Presented at: **ECCV**, Cambridge, England, 1996.

Direct Differential Range Estimation Using Optical Masks

Eero P. Simoncelli and Hany Farid

GRASP Laboratory
University of Pennsylvania, Philadelphia, PA 19104, USA

Abstract. We describe a novel formulation of the range recovery problem, based on computation of the differential variation in image intensities with respect to changes in camera position. The method uses a single stationary camera and a pair of calibrated optical masks to directly measure this differential quantity. The subsequent computation of the range image is simple and should be suitable for real-time implementation. We also describe a variant of this technique, based on direct measurement of the differential change in image intensities with respect to aperture size. These methods are comparable in accuracy to other single-lens ranging techniques. We demonstrate the potential of our approach with a simple example.

1 Introduction

Visual images are formed via the projection of light from the three-dimensional world onto a two-dimensional sensor. In an idealized pinhole camera, all points lying on a ray passing through the pinhole will be imaged onto the same image position. Thus, information about the distance to objects in the scene (i.e., *range*) is lost. Range information can be recovered by measuring the change in appearance of the world resulting from a change in viewing position. Traditionally, this is accomplished via simultaneous measurements with two cameras (binocular stereo), or via a sequence of measurements collected over time from a moving camera (structure from motion).

The recovery of range in these approaches frequently relies on an assumption of brightness constancy, which states that the brightness of the image of a point in the world is constant when viewed from different positions [Horn 86]. Consider the formulation of this assumption in one dimension (the extension to two dimensions is straightforward). Let $f(x;v)$ describe the intensity function measured through a pinhole camera system. The variable v corresponds to the pinhole position (along the direction perpendicular to the optical axis). The variable x parameterizes the position on the sensor. This configuration is illustrated in Figure 1. According to the assumption, the intensity function $f(x;v)$ is of the form:

$$f(x;v) = I\left(x - \frac{v\,d}{Z}\right), \tag{1}$$

where $I(x) = f(x;v)\,|_{v=0}$, d is the distance between the pinhole and the sensor and Z is the range (distance from the pinhole to a point in the world). Note

This work was supported by ARO Grant DAAL03-89-C-0031PRI, NSF Grant STC SBR-89-20230, and ARO/MURI Grant DAAH04-96-1-0007.

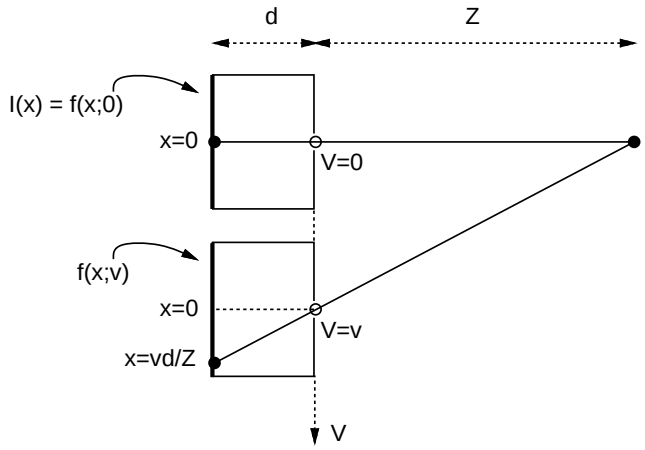

Fig. 1. Geometry for a binocular stereo system with pinhole cameras. The variable V parameterizes the position of the camera pinholes. According to the brightness constancy constraint, the intensity of a point in the world, as recorded by the two pinhole cameras, should be the same.

that this assumption will typically be violated near occlusion boundaries, where points visible from one viewpoint are invisible from another.

Several complications arise in these approaches. The degree to which the brightness constancy assumption holds will, in general, decrease with increasing camera displacement. This is due to larger occluded image regions, and increased effects of the non-Lambertianity of surface reflectances. Violations of the brightness constancy assumption lead to difficulties in matching corresponding points in the images (the so-called "correspondence problem"). Furthermore, a two-camera stereo system (or a single moving camera) requires careful calibration of relative positions, orientations, and intrinsic parameters of the camera(s).

