Blind deconvolution of depth-of-field limited full-field lidar data by determination of focal parameters

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

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

ABSTRACT

We present a new two-stage method for parametric spatially variant blind deconvolution of full-field Amplitude Modulated Continuous Wave lidar image pairs taken at different aperture settings subject to limited depth of field. A Maximum Likelihood based focal parameter determination algorithm uses range information to reblur the image taken with a smaller aperture size to match the large aperture image. This allows estimation of focal parameters without prior calibration of the optical setup and produces blur estimates which have better spatial resolution and less noise than previous depth from defocus (DFD) blur measurement algorithms. We compare blur estimates from the focal parameter determination method to those from Pentland's DFD method, Subbarao's S-Transform method and estimates from range data/the sampled point spread function. In a second stage the estimated focal parameters are applied to deconvolution of total integrated intensity lidar images improving depth of field. We give an example of application to complex domain lidar images and discuss the trade-off between recovered amplitude texture and sharp range estimates.

Keywords: Lidar, Blind Deconvolution, Depth From Defocus, Focal Parameter Determination, Depth of Field, Mixed Pixels, AMCW

1. INTRODUCTION

Full-field range imagers work on the time-of-flight (TOF) principle, acquiring range data for every pixel simultaneously. Whereas in the past imaging lidar devices have used actuated point sensors acquiring data sequentially,[1] the use of full-field sensors[2–4] allows much simpler mechanical design and mass production leading to a reduction in costs. However the use of full-field sensors comes at a cost; many new problem are created such as range-intensity coupling and finite depth-of-field (DOF). Mixed pixels, also known as multipath interference, occur in full-field Amplitude Modulated Continuous Wave rangers and are caused by a pixel integrating light from more than one source.[5,6] This can be due to light scattering within an optical system, defocus blur or unavoidably around the edges of objects. As a result the measured ranges are erroneous, often appearing to have little or no connection to the range to the component light sources. While a number of correction/separation algorithms have been proposed,[7–9] in general mixed pixels remain a fundamentally unsolved problem.

As the resolution of sensors and the necessity for system flexibility increases, DOF issues become problematic. Firstly there are calibration issues – commercial full-field lidar sensors come with fixed focal length wide aperture lenses because if focal length were to change the calibration required to produce accurate point clouds would be invalidated. Additionally, wide apertures result in greater ranging precision due to an increase in the signal-to-noise ratio. This means that when imaging an object which is either very close to the camera or a long way away, the range image is out of focus. In practice this results in a blurry amplitude image and the formation of a wide band of mixed pixels around the edges of objects; thus much of the recorded range data is unreliable. Secondly, even if it were possible to change focus all optical systems suffer from finite DOF – meaning that certain scenes will always suffer from some defocus. One way to mitigate this is to take multiple images at different focal lengths and combining the most in-focus regions of each,[10] instead in this paper we choose to approach this problem by deconvolving range images to extend the DOF. In order to make our algorithm as flexible as possible we determine the defocus parameters of the optical system from the captured range data – thereby allowing an operator to freely refocus the range imager and even change the lens without needing to resample the point spread function (PSF). While we do not deal with other aspects of automatic calibration, we believe that this level of flexibility is going to become more important as full-field lidar technology develops further.

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

1.1 Previous Work

Previous work on single frame blind deconvolution has primarily focussed on the isoplanatic case;[11] this is largely because the problem is ill-posed and it is difficult to quantify how blur changes across a single isolated image. Single image spatially variant deblurring has generally involved complex blur calibration rather than blind deconvolution.[12] In other applications blurring issues have been avoided by use of special hardware such as adaptive optics for telescopes.[13] In the case of limited DOF, high quality restoration requires implicit range information of some sort.

The use of image sequences opens up a variety of new possibilities. For example[14] used an isoplanatically blurred image and a noisy image of the same scene to produce an unblurred denoised image and an estimate of the PSF. Other work on atmospheric turbulence has involved either breaking anisoplanatic images into smaller isoplanatic regions or dewarping the images before deconvolution in order to make the problem tractable,[15] other work has included the deblurring of range-gated lidar.[16] In a similar manner, depth from defocus (DFD) methods measure the change in blurring across a pair of images at different known focal settings in order to determine range to objects in a scene. In this paper we analyse the spectral method of Pentland[17] and the spatial S-Transform method of Subbarao.[18] Once the amount of blurring at each pixel across an image has been determined, the deconvolution problem becomes much simpler. Many other DFD methods have been proposed:[19–21] one of the most promising is the use of coded apertures to make the deconvolution problem more well-posed.[22]

There are several aspects which all previous DFD methods have in common: firstly they process positive intensity images, they produce range estimates from blur by a complex camera calibration process and they require high levels of scene texture/patterning in order to produce correct blur – thus range – estimates. In comparison, the method proposed in this paper uses range images in the complex domain, determines focal parameters from a combination of range and blur data, and can produce blur estimates in areas with no texture as long as there is adequate texture in other regions of the image.

1.2 Overview

We present a new spatially variant parametric two-stage blind deconvolution algorithm. The first stage consists of a Maximum Likelihood based method for determining the focal parameters of a range imager optical system, including the amount of blurring and the distance at which objects are in-focus, from pairs of range images taken at different aperture settings. The algorithm works by reblurring an image taken with a smaller aperture to match an image taken with a large aperture using range data. Whereas previous DFD algorithms have been able to determine the amount of blur at each pixel and thus range from intensity image pairs, since full-field lidar systems also produce range data it is possible to dynamically determine how defocus blur varies with range. Since images taken at small aperture settings are still limited by DOF, we then estimate the amount of blur at each pixel and apply spatially variant deconvolution to extend DOF.

1.3 University of Waikato Range Imager

The examples in the paper were generated using the University of Waikato (UOW)[4] range imager. The UOW range imager uses the time-of-flight (TOF) principle to produce 2 dimensional matrices of range, amplitude and total integrated intensity data. A scene is illuminated with modulated laser light and the system measures the TOF induced phase shift in the illumination modulation, thus the distance to the objects in the scene. While commercial systems use modulated CMOS sensors,[2] modulated image intensifiers currently offer higher spatial resolutions. The UOW range imager has a resolution of 512x512 whereas the highest resolution modulated CMOS sensor has a resolution of 204x204 pixels (PMD PhotonICs®41k-S). In a heterodyne system, either the sensor or an image-intensifier is gain modulated with a signal at a very slightly different frequency from the illumination modulation. This results in the formation of a much lower frequency beat waveform which can be recorded over time by a CCD camera or CMOS device; this beat waveform shape is a correlation of the illumination modulation waveform with the sensor/intensifier response waveform. By taking a temporal Fourier transform of the data, it is possible to determine the phase and amplitude of the beat waveform for each pixel, which is the raw range data we concentrate on in this paper.

