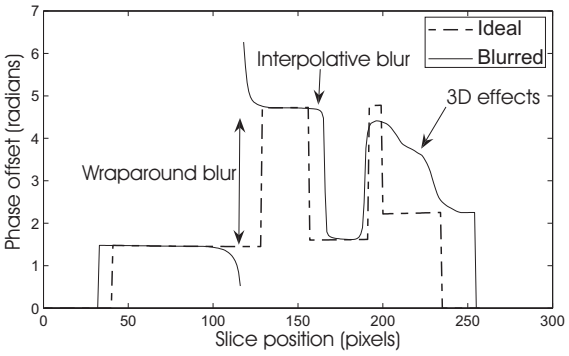


Figure 1: A slice through a simulated range image showing the effects of spatially variant blurring for a sinusoidal beat waveform. Interpolative blur manifests when two objects are within π radians of each other, this is a smooth transition between the two ranges. Wraparound blur manifests when two objects are separated by more than π radians, this can result in implausible ranges. For clarity, a substantial blur has been applied.

Although defocus is capable of producing erroneous data there has been little discussion of the effects of defocus on full-field rangers – an example is shown in figure 1. Defocus results in a spatially variant point spread function (PSF), which is dependant upon the physical location of objects within a scene. Previously, depth from defocus methods (3; 4; 5) have allowed range measurements, thus production of refocused images. These methods

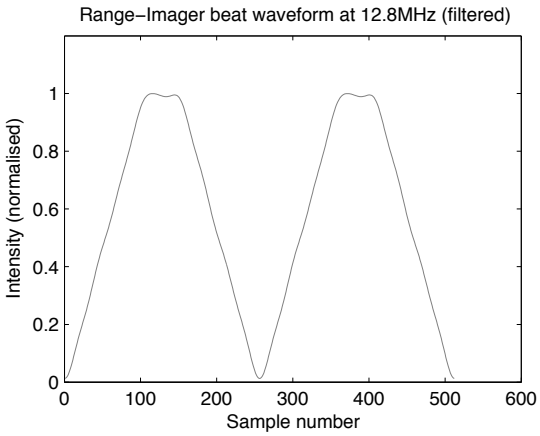


Figure 2: The range-imager beat waveform at 12.8MHz.

have relied upon known characteristics of the modulation transfer function (MTF) of an optical system, one example (5) used a coded aperture with known zeros in the MTF and spatially-variant deblurring techniques to create an in-focus intensity image. However, these systems have usually been limited in their resolution and ability to handle non-patterned surfaces. Also, they have ignored multiple returns within the same pixel.

Identification of multiple returns has been achieved with pulsed range-gating techniques (6), however heterodyne technology requires different methods. One possible method is to use multiple captures at different modulation frequencies to discriminate multiple sources. Multiple superimposed modulation frequencies has been suggested as a method of alleviating the range ambiguity problem, but requires the implementation of more complex hardware (2).

For the purposes of this paper we have used the natural harmonic content of the beat waveform. Since the Waikato ranger uses square wave modulation for both the lasers and intensifier, in an ideal situation the beat waveform is a triangle wave. This is similar to taking several images at different frequencies and allows the production of a voxel based model of signal returns via a positively constrained deconvolution. This signal return model has voxels of irregular volume, due to the radial nature of the range data, and can be used to produce deblurred range and intensity data.

In this paper we apply the Levy-Fullagar sparse spike deconvolution algorithm (7) and the Richardson-Lucy algorithm (8; 9) to simulated ranger data to show that reconstruction of defocused range data is possible. We use an experimentally derived beat waveform rather than a triangle wave because it is asymmetric and reduces cancellation effects (figure 2).

2 Theory

2.1 Simple sinusoidal case

Consider two objects in a scene that appear in adjacent pixels in a ranger movie, but are at different ranges. For a simple full-field heterodyne ranger with a sinusoidal waveform, defocus manifests as a sinusoidal waveform with an additional fixed offset due to the superposition of the blurred neighbouring pixels. If we represent the sinusoidal components as phasors, we can regard the blurring as a 2D complex domain spatially variant blurring problem, where the variant PSF is defined by the unblurred data. If we extend this model to allow for pixels near the edges of objects, which integrate signals from an unknown, but small, number of sources then it becomes practically unsolvable without further information. If we do not handle multiple signal returns then the PSF is miscalculated around the edges of objects, leading to degraded restoration quality.

