

# Optimal Camera Parameters for Depth from Defocus

Fahim Mannan and Michael S. Langer  
 School of Computer Science, McGill University  
 Montreal, Quebec H3A 0E9, Canada.  
 {fmannan, langer}@cim.mcgill.ca

## Abstract

*Pictures taken with finite aperture lenses typically have out-of-focus regions. While such defocus blur is useful for creating photographic effects, it can also be used for depth estimation. In this paper, we look at different camera settings for Depth from Defocus (DFD), the conditions under which depth can be estimated unambiguously for those settings and optimality of different settings in terms of lower bound of error variance. We present results for general camera settings, as well as two of the most widely used camera settings namely, variable aperture and variable focus. We show that for variable focus, the range of depth needs to be larger than twice the focal length to unambiguously estimate depth. We analytically derive the optimal aperture ratio, and also show that there is no single optimal parameter for variable focus. Furthermore we show how to choose focus in order to minimize error variance in a particular region of the scene.*

## 1. Introduction

Depth from Defocus work by estimating defocus blur at every pixel from one or more defocused images. Any finite aperture or non-pinhole camera will produce defocus. This is more prominent when pictures are taken with long focal length and wide aperture lenses. In such cases, when the lens is focused on a subject, everything outside the subject's distance is defocus blurred. This has been used by photographers to shift the viewer's attention or even to produce background effects like Bokeh<sup>1</sup>. From a mathematical viewpoint, the amount of blur depends on the camera parameters namely focal length ( $f$ ), f-number ( $N$ ), and focus position, and the depth of the scene. When this relationship is bijective, we can recover depth by measuring the defocus blur. Pentland first proposed to estimate depth from defocused images [10] in the 1980s. There is a closely related method known as Depth from Focus (DFF) (e.g. [5, 6])

<sup>1</sup>Bokeh is the blur pattern created by bright out-of-focus scene points.

where a set of pictures are taken by changing the focus at discrete intervals. Depth at a pixel/region is determined by finding the focus position where contrast is highest. DFF requires a large number of images and the depth inference is a forward problem where the assigned depth is equal to the focus distance. DFD on the other hand typically requires one or two images and the depth is estimated by solving an inverse problem.

There are a number of pinhole model based methods for estimating depth from single and multiple images, e.g. Structure from Motion (SfM), Stereo, etc. However, defocus blur is unavoidable because there is no true pinhole camera. Furthermore, today's high resolution sensors are more sensitive to defocus blur. So if there is unavoidable defocus blur then we might as well use it. In terms of depth resolution Schechner and Kiryati [14] showed that there are no inherent limitations of DFD in discriminating depth. Compared to methods like SfM, DFD can produce a dense depthmap with only a few images. However DFD does require careful camera calibration (e.g. focus distance and PSF calibration) as well as very good alignment between multiple defocused images.

For getting the most out of DFD, we need to know under what conditions we can estimate depth and what camera settings give us the best performance. The main contributions of this paper are: 1) Finding the conditions under which depth can be estimated unambiguously from defocused images, 2) analytically deriving the lower bound of error variance for different camera settings and, 3) experimentally verifying the model using synthetic and real defocused images. The paper is organized as follows. Sec. 2 gives an overview of the relevant background in DFD, discusses related works and our contributions. Sec. 3 presents three general categories of camera parameters and derives their operating range or limits of unambiguous depth estimation. Sec. 4 investigates the theoretical lower bound of error variance for blur, inverse depth, and depth for different lens settings. These theoretical results are experimentally verified in Sec. 5. Finally Sec. 6 concludes the paper with a summary and some possible applications of our work.

## 2. Background

In this section, we present some of the fundamental ideas behind relative blur based Depth from Defocus that this paper relies on. We start by presenting the blurred image formation model followed by how a pair of defocus blurs are related. Finally we look at modeling DFD for a pair of defocused images using the relative blur model.

### 2.1. Blurred Image Formation

An ideal pinhole camera makes a sharp projection of the scene onto the image plane. However a finite aperture camera with a thin lens can only focus at a single plane (parallel to the sensor plane) in the scene. The amount of defocus outside this plane depends on a number of camera parameters and can be derived using the thin lens model.

Fig. 1 shows how a scene point at distance  $u$  is imaged by a lens of focal length  $f$  and aperture diameter  $A$ . Light rays emanating from the scene point fall on the lens and converge at distance  $v$  on the sensor side of the lens. The relationship between these variables is specified by the thin lens model as:

$$\frac{1}{u} + \frac{1}{v} = \frac{1}{f}. \quad (1)$$

If the imaging sensor is at distance  $s$  from the lens then the imaged scene point creates a circular blur pattern (the exact shape will depend on the shape of the aperture) of radius  $r$  as shown in the figure. The thin lens model (Eq. 1) and similar triangles from Fig. 1 give the radius of the blur in pixels:

$$\sigma = \rho r = \rho \frac{fs}{2N} \left( \frac{1}{f} - \frac{1}{u} - \frac{1}{s} \right). \quad (2)$$

In the above equation the ratio of focal length ( $f$ ) and f-number ( $N$ ) is used instead of the aperture (i.e.  $A = f/N$ ). The variable  $\rho$  is used to convert from physical to pixel dimension. In the rest of this paper we will use  $\sigma$  to denote blur radius in pixels. Note that the blur can be positive or negative depending on which side of the focus plane a scene point resides. However for circularly symmetric aperture the sign of the blur has no effect in the blurred image formation process.