1.4 Conventions

We represent convolution by $\star$, spatially variant convolution by $\star_{sv}$. The Fourier transform of a function $f(x)$ is notated in capitalised form $F(u)$.

To ensure clarity, the authors advise viewing images from this paper on a computer screen or printing in colour.

2. RANGE DATA FORMATION

2.1 Range Data

We represent the ideal, unblurred range data by a vector $\wp \in \{\wp_0, \ldots, \wp_{n-1}\}$, where n is the number of pixels in the 2D image and $\wp_i \in \mathbb{C}$ is the the ith pixel. The range datum $\wp_i$ is formed by

$$\wp_i = a_i e^{4\pi j r_i f_m / c}, \tag{1}$$

where $a_i \in \mathbb{R}^+$ is the amplitude, $r_i \in \mathbb{R}^+$ is the range to the object, f_m is the illumination modulation frequency, c is the speed of light. For convenience, we define a function $\wp = \Lambda(a, r)$ that produces range data from vectors of amplitude and range data respectively, which we utilise in section 4.1.

If we consider a model of active signal returns versus range for a pixel, $f_i(r) \in \mathbb{R}^+$, we can regard a measurement at a particular modulation frequency as sampling a particular spatial frequency of the model of signal returns i.e. $\wp_i = F_i(-c/2f_m)$. Conventional processing estimates amplitude and range by

$$a_i = |\Xi_i| = |\wp_i + \epsilon_i + \lambda_c|, \tag{2}$$
$$r_i = \frac{c}{4\pi f_m} \arg(\Xi_i) = \frac{c}{4\pi f_m} \arg(\wp_i + \epsilon_i + \lambda_c), \tag{3}$$

where Ξ_i is the recorded noisy data at pixel i, ϵ_i is noise and λ_c is light from other objects – which results in mixed pixels. Lenses, image intensifiers and reflective sensors tend to result in light scattering within full-field range imagers. This is particularly important if an image has a high dynamic range as range-intensity coupling can result. The resulting errors are one of the dirty secrets of the range-imaging community – it means that range measurements change depending on reflectivity. A flat board can appear to be spread across multiple ranges because it is patterned. We now model the effect that limited depth of field has on range data.

2.2 Depth of Field

The thin lens equation is

$$\frac{1}{f} = \frac{1}{d_o} + \frac{1}{d_i}, \tag{4}$$

where f is the focal length, d_o is the distance to the object projected onto the optical axis and d_i is the distance to the image, similarly projected. Since in most cameras the distance to the imaging plane is fixed, this means that the distance of the correctly formed image from the imaging plane varies across an image depending on the distance along the optical axis of camera to the object. While the focal length can be adjusted, it is impossible for multiple objects at different discrete ranges to all be in perfect focus with a real physical camera, although a single flat object at an angle can be imaged according to the Scheimpflug principle.[23] The properties of defocus blur are well understood; data subject to limited depth-of-field can be regarded as being convolved with a spatially variant point spread function.

There are two commonly used theoretical models for defocus blur: the Gaussian and pillbox models. The Gaussian model is more appropriate for systems where diffraction/aberration effects dominate, particularly at small aperture sizes. The pillbox model is more representative of systems where the aperture size is larger and diffraction effects are less significant. Whether the illumination is polychromatic or monochromatic can also impact on the suitability of a particular model as the superimposed diffraction patterns at different wavelengths in white light can be approximated by a Gaussian. In general, the PSF of the UOW range imager is best approximated by a pillbox model. While most previous work has focussed on Gaussian blurs because of the nice properties in the Fourier domain, most algorithms can be applied to either sorts of blurring with the inclusion of a correction factor.

We notate the blur radius at a particular point as $\sigma_i \in \mathbb{R}^+$; unlike previous work, we avoid the awkward concept of negative blur radii. We use the same symbol, σ, for both Gaussian and pillbox blur radii. If we ignore the projection onto the optical axis by assuming a sufficiently small field-of-view and instead approximate by using radial distance to the object, from geometric optics it can be determined that

$$\sigma_i \propto D \left| \frac{1}{f} - \frac{1}{d_i} - \frac{1}{r_i} \right| \tag{5}$$

where D is the aperture diameter.[21] This approximation allows us to avoid the lens calibrations required to correctly project onto the Z-axis. The equation can be rewritten as

$$\sigma_i = \alpha \left| 1 - \frac{\beta}{r_i} \right|, \tag{6}$$

where $\alpha \in \mathbb{R}^+$ is a scaling constant and $\beta \in \mathbb{R}^+$ is the distance from the camera at which objects are in focus. We consider the tuple $\Theta = (\alpha, \beta, r)$ to be the focal parameters of a system, where $r \in \{r_0, \ldots, r_{n-1}\}$ is the range to each pixel in the scene. This information completely defines our spatially variant blur model; in general we can regard the image formation process as

$$\Xi = \varrho \star_{sv} \sigma(\Theta) + \epsilon, \tag{7}$$

where $\Xi \in \mathbb{C}^n$ is the recorded range data and $\epsilon \in \mathbb{C}^n$ is noise. In theory, given a good estimate of Θ it ought to be possible to restore a full-field lidar image subject to limited depth of field.

In our system a complicating factor is that there are additional optics, which we cannot account for with this model. These optics couple the CCD camera to the image intensifier – since the Z-distance is fixed, this results in convolution by an additional PSF, albeit very small. For the purposes of this paper, we ignore this effect.