2.2 Signal return models

In this section we first analyse the ideal pinhole camera that has no depth of field problem, we then generalise this to finite apertures in section 2.3.

For an ideal pinhole camera, we can assume that each point on the imaging plane receives light from a single source point in the scene. We can notate this intensity as $a(x, y)$ where $(x, y) \in \mathbb{R}^2$ are coordinates on the imaging plane.

We can extend this model to include range as a third dimension by considering the signal intensity at the imaging plane incident from a particular radial distance

$$I(x, y, r) = a(x, y)\delta(r - r_0(x, y)), \qquad (1)$$

where $r \in \mathbb{R}^+$ is the radial distance from the pinhole to the object, $r_0(x, y)$ is the distance of the object projected onto the image plane at the point (x, y) and δ is the Dirac delta function. We call the function I the signal return model.

We can notate the intensity incident at the point (x, y) for a heterodyne ranger at any phase in the beat cycle as

$$f(x, y, \theta) = \int_0^\infty I(x, y, r)\psi(2\pi r/\lambda - \theta)\, dr, \qquad (2)$$

where θ is the phase offset into the beat cycle, ψ is the cyclic beat waveform and λ is the wavelength of the beat signal. This is a convolution of the signal return model with the beat waveform.

Thus for a discrete pixel (i, j) on the imaging plane for a pinhole system, the intensity at any phase in the beat cycle is

$$f_{i,j}(\theta) = \psi * I_{i,j}, \tag{3}$$

with $*$ representing convolution, and

$$I_{i,j}(r) = \int_{i-0.5}^{i+0.5} \int_{j-0.5}^{j+0.5} I(x, y, r)\, dx\, dy, \tag{4}$$

For most practical situations, the determination of the signal return model from the ranger frame sequence can be considered to be a sparse spike train deconvolution problem. However, this does not hold for returns from surfaces near parallel to the optical axis.

2.3 Defocused signal return models

The model can be extended through analysis of defocus. We now show that we can approximate distance along the optical axis by radial distance and thus reduce defocus to a sequence of spatially invariant convolutions. For a conventional 2D representation, defocus is described by a spatially variant blur

$$a_{blurred}(x, y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} h(x - \xi, y - \eta, \xi, \eta) a(\xi, \eta)\, d\xi\, d\eta, \tag{5}$$

where the blurring function, h, is scene dependent. For a thin lens model, we can assume that only the distance along the optical axis is important for PSF calculation. By geometry the projection of a radial distance onto the optical axis is

$$proj_z(x, y, r) = r \frac{d_p}{\sqrt{x^2 + y^2 + d_p^2}}, \tag{6}$$

for a perspective projection plane at z distance d_p

$$d_p = \frac{x_{dim}}{2 \tan(\theta_x / 2)}, \tag{7}$$

where x and y are measured in pixels, r is measured in metres, x_{dim} is the width of the field of view in pixels and θ_x is the corresponding angular width. This is shown in figure 3. We make this assumption for the generation of all synthesised data for restoration therein. If we assume that θ_x is sufficiently small, we can make the paraxial assumption

$$proj_z(x, y, r) \approx r, \tag{8}$$

Thus for a signal return model, the problem can be approximated as

$$I_{blurred}(x, y, r) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} h_r(x - \xi, y - \eta) I(\xi, \eta, r)\, d\xi\, d\eta, \tag{9}$$

which is a spatially invariant convolution across each spherical surface of constant radius from the ranger.

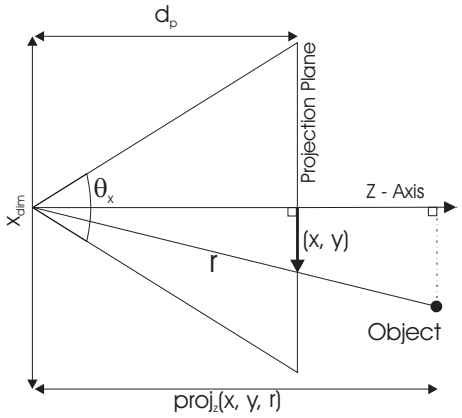


Figure 3: A diagrammatic representation of the projection of a radial distance, r, onto the optical axis for a pixel in the range image at (x, y). Since the range image is subject to perspective projection, we place the perspective projection plane at a distance d_p such that pixel units and metres are equivalent – see equations 6 and 7. The y-axis is out of the page.