These problems are partially alleviated in techniques utilizing a single stationary camera. A number of these techniques are based on estimation of blur or relative blur from two or more images (e.g., [Krotkov 87, Pentland 87, Subbarao 88, Xiong 93, Nayar 95]). Adelson [Adelson 92] describes an unusual method in which a lenticular array is placed over the sensor, effectively allowing the camera to capture visual images from several viewpoints in a single exposure. Dowski [Dowski 94] and Jones [Jones 93] each describe range imaging systems that use an optical attenuation mask in front of the lens. By observing local spectral information in a single image, they are able to estimate range. Both techniques rely on power spectral assumptions about the scene.

In this paper, we propose a single-camera method which avoids some of the computational and technical difficulties of the single-camera approaches discussed above. In particular, we propose a "direct" differential method for range estimation which computes the image derivative with respect to viewing position using a single stationary camera and an optical attenuation mask. We

also present a variation based on the derivative with respect to aperture size (i.e., differential range-from-defocus). These approaches avoid the correspondence problem, make no spectral assumptions about the scene, are relatively straightforward to calibrate, and are computationally efficient.

2 Direct Viewpoint Derivatives

For the purpose of recovering range, we are interested in computing the change in the appearance of the world with respect to change in viewing position. It is thus natural to consider differential measurement techniques. Taking partial derivatives of $f(x;v)$ with respect to the image and viewing positions, and evaluating at $v = 0$ gives:

$$\begin{aligned} I_x(x) &\equiv \tfrac{\partial f(x;v)}{\partial x} \mid_{v=0} \\ &= I'(x), \end{aligned} \tag{2}$$

and

$$\begin{aligned} I_v(x) &\equiv \tfrac{\partial f(x;v)}{\partial v} \mid_{v=0} \\ &= -\tfrac{d}{Z} I'(x), \end{aligned} \tag{3}$$

where $I'(\cdot)$ indicates the derivative of $I(\cdot)$ with respect to its argument. Combining these two expressions gives:

$$I_v(x) = -\tfrac{d}{Z} I_x(x). \tag{4}$$

Clearly, an estimate of the range, Z, can be computed using this equation. Note that in the case of differential binocular stereo (e.g., [Lucas 81]), the derivative with respect to viewing position, I_v, is replaced by a difference, $I_{v_1} - I_{v_2}$. A similar relationship is used in computing structure from motion (for known camera motion), where I_v is typically replaced by differences of consecutive images.

We now show a direct method for measurement of this derivative through the use of an optical attenuation mask. Consider a world consisting of a single point light source and a standard lens-based imaging system with a variable-opacity optical mask, $M(u)$, placed directly in front of the lens (left side of Figure 2). The light striking the lens is attenuated by the value of the optical mask function at that particular spatial location.[1] With such a configuration, the image of the point source will be a scaled and dilated version of the optical mask function:

$$I(x) = \tfrac{1}{\alpha} M(\tfrac{x}{\alpha}), \tag{5}$$

as illustrated in Figure 2. The scale factor, α, is a monotonic function of the distance to the point source, Z, and may be derived from the imaging geometry:

$$\alpha = 1 - \tfrac{d}{f} + \tfrac{d}{Z}, \tag{6}$$

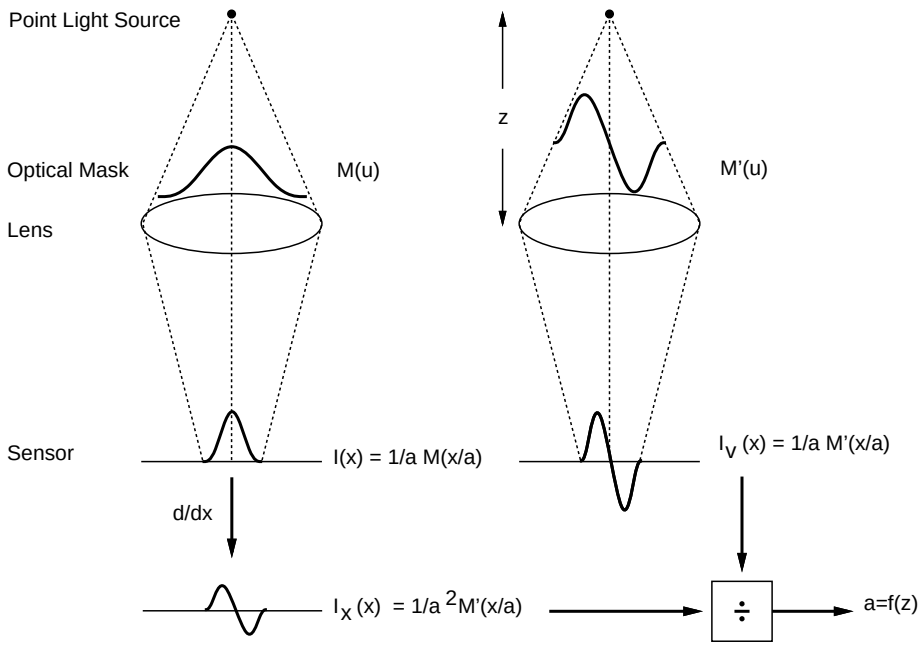

Fig. 2. Direct differential range determination for a single point source. Images of a point light source are formed using two different optical masks, corresponding to the function $M(u)$ and its derivative, $M'(u)$. In each case, the image formed is a scaled and dilated copy of the mask function (by an amount α). Computing the spatial (image) derivative of the image formed under mask $M(u)$ produces an image that is identical to the image formed under the derivative mask, $M'(u)$, except for a scale factor α. Thus, α may be estimated as the ratio of the two images. Range is computed from α using the relationship given in Equation (6).

where d is the distance between lens and sensor, and f is the focal length of the lens.

In the system shown on the left side of Figure 2, the *effective* viewpoint may be altered by translating the mask, while leaving the lens and sensor stationary.[2] The generalized intensity function, for a mask centered at position v is written as:

$$f(x;v) = \frac{1}{\alpha}\, M(\frac{x}{\alpha} - v), \tag{7}$$

assuming that the non-zero portion of the mask does not extend past the edge of the lens.

[1] For our purposes, we assume that the values of such a mask function are real numbers in the range [0,1].

[2] For example, consider a mask which contains a single pinhole; different views of the world are obtained by sliding the pinhole across the front of the lens.

The *differential* change in the image (with respect to a change in the mask position) may be computed by taking the derivative of this equation with respect to the mask position, v, evaluated at $v = 0$:

$$\begin{aligned} I_v(x) &\equiv \tfrac{\partial}{\partial v}\, f(x;v)|_{v=0} \\ &= -\tfrac{1}{\alpha} M'(\tfrac{x}{\alpha}), \end{aligned} \tag{8}$$

where $M'(\cdot)$ is the derivative of the mask function $M(\cdot)$ with respect to its argument. The derivative with respect to viewing position, $I_v(x)$, *may thus be computed directly by imaging with the optical mask* $M'(\cdot)$.[3]

Finally, notice that the spatial derivative of $I(x)$ is closely related to the image $I_v(x)$:

$$\begin{aligned} I_x(x) &\equiv \tfrac{\partial}{\partial x}\, f(x;v)|_{v=0} \\ &= \tfrac{1}{\alpha^2} M'(\tfrac{x}{\alpha}) \\ &= -\tfrac{1}{\alpha}\, I_v(x). \end{aligned} \tag{9}$$

From this relationship, the scaling parameter α may be computed as the ratio of the spatial derivative of the image formed through optical mask $M(u)$, and the image formed through the derivative of that optical mask, $M'(u)$. This computation is illustrated in Figure 2. The distance to the point source can subsequently be computed from α using the monotonic relationship given in Equation (6). Note that the resulting equation for distance is identical to that of Equation (4) when $d = f$ (i.e., when the camera is focused at infinity).