3. DEPTH FROM DEFOCUS

3.1 Pentland's Method

While prior work on depth-from-focus extends back much further, the original concept of DFD can be traced back to Pentland.[17] In this work he proposed two methods of depth recovery: one using sharp edges and one using two images taken at different camera aperture settings. The second method is the ancestor of most modern depth-from-defocus algorithms; we now analyse the isoplanatic case. In the Fourier domain convolving an image, y, by a Gaussian PSF with blur radius σ is equivalent to attenuation in the spatial frequency domain, viz

$$z = y \star \sigma \Leftrightarrow Z[u, v] = Y[u, v] e^{2\pi^2 (u^2 + v^2) \sigma^2}, \tag{8}$$

where z is the blurred image. Given two images, z_α and z_β, of the same scene taken with different focal parameters, we can measure the relative attenuation of the spatial frequencies of the image allowing us to estimate the relative amount of blurring between the two images. For blur radii, σ_α and σ_β, the relative blurring is estimated by

$$g = \sigma_\alpha^2 - \sigma_\beta^2 = \frac{1}{4\pi^2 A_R} \sum_{(u,v) \in R_f} \frac{1}{u^2 + v^2} \ln\left(\frac{|Z_\beta[u, v]|}{|Z_\alpha[u, v]|} \right), \tag{9}$$

where R_f is a region of the spatial frequency domain and A_R is the number of samples from the spatial frequency domain. By applying knowledge of the focal parameters of the optical system, it is possible to recover σ_α and σ_β individually. While in the most general case this involves solving a quadratic equation, in this paper we restrict ourselves to the situation in which the focal length is fixed and only the aperture size changes. For aperture diameters D_α and D_β,

$$\sigma_\alpha^2 = \frac{g}{1 - \frac{D_\beta^2}{D_\alpha^2}}, \tag{10}$$

$$\sigma_\beta^2 = \frac{D_\beta^2}{D_\alpha^2} \sigma_\alpha^2. \tag{11}$$

In the conventional case, given the blur relationships from section 2.2, the range to objects in the scene can be determined. The implementation for this paper divides the image up into 16 by 16 pixel regions and Fourier transforms them – only a small subset of the spatial frequencies are used as most tend to be too excessively attenuated to be useful. A blur estimate is produced for each region using the method described above.
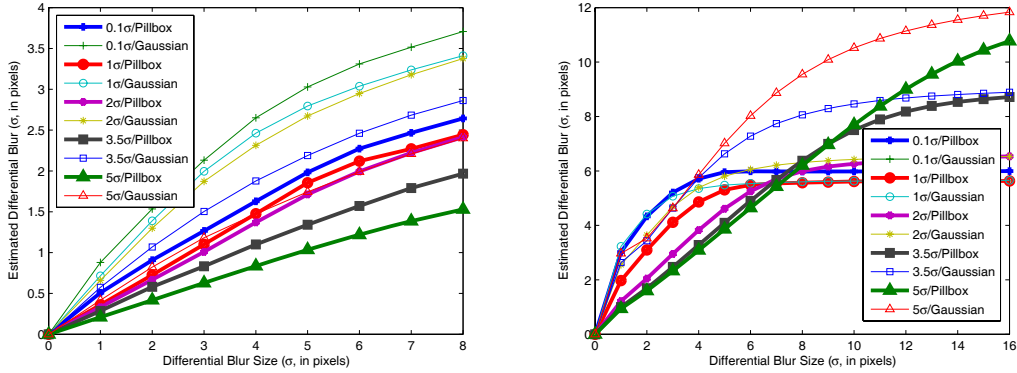
(a) Pentland's frequency domain algorithm response   (b) Subbarao's spatial domain algorithm response($N = 4$)

Figure 1: Depth from defocus algorithm response versus parent data blurring and differential blur type. Key = parent data blurring/differential blur type

3.2 Subbarao's S-Transform Method

Using the S-Transform Subbarao created an alternative method that operates in the spatial domain.[18] The S-Transform models an n-dimensional image as the sum of m-order polynomials. By modelling a local region of an image as a cubic polynomial, Gaussian defocus blur can be expressed as

$$z = y \star \sigma \Leftrightarrow z[x, y] + \frac{\sigma^2}{4} \nabla^2 z[x, y] = y[x, y], \tag{12}$$

where the Laplacian is determined by a Chebyshev polynomial based smoothed differentiation filter. The size of the differentiation filter is parametised by $N \in \mathbb{Z}^+$, giving a filter of size $2N + 1$ pixels. In a similar fashion to eqn. 9, we can estimate g by

$$g^2 = 4 \frac{\sum_{(x,y) \in R_s} (z_\alpha[x, y] - z_\beta[x, y])^2}{\sum_{[x,y] \in R_s} (\nabla^2 z_\alpha[x, y] + \nabla^2 z_\beta[x, y])^2} \tag{13}$$

where R_s defines a spatial domain, rather than a spatial frequency domain. These blur estimates are then denoised via a Parzen windowed histogram over a large spatial region. Although blur estimates can be computed for each individual pixel, they are not statistically independent estimates; for this paper we have produced images at the same resolution as Pentland's algorithm.

3.3 Algorithm Response

In order to test the response of the two DFD algorithms we generated some noiseless test data. These parent data were generated by filtering Gaussian white noise with a Gaussian filter of a particular specified width, thus simulating texture at different spatial scales in the ideal unblurred scene. These parent data were then blurred by another blur function, either pillbox or Gaussian, with a specified width, which we call the differential blur. The parent data and the differentially blurred data were then passed to the algorithms as the unblurred and blurred data respectively and the algorithm response recorded. The results are shown in figs. 1(a) and 1(b). The graph key refers to the Gaussian blur radius for the parent data and the type of differential blur model.