Eq. 2 shows that the blur radius is a linear function of inverse depth. The units of inverse depth is  $m^{-1}$ , or *diopeters* (D). Figures 2a and 2b show how the radius (solid lines) varies with distance for different optical settings (more details about the plots are in Sec. 2.2).

A blurred image is modeled as a convolution of a focused image with a point spread function (PSF) determined by the blur radius. If the PSF function at pixel  $(x, y)$  has radius  $\sigma$  and written as  $h(x, y, |\sigma|)$  (assuming a circularly symmetric PSF we can ignore the sign of the blur and use  $|\sigma|$ ), then for a focused image  $I_0$ , the observed defocused image  $I$  is:

$$I(x, y) = I_0(x, y) * h(x, y, |\sigma|) \quad (3)$$

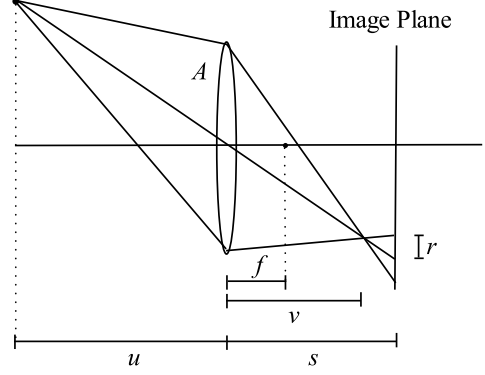


Figure 1: Defocus blur formation

Ideally the PSF is a pillbox function (i.e. a cylindrical shaped function) with radius  $\sigma$ . However, because of diffraction and other unmodeled characteristics of the optical system, the blur kernel is often modeled as a Gaussian with spread  $\sigma_G = \sigma/\sqrt{2}$ , which is what we also use in this paper.

### 2.2. Linearly related defocus blurs

Given a pair of images captured using different optical parameters, the blurs  $\sigma_1$  and  $\sigma_2$  at a corresponding pixel<sup>2</sup> are linearly related (using Eq. 1 and 2) as follows[18]:

$$\sigma_2 = \alpha \sigma_1 + \beta \quad (4)$$

where

$$\alpha \equiv \frac{N_1 f_2 s_2}{N_2 f_1 s_1}, \quad \beta \equiv \rho \frac{f_2 s_2}{2N_2} \left( \frac{1}{f_2} - \frac{1}{f_1} + \frac{1}{s_1} - \frac{1}{s_2} \right). \quad (5)$$

There are two common DFD configurations namely: variable aperture (2A) and variable focus (2F). In the variable aperture case, a pair of images are taken with the same focal length (i.e.  $f_1 = f_2$ ) and focus setting (i.e.  $s_1 = s_2$ ) but with different f-numbers (wlog we assume  $N_1 > N_2$ ), so:

$$\alpha = \frac{N_1}{N_2} > 1, \quad \beta = 0. \quad (6)$$

In the variable focus case, the focal length and f-number are fixed ( $A_1 = A_2$ ) and the focus setting is varied, so  $s_1 \neq s_2$ . If the near and far focal plane distances are  $u_1$  and  $u_2$  respectively, then for  $u_1 < u_2$ , we have  $s_1 > s_2$  and

$$\alpha = \frac{s_2}{s_1} < 1, \quad \beta = \frac{\rho f}{2N} (\alpha - 1) < 0. \quad (7)$$

If the camera settings in each image are known, then  $\alpha$  and  $\beta$  can be estimated. Fig. 2 shows the absolute and relative

<sup>2</sup>The images are geometrically related by a magnification factor  $\frac{s_2}{s_1}$ . Correspondences can be found if this magnification factor is known.

blur in the above two cases. In both cases, the problem of Depth from Defocus is to estimate the blur  $\sigma_1$  (or  $\sigma_2$ ) at each point in the image, and then estimate the depth using the thin lens equation, i.e. Eq. (1). To estimate  $\sigma_1$  or  $\sigma_2$  at each point, one estimates the relative blur (Eq. 13). From Fig. 2 we can see that more than one distance can have the same absolute and relative blur which can result in depth ambiguity. In Sec. 3 we show the conditions under which the relative blur uniquely determines the blurs  $\sigma_1$  and  $\sigma_2$ , and in turn depth  $u$ .

### 2.3. Relative Blur Model

The main objective in DFD is to estimate the blur radius at each pixel and in turn estimate depth using Eq. 2. Many different methods were proposed for this purpose using one or more images, some of which we will review in Sec. 2.4. For now we only consider the problem of estimating depth from two defocused images using the relative blur model. In general, the blur in corresponding regions of two defocused images will vary with the depth and camera parameters. Therefore a region in the image will either be blurrier in one image but sharper in the other or vice versa. For example in the variable focus example shown in Fig. 2b, a point at 1.4D (0.7m) is sharper in one image (red) and another point at 1D is sharper in the other (green). To simplify the explanation of the relative blur model, let us assume a fronto-parallel plane being imaged with two different camera parameters. In this case, one of the images will be sharper ( $I_S$ ) and the other blurrier ( $I_B$ ). The observed image can be modeled as convolution of the hypothetical sharp image  $I_0$  with an appropriate PSF  $h$  and additive noise. Therefore the observed sharper and blurrier images are:

$$I_S = I_0 * h_S + n_S \quad (8)$$

$$I_B = I_0 * h_B + n_B. \quad (9)$$

Let  $h_R$  be the *relative blur* which under the Gaussian PSF assumption is the amount by which the sharper ( $h_S$ ) PSF is blurred to get the blurrier ( $h_B$ ) PSF, that is,

$$h_B = h_S * h_R. \quad (10)$$

Now we make a simplifying assumption that the noise in the sharp image is negligible and let:

$$I_0 * h_B \approx I_S * h_R \quad (11)$$

$$= I_0 * h_S * h_R + n_S * h_R \quad (12)$$

The equation holds with equality when  $n_S = 0$  and very closely approximated as  $h_R$  gets large.