The necessity of the brightness constancy assumption can now be made explicit. For our system, brightness constancy means that the light emanating from a point is of constant intensity across the surface of the mask. A point light source that violates this assumption has a directionally varying light emission, $L(\cdot)$, and when imaged through the pair of optical masks will produce images of the form:

$$I(x) = \tfrac{1}{\alpha} L(\tfrac{x}{\alpha}) \cdot M(\tfrac{x}{\alpha}) \tag{10}$$

$$I_v(x) = \tfrac{1}{\alpha} L(\tfrac{x}{\alpha}) \cdot M'(\tfrac{x}{\alpha}). \tag{11}$$

As before, computing the derivative of $I(x)$ yields:

$$I_x(x) = \tfrac{1}{\alpha^2} \left(L(\tfrac{x}{\alpha}) M'(\tfrac{x}{\alpha}) + L'(\tfrac{x}{\alpha}) M(\tfrac{x}{\alpha}) \right). \tag{12}$$

Thus, if the light is not constant across the aperture (i.e., $L'(\cdot) \neq 0$) then the simple relationship between $I_v(x)$ and $I_x(x)$ (given in Equation (9)) will not hold.

[3] In practice, $M'(u)$ cannot be directly used as an attenuation mask, since it contains negative values. This issue is addressed in Section 4.

3 Range Estimation

Equation (9) embodies the fundamental relationship used for the direct differential computation of range of a single point light source. A more realistic world consisting of a collection of many such uniform intensity point sources imaged through an optical mask will produce an image consisting of a superposition of scaled and dilated versions of the masks. In particular, we can write an expression for the image by summing the images of the visible points, p, in the world:

$$f(x;v) = \int dx_p \; \frac{1}{\alpha_p} M\left(\frac{x-x_p}{\alpha_p} - v\right) L(x_p), \tag{13}$$

where the integral is performed over the variable x_p, the position in the sensor of a point p projected through the center of the lens. The intensity of the world point p is denoted as $L(x_p)$, and α_p is monotonically related to the distance to p (as in Equation (6)). Note again that we must assume that each point produces a uniform light intensity across the optical mask.

Again, consider the derivatives of $f(x;v)$ with respect to viewing position, v, and image position, x:

$$\begin{aligned} \frac{\partial}{\partial v} f(x;v) &= \frac{\partial}{\partial v} \int dx_p \; \frac{1}{\alpha_p} M\left(\frac{x-x_p}{\alpha_p} - v\right) L(x_p) \\ &= -\int dx_p \; \frac{1}{\alpha_p} M'\left(\frac{x-x_p}{\alpha_p} - v\right) L(x_p), \end{aligned} \tag{14}$$

and

$$\begin{aligned} \frac{\partial}{\partial x} f(x;v) &= \frac{\partial}{\partial x} \int dx_p \; \frac{1}{\alpha_p} M\left(\frac{x-x_p}{\alpha_p} - v\right) L(x_p) \\ &= \int dx_p \; \frac{1}{\alpha_p^2} M'\left(\frac{x-x_p}{\alpha_p} - v\right) L(x_p), \end{aligned} \tag{15}$$

where (as before) $M'(\cdot)$ is the derivative of $M(\cdot)$ with respect to its argument.

As in the previous section, the following two partial derivative images are defined:

$$\begin{aligned} I_v(x) &\equiv \frac{\partial}{\partial v} f(x;v) \mid_{v=0} \\ &= -\int dx_p \; \frac{1}{\alpha_p} M'\left(\frac{x-x_p}{\alpha_p}\right) L(x_p), \end{aligned} \tag{16}$$

and

$$\begin{aligned} I_x(x) &\equiv \frac{\partial}{\partial x} f(x;v) \mid_{v=0} \\ &= \int dx_p \; \frac{1}{\alpha_p^2} M'\left(\frac{x-x_p}{\alpha_p}\right) L(x_p). \end{aligned} \tag{17}$$

Equations (16) and (17) differ only in a multiplicative term of $\frac{1}{\alpha_p}$. Unfortunately, solving for α_p is nontrivial, since it is embedded in the integrand and depends on the integration variable. Consider, however, the special case where all points in the world lie on a frontal-parallel plane relative to the sensor.[4] Un-