Since texture scale is proportional to blur radius, fig. 1(a) shows that the Pentland algorithm is extremely sensitive to the scale of texture. There is a systematic bias caused by attenuation of the higher spatial frequencies that pulls the estimate of blur towards zero as the accuracy of the logarithmic term is limited by finite precision and noise. A simple example of this effect is if a spatial frequency does not exist (or only noise is present) in the less blurred image, if this spatial frequency is included in the spectral region being analysed then the blur estimate is erroneously pulled towards zero. More advanced methods that dynamically determine the spectral region are possible, but are not dealt with here. Due to the extreme nature
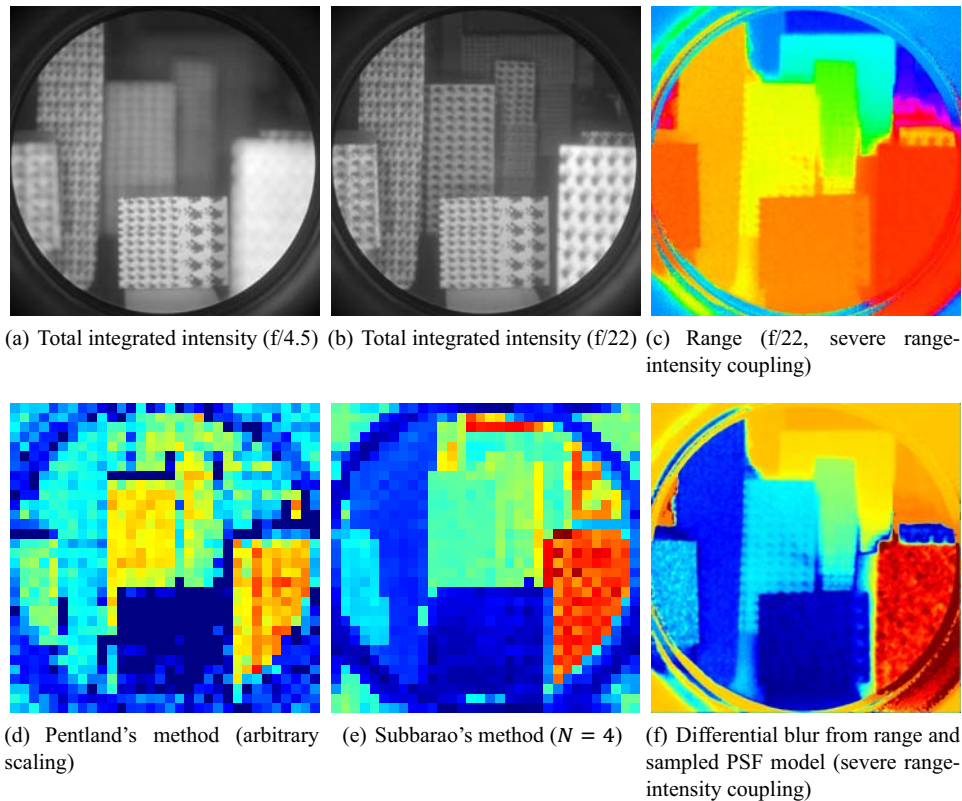
(a) Total integrated intensity (f/4.5)   (b) Total integrated intensity (f/22)   (c) Range (f/22, severe range-intensity coupling)

(d) Pentland's method (arbitrary scaling)   (e) Subbarao's method ($N = 4$)   (f) Differential blur from range and sampled PSF model (severe range-intensity coupling)

Figure 2: Scene A depth from defocus results. Total integrated intensity images use $\gamma = 0.1$.

of this effect, Pentland's algorithm appears to require the addition of an arbitrary scaling coefficient for comparison to Subbarao's algorithm.

Fig. 1(b) shows that the Subbarao Algorithm has a linear response in certain cases, and a non-linear response in others; much of this is due to the spatial scale of the texture. For this experiment, $N = 4$, meaning that the Laplacian was calculated using 9 pixel derivative filters. If the size of the derivative filter does not match the spatial scale of the texture, the algorithm appears to give erroneous results. Parent data which have a much larger initial blur results in a much more linear response gradient – which appears to be close to $\frac{6}{8}$ in the pillbox case and $\frac{8}{6}$ in the Gaussian case.

3.4 Depth From Defocus Results

Three range images were taken of different scenes, each with a different focal length. The complex domain range images were then analysed using the two DFD algorithms and compared to blur estimates created using range data and the known focal parameters. The only difference between processing intensity images and processing complex domain images being that the absolute value of the Laplacian and relative error are taken in the Subbarao algorithm. The results are shown in figs. 2, 3 and 4.

Scene A has an extremely high dynamic range; the total integrated intensity images have been highly gamma compressed using $\gamma = 0.1$. This high dynamic range results in a problem known as range-intensity coupling,[7] where light scattered within the range imager results in ranges that change depending on the reflectivity of the target object. Fig. 2(c) shows range information, where red represents objects closer to the camera and blue/purple objects farther away. Figs. 2(a) and 2(b) show how decreasing the aperture size increases the depth of field, however even at the very small f/22 the depth of field is still finite and some objects appear quite blurry. Passing the range image of the scene to the DFD algorithms gives figs. 2(d) and 2(e). Both algorithms struggle to accurately estimate the blur scale for objects in the background, however Subbarao's produces visibly better output. Pentland's algorithm produced artefacts at several object edges, suffered more
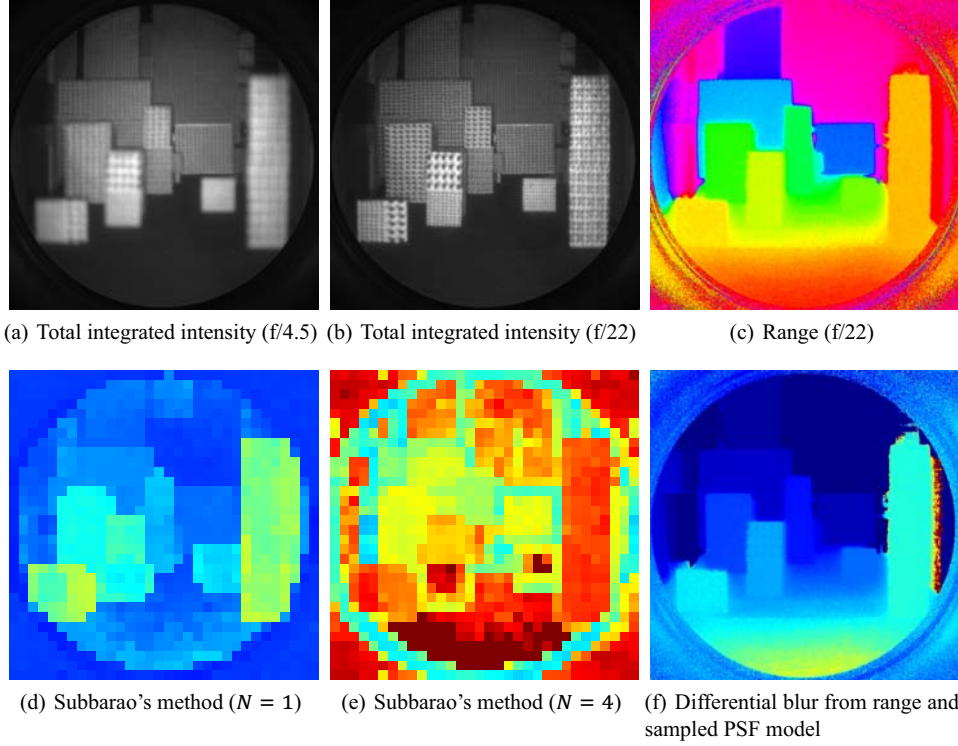
(a) Total integrated intensity (f/4.5) (b) Total integrated intensity (f/22)   (c) Range (f/22)