Let the blur radii of the kernels  $h_B$  and  $h_S$  be  $\sigma_B$  and  $\sigma_S$  respectively. The radius  $\sigma_R$  of the relative blur  $h_R$  in Eq. 10 is:

$$\sigma_R \equiv \sqrt{\sigma_B^2 - \sigma_S^2}. \quad (13)$$

The relative blur estimation problem can be written as the following optimization problem:

$$\arg \min_{h_R} \|I_B - I_S * h_R\|_2^2 \quad (14)$$

Eq. 14 can be solved using least squares. However, this assumes we know in advance which image is the blurrier one at each pixel. In the more general case given two defocused images  $I_1$  and  $I_2$ , we would have to decide which one to blur and by what amount (i.e.  $\sigma_R$ ). In such cases, it is more convenient to use signed relative blur. We can choose the sign of  $\sigma_R$  to be negative when  $|\sigma_1| > |\sigma_2|$  and positive otherwise. Therefore when the sign of relative blur is negative  $\sigma_B = \sigma_1$  and  $\sigma_S = \sigma_2$  and vice-versa when the sign is positive. Experiments for this paper were performed using this approach.

### 2.4. Related Work and Our Contributions

Our emphasis in this paper is not on DFD algorithms per se, but rather on the computational constraints on the DFD problem that apply to any algorithm based on relative blur. Here we briefly review several DFD algorithms. Our goal is to summarize some of the key DFD contributions and, more specifically, to distinguish some of the algorithms by aspects such as the camera configurations used, and whether they have addressed issues like uniqueness and lower bounds on estimates which is what our paper is mainly about.

In the first DFD paper, Pentland [10] used a pair of very small (almost pinhole) and large aperture images, and estimated relative blur using inverse filtering in the frequency domain. Later, Subbarao [18] relaxed the requirement of pinhole aperture and proposed a method for more general camera parameters. He also investigated some of the special cases under which unambiguous depth estimates can be obtained. Ens and Lawrence [2] proposed a more robust relative blur estimation approach in the spatial domain. In their method they first calibrated relative blur kernels for different depth and camera parameters, and used that for depth estimation for arbitrary scenes with the calibrated camera parameters. They only showed results for variable aperture. Subbarao and Surya [19] proposed a spatial domain transformation method to find the relative blur from two defocused images. They also looked at the requirements for unambiguous depth estimation. Watanabe et al. [22] used variable focus with frequency domain filtering. Their approach required the entire scene to be between the two focal planes. Favaro and Soatto [3, 4] formulated an alternative filtering model which uses orthogonal operators rather than convolution for estimating relative blur. These orthogonal operators are either analytically derived or learned. Other works have considered Variational [4], MAP-MRF [12, 9, 8], or Linear Programming [15] based formulations.

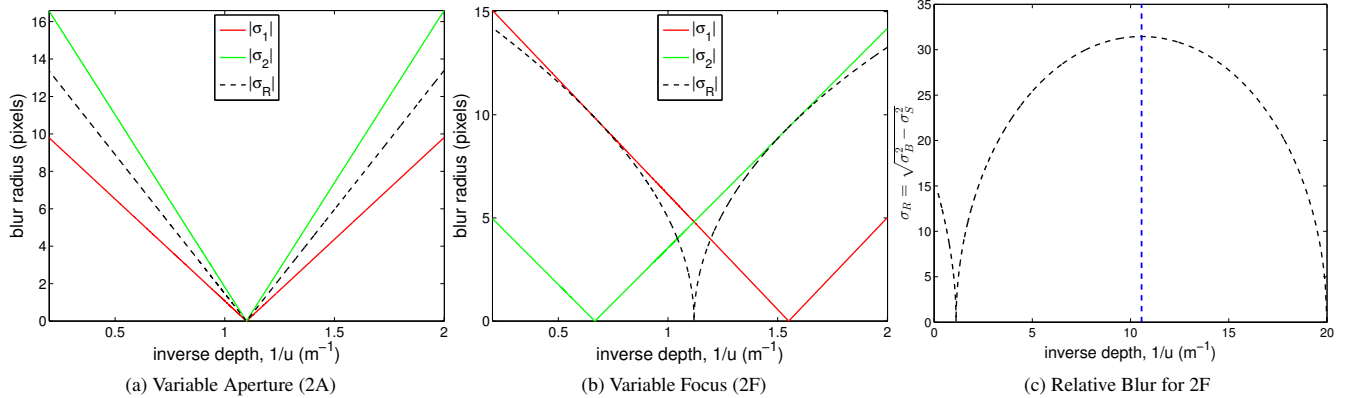


Figure 2: Absolute blur  $|\sigma_1|$  (red),  $|\sigma_2|$  (green) and relative blur  $\sigma_R$  (dashed) from Eq. (13) versus inverse depth  $u^{-1}$  in diopters (D) i.e.  $m^{-1}$ . a) Variable Aperture (2A) configuration, where the two images have the same focal length and focused at the same distance but different f-number, and b) Variable Focus (2F) configuration, where the two images have the same focal length and f-number but focused at two different depths. c) Relative blur for (b) with extended range (i.e. beyond  $2m^{-1}$ ). The blue line marks where ambiguity starts (Eq. 19).