3 Methodology

3.1 Test sequence generation

The 256×256 scene used for testing is shown in figures 5(a) and 5(b). The scene is composed of several surfaces perpendicular to the optical axis at different distances, including several with patterned surfaces. In unblurred form, this scene has only a single signal return per pixel.

A 256×256×256 capture sequence was generated by creating an ideal, single beat, unblurred movie, i.e. a triangle wave for each pixel versus time, continuously variable in phase offset and intensity. This was subsequently spatially variantly blurred across each frame, using the defocus model described in section 3.2, followed by addition of Gaussian white noise. We assume that there are no objects with a phase offset of more than 2π radians, so there are no ambiguous ranges.

3.2 Defocus model

We modelled a thin lens system using geometrical optics and a continuously variable pillbox PSF model. The thin lens equation allows calculation of the image location

$$\frac{1}{d_i} = \frac{1}{f} - \frac{1}{d_o}, \tag{10}$$

where d_i and d_o are distances along the optical axis assumed to be equal to the radial distance for the purposes of image restoration. And it follows by geometry that

$$r_p = \frac{r_a}{d_i} |d_w - d_i|, \tag{11}$$

where r_p is the radius of the PSF, r_a is the radius of the lens and d_w is the distance along the optical axis between the lens and the imaging plane.

3.3 Deconvolution of the signal return model

Deconvolution of the frame sequence to produce the signal return model is a difficult operation. While the convolution is naturally circular, which allows the use of Fourier based techniques, the impulse response is as long as the data. Initial attempts to apply conventional deconvolution methods such as the Landweber and Richardson-Lucy methods required a large number of iterations to converge and were negatively impacted by constant offsets (e.g. ambient light).

Instead we have approached this problem from a sparse spike train perspective using the comparatively simple inequality constraints algorithm of Levy and Fullagar (7). This algorithm is faster than L^1 norm regularization techniques (10) which have commonly been applied to similar problems in Geology, and does not require determining the parameters of a Bernoulli-Gaussian distribution (11; 12).

This algorithm is designed for the deconvolution of sparse spike trains where only a small part of the Fourier spectrum can be recovered. It operates by using the statistical properties of Gaussian noise to set maximum and minimum constraints in the Fourier domain and then using linear programming to find the solution with the minimum L^1 norm. This ensures sparsity of the solution. Unlike many general deconvolution techniques, constant offsets within the convolved data are ignored. The constraints are of the form

$$\pm Re\{A_j\} + \epsilon_j > \pm \sum_{n=0}^{N-1} b_n \cos(2\pi nj/N), \quad (12)$$

$$\pm Im\{A_j\} + \epsilon_j > \pm \sum_{n=0}^{N-1} b_n \sin(2\pi nj/N), \quad (13)$$

where A_j is the Fourier transform of the recovered object for frequency j, b_n is the value of the deconvolved data for position n, and ϵ_j is an estimate of the error bounds given by

$$\epsilon_j = \frac{\beta\sigma\sqrt{N}}{\sqrt{2}|W_j|}, \quad (14)$$

where β defines Gaussian confidence limits and W_j is the Fourier transform of the PSF for frequency j. The values are also positively constrained. For the purposes of this paper, we have used a value of $\beta = 4$ and have only used the first 8 harmonics of the beat waveform. To mitigate noise effects