(d) Subbarao's method ($N = 1$)   (e) Subbarao's method ($N = 4$)   (f) Differential blur from range and
                                                                        sampled PSF model

Figure 3: Scene B depth from defocus results. Total integrated intensity images use $\gamma = 0.5$.

from noise and while retaining the correct relative blur ordering across objects, did not maintain the same ratios. Subbaro's algorithm however performed creditably compared to the estimated actual blurring in fig. 2(f) (blue represents less blurring, red more blurring). The blur estimates are slightly corrupted due to range-intensity coupling.

Scene B has a much smaller dynamic range than scene A; as a result the range data (fig. 3(c)) suffers much less range-intensity coupling. However there are still minor erroneous ranges to the right/top of the rightmost box in the image. Figs. 3(d) and 3(e) show how the value of N affects the quality of range data returned by Subbarao's algorithm. We are currently unsure whether there is a good automated way to estimate the optimal setting. Notably, the textureless table region does not produce accurate blur estimates.

In scene C no value of N was found to produce correct blur estimates from Subbarao's algorithm. Fig. 4(e) shows a blur estimate produced by the Maximum Likelihood method discussed in the next section.

4. DETERMINATION OF FOCAL PARAMETERS

4.1 Theory

Photon shot noise is Poisson distributed; in most imaging situations the mean intensity, μ_i, is sufficiently high that the Poisson distribution can be accurately modelled by a Gaussian distribution where $\sigma_i^2 = \mu_i$. In AMCW lidar noise variance is not necessarily proportional to amplitude ($\sigma_i^2 \not\propto |\varrho_i|$) due to contributions from ambient lighting and partial cancellation at mixed pixels. Instead the total integrated intensity at a pixel, χ_i, takes into account all light sources within a pixel

$$\chi_i = \varphi_i + \int f_i(r)dr, \qquad (14)$$

where φ_i is the contribution from ambient light at pixel i. This allows us to model the noise at a pixel, ϵ_i, as having a zero centred circularly symmetric complex normal distribution, $\epsilon_i \sim \mathcal{N}_{\mathbb{C}}(0, \chi_i)$, resulting in the absolute error at a pixel, $|\epsilon_i|$, being Rayleigh distributed. In general, total integrated intensity images look very similar to amplitude images unless

(a) Total integrated intensity (f/4.5)   (b) Total integrated intensity (f/22)   (c) Range (f/22)

(d) Subbarao's method ($N = 4$)   (e) ML focal parameter determination   (f) Differential blur from range and sampled PSF model
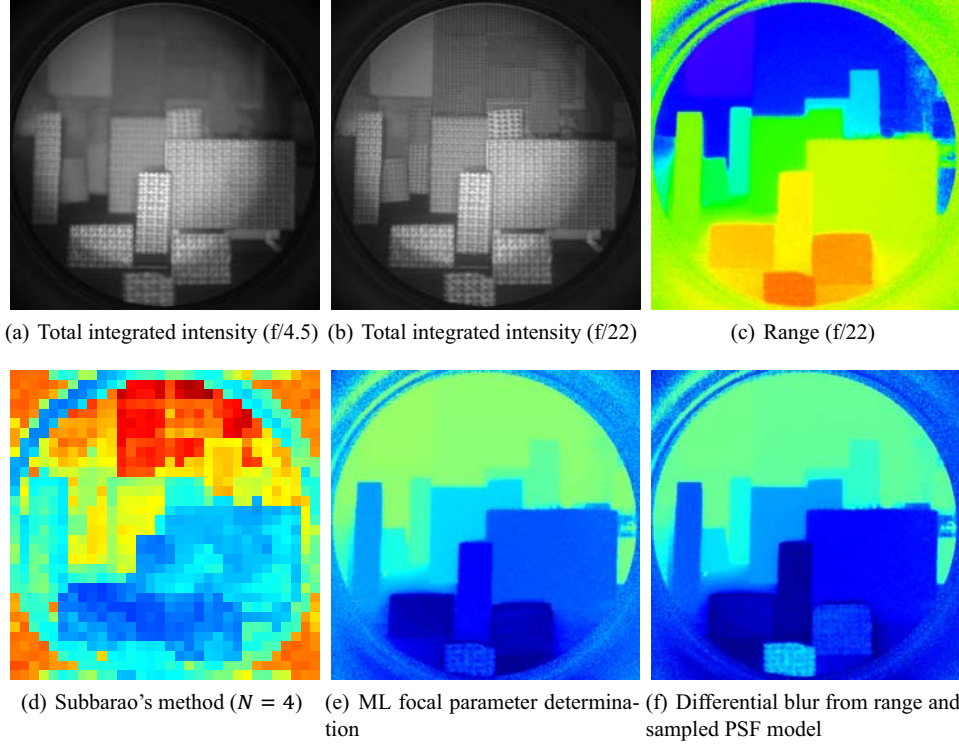
Figure 4: Scene C depth from defocus results. Total integrated intensity images use $\gamma = 0.5$.

there is significant ambient light or partial cancellation at mixed pixels (all data for this paper were generated in a darkened room).