However all of these methods include a relative blur based term in the cost function, and the additional terms incorporate some type of smoothness priors. Compared to these works our main contribution is categorizing different possible camera settings and understanding their operating range or limits within which depth can be estimated unambiguously. Compared to [18, 19] we do not only consider cases that give unique solution. Most notably we present the operating range for variable focus which is widely used in practice.

In the case of optimal camera parameters for DFD, Rajagopalan and Chaudhuri [11] estimated the optimal ratio of blur between two images, which can be used to find the optimal aperture pair. With a similar goal Zhou et al. [23] looked at the problem of optimal coded aperture pair. Subbarao et al. investigated the performance of their spatial domain transform method in [20]. Schechner and Kiryati [13, 14] analyzed how far apart the two focal planes should be to get reliable depth estimates. Their analysis assumes the sharper image to be noisy. However in the relative blur based approaches it is more appropriate to assume noise to be in the blurrier image as the sharper image is synthetically blurred to get the blurrier image. Our analysis considers noise to be in the blurrier image. Shih et al. [16, 17] takes an approach similar to Rajagopalan and Chaudhuri and derives the optimal relative blur between two images from the Cramér-Rao Lower Bound. For single image DFD, Trouvéloux et al. [21] showed a performance measure using the Cramér-Rao Lower Bound. We use Shih et al.’s work as the starting point of our analysis and extend it to finding opti-

mal camera parameters under different conditions. Instead of optimizing the variance lower-bound of relative blur, we emphasize on optimizing the variance lower-bound of blur, inverse depth, and depth. This gives more insight into how to choose the camera parameters and also where in the scene to expect best performance.

### 3. Camera Parameters and Operating Range

In this section, we first look at three general camera configurations, their special cases, and the conditions under which the relative blur  $\sigma_R$  uniquely determines the blur values  $\sigma_1$  and  $\sigma_2$ .

In Sec. 2 we saw how to model DFD using relative blur. Figs. 2a and 2b show that different depths can both have the same absolute and relative blur. In the case of 2A, for a given relative blur radius there are two possible  $(\sigma_1, \sigma_2)$  pairs on either side of the focal plane. For 2F, this ambiguity can be resolved based on which image is more blurred as long as the scenes are limited to one side of the critical depth shown in Fig. 2c. Here we address the more general question of what are the limits of unambiguous depth estimation in relative blur based methods.

We first consider the relative blur equation (Eq. 13) in terms of the linear relationship (Eq. 4) between two different blurs  $\sigma_1$  and  $\sigma_2$  and obtain:

$$\sigma_R = \sqrt{|(\alpha^2 - 1)\sigma_1^2 + 2\alpha\beta\sigma_1 + \beta^2|} \quad (15)$$

The term inside the absolute value is a parabola of variable  $\sigma_1$  and can be considered as the signed squared relative blur.

The parabola has a critical point at

$$\sigma_1^c \equiv \frac{-\alpha\beta}{\alpha^2 - 1}. \quad (16)$$

To unambiguously determine depth, the set of possible blurs needs to be restricted to one of the two sides of  $\sigma_1^c$  in Eq. 16. Next we look at what this means for three general camera configurations, and their special cases.

**Case 1:  $\alpha \neq 1$  and  $\beta = 0$**

In this case the critical point of the parabola is at  $\sigma_1^c = 0$ . The corresponding point in the scene will depend on the actual camera parameter that produces this configuration. In the following we review two such cases.

**Variable Aperture (2A)** The variable aperture case discussed in Sec. 2.2 (Eq. 6) has been around since the very beginning of DFD [10]. The condition  $\sigma_1^c = 0$  and the blur equation Eq. 2 tell us that we need to limit depth estimation to one side of the focal plane. This limitation is well known in the DFD literature and requires either limiting the range of depths or focusing the lens to infinity.

**Infinity focus with variable focal length** This is another special case of  $\beta = 0$  where  $f_1 = s_1$  and  $f_2 = s_2$ . This corresponds to having two different focal lengths and focusing at infinity. From Eq. 5,

$$\alpha = \frac{A_2 f_2}{A_1 f_1}. \quad (17)$$

Like variable aperture, the critical point is at  $\sigma_1 = 0$ , which in this case is at infinity (i.e. unique solution). Note that when  $f_1 = f_2$ , this turns into variable aperture.

**Case 2:  $\alpha = 1$  and  $\beta \neq 0$**

When  $\alpha = 1$ , the signed squared relative blur in Eq. 15 is linear, and therefore  $\sigma_1^c = \infty$  (i.e. unique solution). This condition can be achieved when  $N_1 s_2 f_2 = N_2 s_1 f_1$ . Subbarao et al. [19] used a particular instance of this configuration where the apertures were fixed (i.e.  $f_1/N_1 = f_2/N_2$ ), and the focal lengths ( $f$ ) along with the sensor to lens distances ( $s$ ) were varied. One can however use other configurations to obtain the same effect as long as  $\beta \neq 0$  (see supplementary material<sup>3</sup> for examples).

**Case 3:  $\alpha \neq 1$  and  $\beta \neq 0$**

This configuration is less constrained than the previous ones. In the following we consider a well known special case.