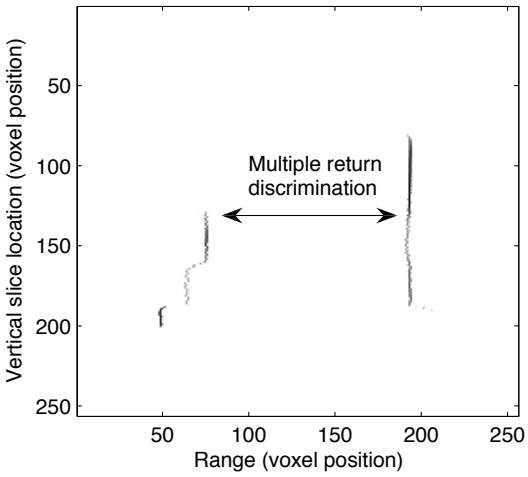


Figure 4: A vertical slice through a signal return model under defocus, showing object surfaces clearly discriminated despite multiple returns. Minor noise is visible. Image was generated from data with a SNR of 1000:1 via the Levy-Fullagar algorithm.

(as explained in section 4), we have then applied a Gaussian blur in the r direction with $\sigma = 6$. This value was determined by trial and error, an optimum blur size is yet to be determined.

3.4 Deconvolution of r-slices

Richardson-Lucy deconvolution (8; 9) was applied to the invariant r-slice deconvolution problem. For an original blur of the form $g = \mathrm{H}f$ the Richardson-Lucy algorithm is

$$\hat{f}_0 = \text{positive initial values}, \quad (15)$$

$$\hat{f}_{n+1} = \hat{f}_n \odot \mathrm{H}^T[g/(\mathrm{H}\hat{f}_n)], \quad (16)$$

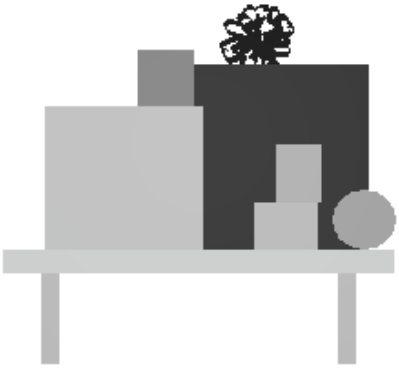
for the original image f, an estimate of the original image $\hat{f}$, the transformation matrix H, a blurred image g, with $\odot$ and / being elementwise multiplication and division respectively. Two hundred iterations were used for each image. Range data was produced by finding the phase of the lowest frequency bin of the Fourier transform versus range from the restored voxel model for each pixel.

4 Results

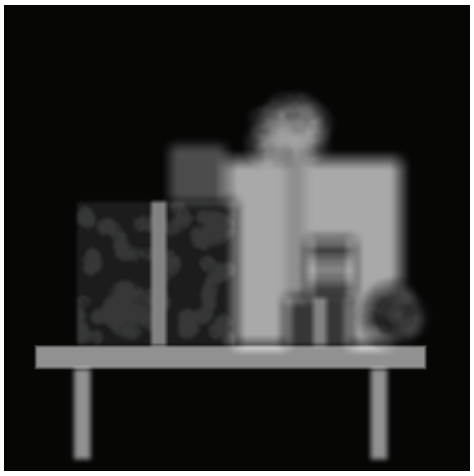
4.1 Multiple return discrimination

Initial results indicate that the Levy-Fullagar method is able to identify multiple returns within the same pixel quite well. However, due to the effects of noise there is a tendency to misplace the signal return by a few voxels – this can be seen in figure 4. Attempting to immediately deconvolve slices of this signal return model would result in a particularly poor result due to the combined effects of this noise and the discrete nature of each

(a) Ideal intensity image of scene

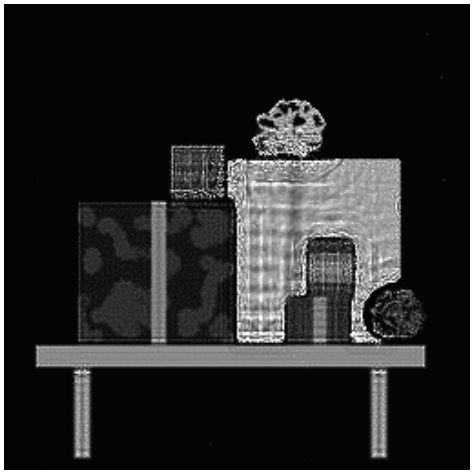


(b) Ideal range image of scene



(c) Intensity image from blurred data

(d) Range image from blurred data



(e) Intensity image from restored data

(f) Range image from restored data

Figure 5: An example range image restoration. Intensity images were produced by summing over the voxel model for each pixel. Figures d and f were produced via Fourier transform, d using the phase of the fundamental frequency of the beat waveform, f by taking the phase of the lowest frequency bin from the Fourier transform of the restored voxel model vs. range. For range images, white represents objects closer to the camera and black is farther away. Range image pixels with an intensity below 0.3 have been set to 0 to limit the influence of noise. This example used a SNR of 1000:1. The dynamic range of the ringing in figure e exceeds that of the image.