If we model z_α and z_β as a convolution of ideal unblurred range data, $\widehat{y} = \Lambda(a_\emptyset, r_\emptyset)$, we can express the log-likelihood of particular focal parameters and unblurred data as

$$
\begin{aligned}
\mathcal{L}(\alpha, \beta, r_\emptyset, a_\emptyset | z_\alpha, z_\beta, \Sigma_\alpha, \Sigma_\beta) \quad = \quad & const - \frac{1}{2}(z_\alpha - h(\alpha, \beta, r_\emptyset)\Lambda(a_\emptyset, r_\emptyset))^* \Sigma_\alpha^{-1}(z_\alpha - h(\alpha, \beta, r_\emptyset)\Lambda(a_\emptyset, r_\emptyset)) \\
& - \frac{1}{2}(z_\beta - h(\alpha\frac{D_\beta}{D_\alpha}, \beta, r_\emptyset)\Lambda(a_\emptyset, r_\emptyset))^* \Sigma_\beta^{-1}(z_\beta - h(\alpha\frac{D_\beta}{D_\alpha}, \beta, r_\emptyset)\Lambda(a_\emptyset, r_\emptyset)) \quad (15)
\end{aligned}
$$

where Σ_α and Σ_β are diagonal matricies consisting of the values from χ_α and χ_β respectively, and $h \colon (\mathbb{R}^+, \mathbb{R}^+, \mathbb{R}^{+n}) \to \mathbb{R}^{n \times n}$ is a function that returns the linear transformation matrix corresponding to a spatially variant convolution such that $h(\alpha, \beta, r)z = z \star_{sv} \sigma(\alpha, \beta, r)$.

A single step approach might attempt to maximise this, most probably with the addition of some sort of regularisation. Instead we assume that the least blurred range estimates r_β are correct and approximate the relationship between z_α and z_β as a spatially variant convolution parametised in the same fashion as before, giving

$$
\begin{aligned}
\mathcal{L}(\alpha, \beta | z_\alpha, z_\beta, \Sigma_\alpha, \Sigma_\beta) \quad = \quad & const - \frac{1}{2}\ln(\det(\Sigma_\alpha + h(\tau\alpha, \beta, r_\beta)\Sigma_\beta h^*(\tau\alpha, \beta, r_\beta))) \quad (16) \\
& - \frac{1}{2}(z_\alpha - h(\tau\alpha, \beta, r_\beta)z_\beta)^*(\Sigma_\alpha + h(\tau\alpha, \beta, r_\beta)\Sigma_\beta h^*(\tau\alpha, \beta, r_\beta))^{-1}(z_\alpha - h(\tau\alpha, \beta, r_\beta)z_\beta),
\end{aligned}
$$

where τ is a constant conversion factor. In an ideal case τ would be confirmed experimentally, however for the purposes of this paper we have assumed that $\tau = \frac{D_\alpha - D_\beta}{D_\alpha}$. By approximating the error covariance by $2\Sigma_\alpha$ we can further simplify the problem, giving a constant, diagonal covariance matrix. An estimate of the remaining focal parameters (α, β) is found by
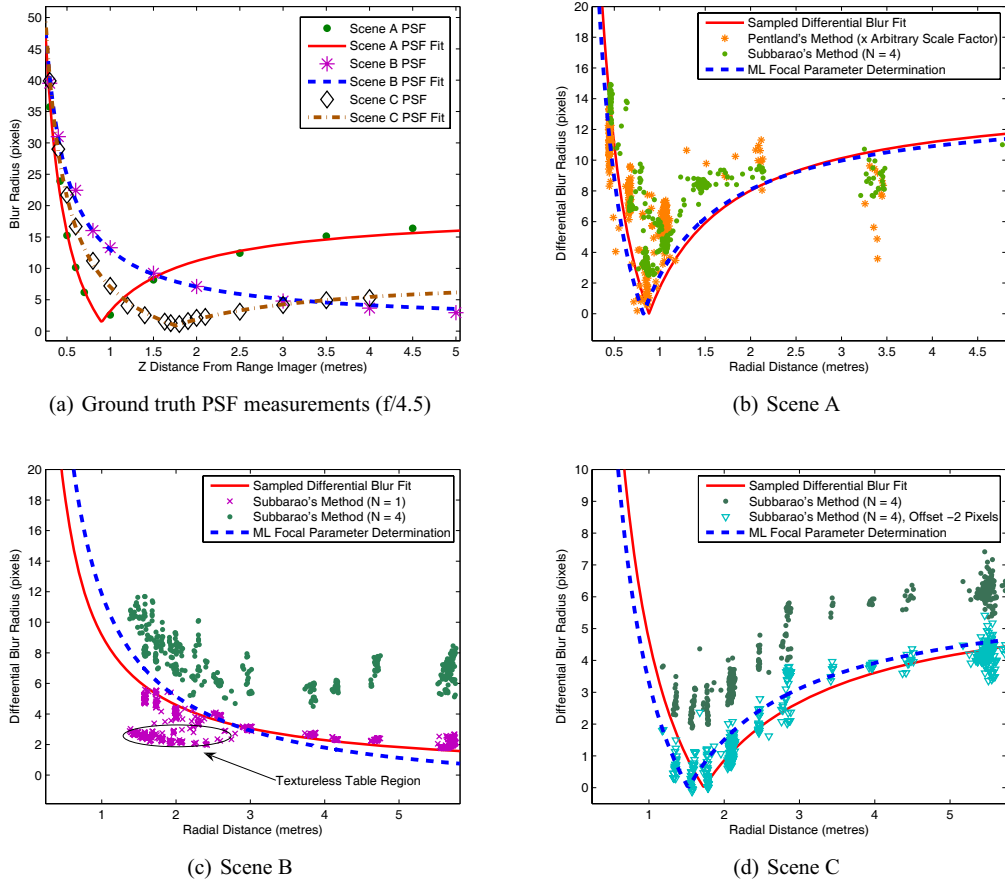
(a) Ground truth PSF measurements (f/4.5)

(b) Scene A

(c) Scene B

(d) Scene C

Figure 5: ML focal parameter determination compared to depth from defocus methods and ground truth. Sampled PSF at f/22 not shown.

numerically maximising

$$\mathcal{L}(\alpha, \beta | z_\alpha, z_\beta, \Sigma_\alpha) = const - \frac{1}{2}(z_\alpha - h(\tau\alpha, \beta, r_\beta)z_\beta)^*(2\Sigma_\alpha)^{-1}(z_\alpha - h(\tau\alpha, \beta, r_\beta)z_\beta). \tag{17}$$

Given an estimate of the focal parameters, a spatially variant deconvolution operation can be performed as a second step.