[4] In actuality, this assumption need only be made locally.

der this condition, the scaling parameter α_p is the same for all points x_p and Equations (16) and (17) can be written as:

$$I_v(x) = \frac{1}{\alpha} \int dx_p \; M'\left(\frac{x - x_p}{\alpha}\right) L(x_p) \tag{18}$$

$$I_x(x) = \frac{1}{\alpha^2} \int dx_p \; M'\left(\frac{x - x_p}{\alpha}\right) L(x_p). \tag{19}$$

The scaling parameter, α, a function of the distance to the points in the world (Equation (6)) can then be computed as the ratio:

$$I_v(x) = -\alpha I_x(x). \tag{20}$$

As in the single-point case, this expression is identical to that of Equation (4) with $d = f$.

In order to deal with singularities (i.e., $I_x = 0$), a least-squares estimator can be used for α (as in [Lucas 81]). Specifically, we minimize $E(\alpha) = \sum_P (I_v + \alpha I_x)^2$, where the summation is performed over a small patch in the image, P. Taking the derivative with respect to α, setting equal to zero and solving for α yields the minimal solution:

$$\alpha = -\frac{\sum_P I_v I_x}{\sum_P I_x^2}. \tag{21}$$

The algorithm easily extends to a three-dimensional world: we need only consider two-dimensional masks $M(u, w)$, and the horizontal partial derivative $M'(u, w) = \partial M(u, w)/\partial u$. For a more robust implementation, the vertical partial derivative mask $\partial M(u, w)/\partial w$ may also be included. The least-squares error function becomes:

$$E(\alpha) = \sum_P (I_u + \alpha I_x)^2 + (I_w + \alpha I_y)^2. \tag{22}$$

Solving for the minimizing α gives:

$$\alpha = -\frac{\sum_P (I_u I_x + I_w I_y)}{\sum_P (I_x^2 + I_y^2)}. \tag{23}$$

4 Aperture Mask Design

Thus far, the only restriction placed on the aperture masks, $M(u, w)$ and $M'(u, w)$, is that the second be the derivative of the first. Figure 3 contains a matched pair of masks based on a two-dimensional Gaussian. In practice, the function $M'(u, w)$ has negative values and thus is not feasible for use as an optical attenuation mask. Furthermore, a positive constant cannot simply be added to $M'(u, w)$, since this will destroy the required derivative relationship between the two masks.

Due to the linearity of the imaging process, however, we can use masks that are linear combinations of the masks $M'(u, w)$ and $M(u, w)$. In particular, a

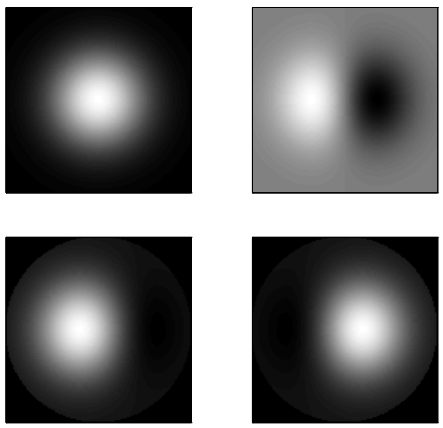

Fig. 3. Gaussian aperture masks. Top left: A two-dimensional Gaussian mask, $M(u,w)$. Top right: Gaussian partial derivative, $M'(u,w)$. Bottom row: Two non-negative aperture masks, $M_1(u,w)$ and $M_2(u,w)$. These are computed from the top masks using Equations (24) and (25).