**Variable Focus (2F)** From Eq. 7 and Eq. 16 we have,

$$\sigma_1^c = \frac{-\rho f s_2}{2N(s_1 + s_2)}. \quad (18)$$

The depth ( $u^c$ ) corresponding to the critical point can be found by plugging Eq. 18 back into the blur equation (Eq. 2), i.e.

$$\frac{1}{u^c} = \frac{1}{f} - \frac{1}{s_1 + s_2}. \quad (19)$$

Eq. 19 suggests that the critical depth  $u^c$  depends on the camera parameters. For a given  $f$ ,  $u^c$  lies between  $f$  and some distance greater than  $f$  that is defined by the case when  $s_1 + s_2$  takes its smallest possible value. But we know from the thin lens model that  $s_1 + s_2 > 2f$ . It follows that  $u^c$  can be at most  $2f$ . In other words, as long as we are estimating depth beyond  $2f$  there can be no ambiguity in depth. Fig. 2c shows an example of this critical point.

## 4. Relative Blur Error Analysis

In this section we consider how well relative blur and depth (or inverse depth) can be estimated from two defocused images. More specifically, which camera parameters yield better estimates? Recall that in the blurred image formation model we assumed additive noise with a known probability distribution (Gaussian in our case). Therefore the problem of estimating error in relative blur and depth can be treated in the framework of statistical estimation. Given a DFD problem how do we evaluate the performance of the estimator? In this work we consider a minimum variance unbiased estimator [7]. For an unbiased estimator with known probability distribution, the theoretical lower bound of the estimator variance is given by the Cramér-Rao Lower Bound (CRLB). This approach has been previously used in [11, 16, 17]. We build on these works and evaluate the performance under different camera configurations.

Relative blur estimation requires minimizing the objective function specified in Eq. 14. For a fronto-parallel scene this is essentially a regression problem where we find the radius of the blur that minimizes the error. Shih et al., in [16, 17] analytically derived the lower bound of the variance in relative blur estimates assuming a Gaussian noise distribution in the blurred image and negligible noise in the sharp image. This is a reasonable assumption to make since blurring the sharp image will result in the noise variance going down very rapidly as relative blur increases. The lower bound of the variance of  $\hat{\sigma}_R$  (the  $\hat{\cdot}$  operator denotes estimate) that they derived for noise variance  $\sigma_n^2$  is:

$$\text{var}\{\hat{\sigma}_R\} \geq \frac{(\sigma_S^2 + \sigma_R^2)^3}{K\sigma_R^2} = \frac{\sigma_B^6}{K\sigma_R^2} \quad (20)$$

where  $K = f(\sigma_n, I_0)$  depends on the original pinhole image  $I_0$  (i.e. the true sharp image) and noise. Therefore,

<sup>3</sup><http://cim.mcgill.ca/~fmannan/3dv/DFD.html>

given a scene and noise variance,  $K$  will be a constant scaling factor for the estimator variance. Also notice that the numerator is the defocus radius of the blurrier image. Therefore it is desirable to minimize the blur of the blurrier image to obtain better estimates.

We can derive a number of results from the CRLB. First, if we want to minimize the variance of  $\hat{\sigma}_R$  with respect to relative blur  $\sigma_R$ , the optimal relative blur is (as found by Shih et al.):

$$\sigma_R = \frac{\sigma_S}{\sqrt{2}} \quad (21)$$

However in practice, we are mainly interested in minimizing the variance of depth or inverse depth estimates. For an intuitive motivation of why variance of inverse depth is more helpful than that of relative blur, consider how the relative blur to depth conversion differs in 2F and 2A. For the 2F case in Fig. 2b, the relative blur changes rapidly at the intersection of the two absolute blur curves. Therefore large variation in relative blur estimates in that region will have a small effect in inverse depth estimates. However for the 2A case (Fig. 2a), variation in relative blur estimate results in a constant (actual value will depend on the camera settings) variation in inverse depth estimate everywhere. Therefore we should expect the variance of inverse depth (or blur and depth) to behave differently. In addition to the lower bound of variance<sup>4</sup>, we want to find the optimal parameters (in this work the combination of parameters given by  $\alpha$  and  $\beta$ ), or the position (in terms of depth or inverse depth) in the scene where error variance is lowest. In the following we derive these relationships from first principles (see supplementary material<sup>3</sup> for details) and experimentally verify them. To our knowledge this has not been previously addressed.

Using the CRLB theorem for general functions of the estimate, we can derive the lower bound on the variance of blur in one of the images as follows (see supplementary material<sup>3</sup>):

$$\text{var}\{\hat{\sigma}_1\} \geq \left(\frac{d\sigma_1}{d\sigma_R}\right)^2 \text{var}\{\hat{\sigma}_R\} = \frac{\sigma_B^6}{K((\alpha^2 - 1)\sigma_1 + \alpha\beta)^2}. \quad (22)$$

For inverse depth estimate  $1/\hat{u}$ , we get,

$$\text{var}\left\{\frac{1}{\hat{u}}\right\} \geq \left(\frac{d\frac{1}{u}}{d\sigma_1}\right)^2 \text{var}\{\hat{\sigma}_1\} = \left(\frac{2N_1}{\rho f_1 s_1}\right)^2 \text{var}\{\hat{\sigma}_1\} \quad (23)$$

and for depth estimate  $\hat{u}$ :

$$\text{var}\{\hat{u}\} \geq \left(\frac{d\frac{1}{u}}{d1/u}\right)^2 \text{var}\left\{\frac{1}{\hat{u}}\right\} = u^4 \text{var}\left\{\frac{1}{\hat{u}}\right\}. \quad (24)$$

<sup>4</sup>In the rest of the paper we will use the term variance to refer to variance of blur, inverse depth and depth.