4.2 Focal Parameter Determination Results

Results from application of the ML Focal Parameter Determination algorithm to the three scenes are shown in fig. 5. Fig. 5(a) shows some ground truth measurements of the PSF blur radius as a function of Z distance from the range imager; measurements at f/22 and f/4.5 were combined in order to estimate the differential blur which is used as a reference in each of the following graphs. Fig. 5(b) shows how the DFD results compare to the actual differential blur and the results of ML focal parameter determination. In this case, both DFD methods suffer from noise issues – although the range information is corrupted slightly by range-intensity coupling, each discrete object should still have a consistent blur level, which they do not. Both methods do appear to identify roughly the correct in-focus range. The ML estimate of focal parameters very closely matches the sampled differential blurring. Fig. 5(c) contains the results from Scene B. The graph shows how N affects the quality of blur estimates from Subbarao's algorithm; in the $N = 1$ case the blur estimates are very close to the actual differential blurring everywhere except for the featureless table region. Scene C results in an interesting problem with the Subbarao algorithm; fig. 5(d) shows that the blur estimates appear to be offset by 2 pixels from the correct values. The ML focal parameter determination algorithm, while not producing inordinately false results, has misestimated the in-focus distance by ~ 30 cm.
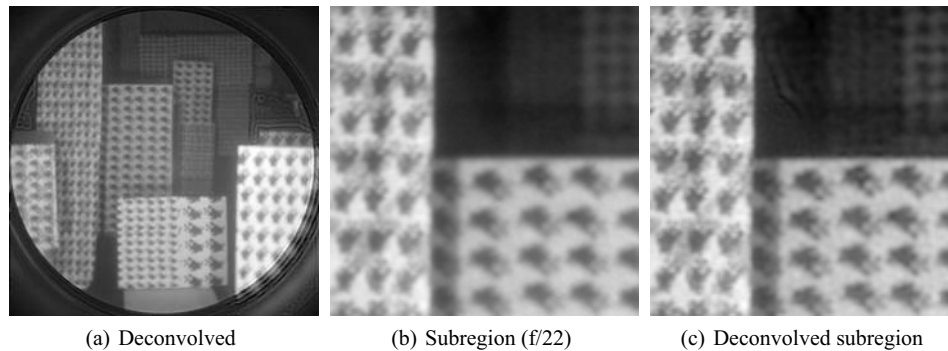
| (a) Deconvolved | (b) Subregion (f/22) | (c) Deconvolved subregion |

Figure 6: Scene A deconvolved total integrated intensity image

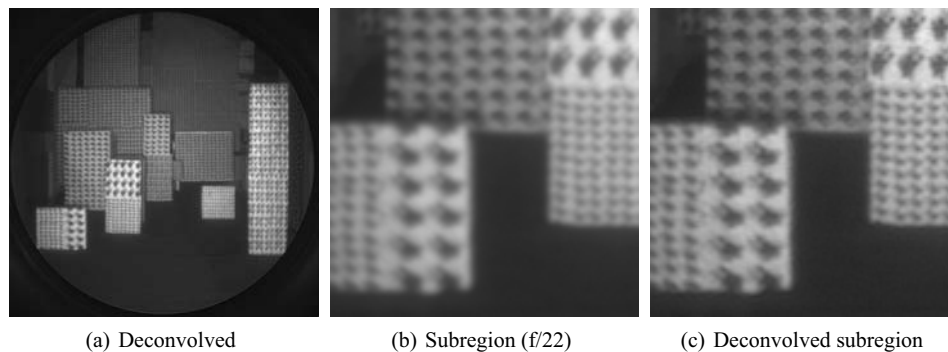| (a) Deconvolved | (b) Subregion (f/22) | (c) Deconvolved subregion |

Figure 7: Scene B deconvolved total integrated intensity image

5. DECONVOLUTION OF LIDAR IMAGES

5.1 Theory

Deconvolution was performed by a spatial derivative regularised, variant Landweber algorithm.[24] The Landweber algorithm is an iterative Maximum Likelihood algorithm in the case of additive Gaussian noise, although it can be applied to other noise distributions. Each iteration is defined by

$$\widehat{y_i} = \widehat{y_{i-1}} + \zeta(h^T(\alpha, \beta, r_\beta)(z_\beta - h(\alpha, \beta, r_\beta)\widehat{y_{i-1}}) - \lambda L\widehat{y_{i-1}}), \qquad (18)$$

where $\widehat{y_i}$ is the ith estimate of the deconvolved data, ζ is gain, L is a Laplacian filter and λ is a regularisation parameter. We implement the spatially variant convolution by generating pillbox blur models at blur radii intervals of one-third of

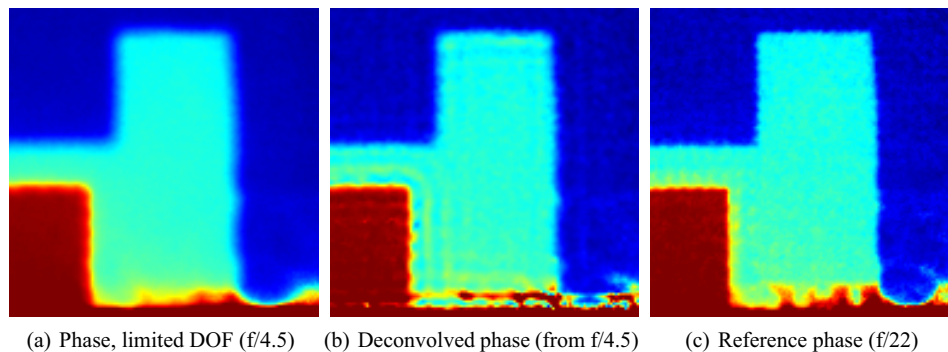| (a) Phase, limited DOF (f/4.5) | (b) Deconvolved phase (from f/4.5) | (c) Reference phase (f/22) |

Figure 8: Scene C phase of deconvolved complex domain range image subregion

a pixel and modelling intermediate blur radii by linear interpolation. This leads to efficient computation, although linear interpolation is more suited to smooth functions. The transpose of the spatially variant convolution matrix, $h^T(\alpha, \beta, r_\emptyset)$ corresponds to a spatially variant correlation operation which can be implemented in a similar manner to the spatially variant convolution. A good overview of the implementation of spatially variant convolutions is provided by.[25] Unfortunately, the finite size of the image intensifier crops the images resulting in edge effects in deconvolved images.