source. In order to mitigate these effects we apply a Gaussian blur to the data (see 3.3) before RL deconvolution, this turns each single pixel peak into a larger, more diffuse peak that results in less noisy output. This comes at the cost of incorrect deconvolution, especially for objects particularly close to the camera where a small change in range can result in a much larger change in the PSF.

4.2 Deblurring of a range image

Figure 5 shows the results of deblurring a simulated range image capture sequence. The deblurred range image figure 5(f) has better defined edges than the blurred image 5(d) and some of the detail in the intensity image 5(e) has been recovered. However, noise has been added by multilayered ringing effects and some regions have no range information. Where multiple returns are separated by near π radians, reduced harmonic information is available due to cancellation. In the case of a triangle wave this can result in complete cancellation, but because the beat waveform contained even harmonics this mostly resulted in attenuation rather than complete loss of data. This attenuation is largely responsible for the ringing. Artefacts are also visible on the front of the table in the intensity image, due to negative effects from the Gaussian blur.

5 Summary

We have shown that it is possible to restore simulated full-field heterodyne range images in the presence of noise by identifying multiple returns for each pixel. This algorithm appears to be able to restore range data which have been blurred by simulated defocus, but ringing appears to limit the quality of restored intensity images. A more suitable sparse spike deconvolution method needs to be identified, which reduces attenuation effects and range displacement error. Possible future research includes use of multiple captures, custom maximum likelihood algorithms, algorithms with reduced computational complexity and extensions to Poisson noise.

6 Acknowledgements

Acknowledges the support of a Tertiary Education Commission Top Achiever Doctoral Scholarship.

References

[1] M. J. Cree, A. A. Dorrington, R. M. Conroy, A. D. Payne, and D. A. Carnegie, "The Waikato range imager," in *Image and Vision Computing New Zealand (IVCNZ'06)*, Gt. Barrier Is. New Zealand, 2006, pp. 233–238.

[2] 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," *Measurement Science and Technology*, vol. 18, no. 9, pp. 2809–2816, 2007.

[3] A. P. Pentland, "A new sense for depth of field," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 9, no. 4, pp. 523–531, 1987.

[4] S. Hiura and T. Matsuyama, "Depth measurement by the multi-focus camera," in *CVPR '98: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. Washington, DC, USA: IEEE Computer Society, 1998, p. 953.

[5] A. Levin, R. Fergus, F. Durand, and W. T. Freeman, "Image and depth from a conventional camera with a coded aperture," in *Proc. SIGGRAPH 07*, 2007.

[6] B. W. Schilling, D. N. Barr, G. C. Templeton, L. J. Mizerka, and C. W. Trussell, "Multiple-return laser radar for three-dimensional imaging through obscurations," *Applied Optics*, vol. 41, pp. 2791–2799, 2002.

[7] 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.

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

[9] L. B. Lucy, "An iterative technique for the rectification of observed distributions," *Astronomical Journal*, vol. 79, p. 745, 1974.

[10] F. Santosa and W. W. Symes, "Linear inversion of band-limited reflection seismograms," *SIAM Journal Sci. Stat. Comput.*, vol. 7, no. 4, pp. 1307–1330, 1986.

[11] F. Champagnat, Y. Goussard, and J. Idier, "Unsupervised deconvolution of sparse spike trains using stochastic approximation," *IEEE Trans. Signal Proc.*, vol. 44, no. 12, pp. 2988–2998, 1996.

[12] K. F. Kaaresen, "Deconvolution of sparse spike trains by iterated window maximisation," *IEEE Trans. Signal Proc.*, vol. 45, no. 5, pp. 1173–1183, 1997.