There are a few things to note from the above equations. Relative blur  $\sigma_R$  and blur  $\sigma_1$  have non-linear relationship and as a result the two variances have a non-linear relationship. Blur and depth have a linear relationship and therefore variance of inverse depth is a scalar multiple of variance of blur. Finally inverse depth and depth are non-linearly related which is also the case for their variance. In the rest of the paper, we consider optimizing  $\alpha$  and  $\beta$  for blur or inverse depth, for the different cases discussed in the previous section. Also, for a given set of parameters, we find the position (or inverse depth) where the lower bound of variance is lowest.

### Case 1: $\alpha \neq 1$ and $\beta = 0$

For this configuration we have two cases,  $\alpha < 1$  and  $\alpha > 1$ . In the first case,  $\sigma_1 > \sigma_2$  and so  $\sigma_B = \sigma_1$ . And in the second case,  $\sigma_1 < \sigma_2$  and  $\sigma_B = \sigma_2 = \alpha\sigma_1$ . For these two cases:

$$\text{var}\{\hat{\sigma}_1\} \geq \begin{cases} \frac{\sigma_1^4}{K(\alpha^2 - 1)^2} & , \alpha < 1 \\ \frac{\sigma_1^4 \alpha^6}{K(\alpha^2 - 1)^2} & , \alpha > 1 \end{cases} \quad (25)$$

The optimal  $\alpha$  for the two cases are as follows:

$$\alpha = \begin{cases} 0 & , \alpha < 1 \\ \sqrt{3} & , \alpha > 1 \end{cases} \quad (26)$$

In the above equation, choosing  $\alpha < 1$  may result in a smaller variance compared to the  $\alpha > 1$  case. However one needs to consider how large  $\sigma_1$  can get when choosing between the two cases. In the following we look at two special cases.

**Variable Aperture (2A)** In this special case, we have  $\beta = 0$ ,  $\alpha = N_1/N_2$ . Without loss of generality we can choose  $N_1 > N_2$ . This restricts us to the second case in Eq. 25. Therefore the optimal aperture ratio is  $\alpha = \sqrt{3}$ . The position in the scene where lower bound of the variance is minimized is at  $\sigma_B = 0$  which is at the focus distance.

It is interesting to compare this result with Eq. 21. If we are considering variable aperture then according to Eq. 21, the optimal ratio of the two apertures should be  $\sqrt{3}/2$ . However that is not the optimal if we are minimizing the variance of estimated blur. This ratio was also found experimentally by Chaudhuri and Rajagopalan [11, 1]. Note that in their paper they consider the ratio of blurs, which is equivalent to aperture ratio for variable aperture case. Furthermore from their work it is unclear as to how to achieve that ratio in general.

**Infinity Focus with Variable Focal Length** Depending on the focal length and aperture both cases of Eq. 26 are

possible. Here one can find the optimal combination of all four parameters in Eq. 17. However one simple choice is to set  $N_1 = N_2$  and  $f_1 > f_2$ . In this case, we are restricted to  $\alpha < 1$  and variance of the estimate improves as  $\alpha \rightarrow 0$ . Note that as  $\alpha \rightarrow 0$ , Eq. 25 depends on  $\sigma_1$ . Therefore when choosing camera parameters for  $\alpha < 1$  one needs to ensure that the resulting  $\sigma_1$  does not get too large.

### Case 2: $\alpha = 1$ and $\beta \neq 0$

In this case, the lower bound of the variance depends on the specific optical parameters that are chosen. The position where the variance lower bound is minimized is given by Eq. 27 which will be derived below in case 3.

### Case 3: $\alpha \neq 1$ and $\beta \neq 0$

**Variable Focus (2F)** Using the relationship from Eq. 7 and minimizing Eq. 22 wrt  $\alpha$ , it can be shown that optimal  $\alpha$  is a function of  $\sigma_1$  (see supplementary material). In other words there is no single  $\alpha$  (i.e. ratio of sensor distances) that minimizes the variance for all depths. For a given  $\alpha$ , the scene depth with the lowest variance can be found at the intersection of the two blur curves i.e. when  $\sigma_R = 0$ . Equating the two blurs give (see supplementary material),

$$\sigma_1 = -\beta/(\alpha + 1). \quad (27)$$

Therefore, given an  $\alpha$ , the depth halfway between the two focal planes (in diopters) has the lowest variance. Note that the lower bound of variance at this point is not zero. If we want to find an  $\alpha$  that minimizes the lower bound at this point, we can do that by substituting  $\sigma_1$  from Eq. 27 into the variance equations to get:

$$\text{var}\{\hat{\sigma}_1\} \geq \left(\frac{\rho f}{2N}\right)^2 \frac{(\alpha - 1)^4}{(\alpha + 1)^6}. \quad (28)$$

As  $\alpha \rightarrow 1$ , the lower bound of the variance decreases. This implies that as the two focal planes move closer to each other, we can expect the lower bound of error variance to get smaller at the intersection of the two blurs. Also as  $\alpha \rightarrow 1$ ,  $\sigma_R \rightarrow 0$ . As a result, the negligible noise in the sharp image assumption will not hold and the model will not be a good approximation. In this case, we can expect the overall variance to increase. These observations are experimentally verified later in the paper (see Fig. 3).