5.2 Results

Figs. 6 and 7 show the resultant deconvolved total integrated intensity images from scenes A and B respectively. Using total integrated intensity images instead of amplitude allows us to avoid any confounding effects of phase in the complex domain. There is an improvement in depth of field, but with a difficult trade-off between recovering detail and noise amplification. The high level of gamma compression in fig. 6(a) shows significant artefacts in regions where the blur scale has been misidentified due to mixed pixels. For example, there is a region on the far right above the very bright defocussed object where severe blur misestimation due to mixed pixels has resulted in irregular wave patterns.

Fig. 8 shows the results from attempting to deconvolve a complex domain range image. An image taken at f/4.5 (fig. 8(a)) is highly blurred due to limited depth of field – thus forming a large band of mixed pixels around the edge of each object, each mixed pixel being an erroneous range measurement. Deconvolution reduces the size of the regions containing mixed pixels (fig. 8(b)), however it also results in ringing – contributed to by the highly bandlimited nature of large pillbox PSFs (Gibbs oscillations). A comparison image at f/22 is given in fig. 8(c). While there is some PSF misestimation, similar ringing occurs if simulated isoplanatic data with a hard discontinuity is generated and deconvolved with the correct PSF. Designing a good deconvolution algorithm involves determining how to synthesise these missing frequencies, without which the image cannot be properly restored. These oscillations are very problematic; because the image is in the complex domain, it contains both amplitude and range information. The range information is best regularised to ensure sharper edges but less detail, yet at the same time we need to retain (and improve) the highly textured amplitude information. This suggests that it may be necessary to develop a more advanced regularisation technique: for example, one that combines Harr wavelet based regularisation in the phase/range domain and spatial derivative regularisation in the amplitude domain. We leave this to future work.

Apart from the regularisation trade-off, the biggest problem with the current algorithm is mixed pixels. Range-intensity coupling due to scattered light introduces nonexistent discontinuities in blur estimates for patterned surfaces and in dark regions can result in completely erroneous blur estimates due to wraparound effects. Edge induced mixed pixels cause similar problems, although are more predictable – resulting in artefacts around the edges of objects. Fundamentally, there is no simple solution to this problem because each individual pixel can contain multiple component signal returns, each with a different range and thus subject to a different amount of blur.

6. CONCLUSION

We have presented a new method for parametric spatially variant blind deconvolution of full-field lidar image pairs taken at different aperture settings. The method allows estimation of focal parameters without prior calibration of the optical setup and produces blur estimates which have better spatial resolution and less noise than previous depth from defocus blur measurement algorithms under most circumstances. The estimated focal parameters were applied to deconvolution of total integrated intensity images improving depth of field. Application to complex domain images of multiple objects requires the development of more appropriate regularisation methods due to the trade-off between different levels of amplitude and range texture in the same image. There remain issues with correct PSF determination, especially around the edges of objects and in regions affected by range-intensity coupling.

ACKNOWLEDGMENTS

REFERENCES

[1] Blais, F., "Review of 20 years of range sensor development," Videometrics VII, Proc. SPIE vol. 5013, 62-76 (2003).

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

[3] Ringbeck, T. and Hagebeuker, B., "A 3D time of flight camera for object detection," Proc. of Optical 3-D Measurement Techniques, (2007).

[4] 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).

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

[6] 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).

[7] Godbaz, J. P., Cree, M. J., and Dorrington, A. A., "Undue influence: Mitigating range-intensity coupling in AMCW 'flash' lidar using scene texture," Proc. IVCNZ 2009, (2009).

[8] Larkins, R. L., Cree, M. J., Dorrington, A. A., and Godbaz, J. P., "Surface projection for mixed pixel correction," Proc. IVCNZ 2009, (2009).

[9] Godbaz, J. P., Cree, M. J., and Dorrington, A. A., "Multiple return separation for a full-field ranger via continuous waveform modelling," Image Processing: Machine Vision Applications II, Proc. SPIE vol. 7251, (2009).

[10] Goshtasby, A. A., "Fusion of multifocus images to maximize image information," Proc. SPIE vol. 6229, (2006).

[11] Kundur, D. and Hatzinakos, D., "Blind image deconvolution," IEEE Signal Processing Magazine, 13(3), 43–64, (1996).

[12] Boden, A. F., Redding, D. C., Hanisch, R. J., and Mo, J., "Massively parallel spatially-variant maximum likelihood image restoration," Proc. Astronomical Data Analysis Software and Systems V, 131–134 (1996).

[13] Hardy, J. W., [Adaptive Optics for Astronomical Telescopes] (July 1998).

[14] Yuan, L., Sun, J., Quan, L., and H., S., "Image deblurring with blurred/noisy image pairs," Proc. SIGGRAPH07, (2007).

[15] Fraser, D., Lambert, A., Jahromi, M. R. S., Tahtali, M., and Clyde, D., "Anisoplanatic image restoration at ADFA," Proc. Digital Image Computing: Techniques and Applications (2003).

[16] MacDonald, A. and Cain, S. C., "Parameterized blind deconvolution of laser radar imagery using an anisoplanatic optical transfer function," Optical Engineering 45(11) (2006).

[17] Pentland, A. P., "A new sense for depth of field," IEEE Trans. Pattern Anal. Mach. Intell. 9(4), 523–531 (1987).

[18] Subbarao, M. and Surya, G., "Depth from defocus: a spatial domain approach," Int. J. Comput. Vision 13(3), 271–294 (1994).

[19] Subbarao, M., "Parallel depth recovery by changing camera parameters," in Proc. Conference on Computer Vision, (1988).

[20] Rajagopalan, A. N. and Chaudhuri, S., "Space-variant approaches to recovery of depth from defocused images," Comput. Vis. Image Underst. 68(3), 309–329 (1997).

[21] Chaudhuri, S. and Rajagopalan, A. N., [Depth from defocus: a real aperture imaging approach], Springer Verlag (1999).

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

[23] Scheimpflug, T., "Improved method and apparatus for the systematic alteration or distortion of plane pictures and images by means of lenses and mirrors for photography," GB Patent No. 1196, (1904).

[24] Landweber, L., "An iteration formula for fredholm integral equations of the first kind," Am. J. of Math. 73(3), 615–624 (1951).

[25] Nagy, J. G. and O'Leary, D. P., "Fast iterative image restoration with a spatially-varying psf," Tech Report 270642, University of Maryland at College Park (1997).