## 5. Experimental Results

We now experimentally verify some of the results derived in Sec. 4 using synthetic and real defocused images. In both cases we use a fractal texture (1/frequency amplitude spectrum). For synthetic experiments, this texture is synthetically blurred using Gaussian PSFs. For a given depth and a pair of camera parameters, the blur radii pair

$(\sigma_1, \sigma_2)$  are determined using Eq. 2. Then for each pair the texture is synthetically blurred and Gaussian noise is added. Finally Eq. 14 is used to estimate the relative blur which is then converted to inverse depth estimates.

Fig. 3a plots the Root Mean Squared Error (RMSE) for depth range 50 cm to 5 m, for 6 different camera configurations (3 variable focus, 2 variable aperture and 1 infinite focus with variable focal length). For unbiased estimation, the RMSE is equal to the standard deviation. Fig. 3b shows the corresponding plots obtained using theory (with  $K = 1$ ).

For experiments with real images we placed a camera (Nikon D90 with 50 mm, f/1.8 lens) in front of an LED Cinema Display (94.3 ppi). The object to sensor distance (or focus mark) was between 0.61 to 1.5 m at approximately 5 cm intervals. The pre-processing pipeline involve vignetting correction, and alignment for multifocus images. We took images of the same fractal noise texture as in the synthetic experiment. Variable aperture images were taken with f/22 and f/11 with two different focus setting. Variable focus images were taken with f/16 with focus 0.61 and 1.5 m. The raw images were linearized using a conversion table provided by the camera, followed by exposure normalization and vignetting correction. For variable focus experiments we also aligned the images. For vignetting correction we displayed a uniform color on the monitor and took images for every camera settings. We then fit a higher-order polynomial model with center estimation. For alignment we took images of a grid of disks, estimated the center of the disks and fit an affine model that aligns the centers for differently focused disk images. For relative blur estimation, we first ensured that both images had the same mean value by dividing by their respective mean. Then we took  $201 \times 201$  sized patches and solved Eq. 14 for  $\sigma_R$  using continuous least squares. For variable focus, we specified the sign of  $\sigma_R$  and added the constraint  $|\sigma_R| \geq 0$ . This reduces the errors in the derivative near  $\sigma_R = 0$  due to discretization and image noise. Results for the real image based experiments are shown in Fig. 3c.

Fig. 3 shows that both synthetic and real experiments produce variance curves that are consistent with the theoretical lower bound of inverse depth variance. For variable aperture, best performance is at the focus region. For variable focus, best performance is near  $\sigma_R = 0$ . The performance improves at the blur intersection as the focal planes get closer to each other up to a certain limit (e.g. 0.88D - 1.33D vs. 0.2D - 2D).

It should be noted that in the synthetic experiments we have noise in both images unlike the simplifying assumption made in theory. The effect of this noisy sharp image is seen when the relative blur in the two images are small (i.e.  $\sigma_R$  near 0). In this case, synthetically blurring the sharp image by a small amount does not reduce the noise, and as a result, the theory plot differs from experiment. Therefore

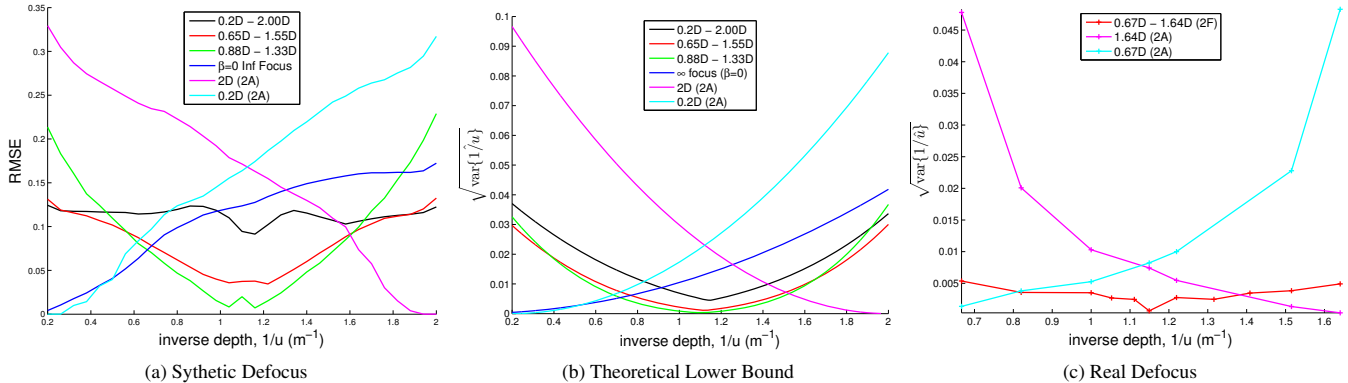


Figure 3: a) The RMS errors with respect to inverse depth for synthetic experiments. b) The lower-bound theoretical variance for (a). Note that the y-axis has arbitrary scale (because of the constant  $K$  in Eq. 20). For (a) and (b), the red, green and black curves correspond to variable focus or 2F (e.g. 0.2 D - 2.00 D corresponds to focus distances at 5 m and 0.5 m), magenta and cyan correspond to variable aperture or 2A, and blue is variable focal length with focus at infinity. c) Estimated variance from real images using variable aperture (2A) and focus (2F).

Table 1: Summary of different camera configurations, range of unambiguous depth and accuracy.

Configuration	$\sigma_1^c$	$u^c$	Opt. Params	Opt. Dist.
Case 1: $\alpha \neq 1$ and $\beta = 0$ - Variable Aperture (2A) - Inf. Focus with Variable Focal Length	0	focus dist.	$\alpha = \sqrt{3}$ $\alpha \rightarrow 0$ and $\sigma_1$ small, or $\alpha = \sqrt{3}$	focus dist.
Case 2: $\alpha = 1$ and $\beta \neq 0$	$\infty$	$\infty$	Eqs. 22, 23, 24	Intersection of blur curves (Eq. 27)
Case 3: $\alpha \neq 1$ and $\beta \neq 0$ - Variable Focus (2F)	Eq. 16 Eq. 18	Eq. 2 with $\sigma_1^c$ > $2f$	Eqs. 22, 23, 24	Intersection of blur curves (Eq. 27)

we should expect larger error in relative blur based methods as  $|\sigma_R|$  gets smaller.

## 6. Conclusion

In this paper, we have shown under what conditions depth can be unambiguously estimated from relative blur. We have also analyzed the error variance for different configurations and the optimal parameters  $(\alpha, \beta)$  that reduce the variance. These results can be used to compute the lower bound of estimator variance for any given camera parameters. Our findings are summarized in Table 1. The first column shows the different configurations and some of the special cases addressed in this paper. The second and third columns give the critical points for unambiguous depth in terms of blur and depth, and the last two columns give the optimal parameters and the location where the lowest variance occurs.

The main application of this work is in finding optimal

camera parameters that reduce the error variance over a range of depths or inverse depths. Furthermore we can find combination of different parameters for specific scenes and thereby allow us to improve reconstruction without taking a large number of images like Depth from Focus.

## Acknowledgements

We would like to thank Stéphane Kaufmann for discussions on an early version of this work. This work was supported by grants from the Natural Sciences and Engineering Research Council of Canada (NSERC).

## References

- [1] S. Chaudhuri and A. N. Rajagopalan. *Depth from defocus - a real aperture imaging approach*. Springer, 1999. 6
- [2] J. Ens and P. Lawrence. An investigation of methods for determining depth from focus. *PAMI*, 15(2):97–108, 1993. 3



- [3] P. Favaro and S. Soatto. A geometric approach to shape from defocus. *PAMI*, 27(3):406–417, march 2005. 3
- [4] P. Favaro and S. Soatto. *3-D Shape Estimation and Image Restoration - Exploiting Defocus and Motion Blur*. Springer, 2007. 3
- [5] P. Grossmann. Depth from focus. *Pattern Recognition Letters*, 5(1):63–69, 1987. 1
- [6] S. W. Hasinoff and K. N. Kutulakos. Confocal stereo. *IJCV*, 81(1):82–104, 2009. 1
- [7] S. Kay. *Fundamentals of Statistical Signal Processing: Estimation Theory*. Number v. 1 in *Fundamentals of Statistical Signal Processing*. Prentice-Hall PTR, 1998. 5
- [8] F. Li, J. Sun, J. Wang, and J. Yu. Dual-focus stereo imaging. *Journal of Electronic Imaging*, 19:043009, 2010. 3
- [9] V. Namboodiri, S. Chaudhuri, and S. Hadap. Regularized depth from defocus. In *ICIP*, pages 1520–1523, oct. 2008. 3
- [10] A. P. Pentland. A new sense for depth of field. *PAMI*, 9:523–531, July 1987. 1, 3, 5
- [11] A. Rajagopalan and S. Chaudhuri. Optimal selection of camera parameters for recovery of depth from defocused images. In *CVPR*, pages 219–224, jun 1997. 4, 5, 6
- [12] A. Rajagopalan and S. Chaudhuri. An MRF model-based approach to simultaneous recovery of depth and restoration from defocused images. *PAMI*, 21(7):577–589, jul 1999. 3
- [13] Y. Y. Schechner and N. Kiryati. The optimal axial interval in estimating depth from defocus. In *ICCV*, pages 834–838, 1999. 4
- [14] Y. Y. Schechner and N. Kiryati. Depth from defocus vs. stereo: How different really are they? *IJCV*, 39:141–162, September 2000. 1, 4
- [15] S. M. Seitz and S. Baker. Filter flow. In *ICCV*, pages 143–150, 29 2009–oct. 2 2009. 3
- [16] S.-W. Shih, P.-S. Kao, and W.-S. Guo. Error analysis and accuracy improvement of depth from defocusing. In *CVGIP*, 2003. 4, 5
- [17] S.-W. Shih, P.-S. Kao, and W.-S. Guo. An error bound of relative image blur analysis. In *ICPR*, volume 4, pages 100–103 Vol.4, 2004. 4, 5
- [18] M. Subbarao. Parallel depth recovery by changing camera parameters. In *ICCV*, pages 149–155, dec 1988. 2, 3, 4
- [19] M. Subbarao and G. Surya. Depth from defocus: A spatial domain approach. *IJCV*, 13(3):271–294, 1994. 3, 4, 5
- [20] M. Subbarao and J.-K. Tyan. Noise sensitivity analysis of depth-from-defocus by a spatial-domain approach. In *Proc. SPIE 3174*, pages 174–187, 1997. 4
- [21] P. Trouvé-Peloux, F. Champagnat, G. L. Besnerais, and J. Idier. Theoretical performance model for single image depth from defocus. *J. Opt. Soc. Am. A*, 31(12):2650–2662, Dec 2014. 4
- [22] M. Watanabe and S. K. Nayar. Rational filters for passive depth from defocus. *IJCV*, 27:203–225, May 1998. 3
- [23] C. Zhou, S. Lin, and S. Nayar. Coded Aperture Pairs for Depth from Defocus and Defocus Deblurring. *International Journal on Computer Vision*, 93(1):53, May 2011. 4