scalar multiple of $M(u,w)$ can be added to $M'(u,w)$ in order to form a mask function that is entirely positive. The new mask, $M_1(u,w)$, shown in Figure 3, is given by:

$$M_1(u,w) = \beta M(u,w) + \gamma M'(u,w), \tag{24}$$

where β, γ are scaling constants chosen to force the function $M_1(u,w)$ to fill the range $[0,1]$. A second symmetrical mask can be formed by *subtracting* $M'(u,w)$ from $M(u,w)$:

$$M_2(u,w) = \beta M(u,w) - \gamma M'(u,w). \tag{25}$$

Note that $M_2(u,w)$ is equal to $M_1(u,w)$ rotated 180 degrees about its center, and that

$$M(u,w) = \frac{M_1(u,w)+M_2(u,w)}{2\beta} \tag{26}$$

$$M'(u,w) = \frac{M_1(u,w)-M_2(u,w)}{2\gamma}. \tag{27}$$

Again, by linearity of the imaging process, the images that would have been obtained with the masks $M(u,w)$ and $M'(u,w)$ can be recovered from images obtained with two masks $M_1(u,w)$ and $M_2(u,w)$. In particular, let $I_1(u,w)$ be the image obtained through the mask M_1, and $I_2(u,w)$ the image obtained through the mask M_2. Then:

$$I(u,w) = \frac{I_1(u,w)+I_2(u,w)}{2\beta} \tag{28}$$

$$I_v(u,w) = \frac{I_1(u,w)-I_2(u,w)}{2\gamma}, \tag{29}$$

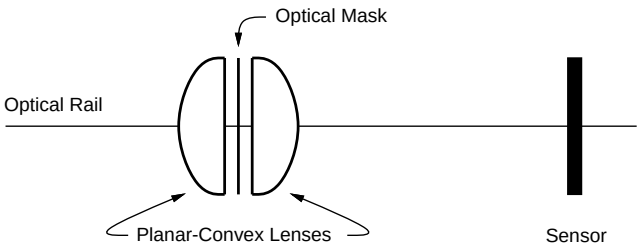

Fig. 4. Experimental camera configuration. Shown are a pair of planar-convex lenses, placed back-to-back, with an optical attenuation mask sandwiched between them. The imaging sensor is mounted on a rail that is aligned with the optical axes of the lenses.

where $I(u,w)$ and $I_v(u,w)$ are the desired quantities for estimating the range image using Equation (21).

There are still several mask design issues that need to be resolved. First, our example of Gaussian-based optical masks was somewhat arbitrary. A pair of masks should be designed from a set of optimality constraints based on derivative accuracy, effective baseline, light transmittance, etc. Once an optimal function is determined, the construction of the actual optical masks must be calibrated to include nonlinearities in the printing process (e.g., halftoning), and the effects of the intrinsic point spread function of the camera. In particular, the image of a point light source recorded by the camera with mask $M'(u,v)$ must be equal to the spatial derivative of the image recorded with mask $M(u,v)$. Finally, noise in the image measurements, $I_1(u,w)$ and $I_2(u,w)$, will be amplified by the computations in Equation (28) and Equation (29): small values of β or γ are thus undesirable.

5 Results

We have constructed a preliminary system for computing aperture derivatives and estimating range. The configuration consists of a pair of planar-convex lenses (50mm diameter, 50mm focal length), a Gaussian-based non-negative optical mask (Figure 3) printed onto a transparency from a 600 dpi laser printer, and a standard CCD sensor array (SONY XC-77R). The optical mask was sandwiched between the pair of lenses, placed back-to-back, and mounted along an optical rail in front of the CCD array (Figure 4).

A pair of images $I_1(u,w)$ and $I_2(u,w)$ are acquired by imaging with the masks M_1 and M_2. A linear combination of the images (Equations (28) and (29)) yields the required $I(u,w)$ and $I_v(u,w)$. The range image is computed from a ratio of $I_v(u,w)$ and the spatial derivative, $I_x(u,w)$, of $I(u,w)$ (Equation (21)). Spatial derivatives were computed using optimized 5×5 derivative filters [Simoncelli 94].

Illustrated in Figure 5 are results from a frontal-parallel target 250 mm from the lens, consisting of an intensity step edge. Shown are the pair of intensity images, $I_1(u,w)$ and $I_2(u,w)$, and the images $I_v(u,w)$ and $I_x(u,w)$. Shown also

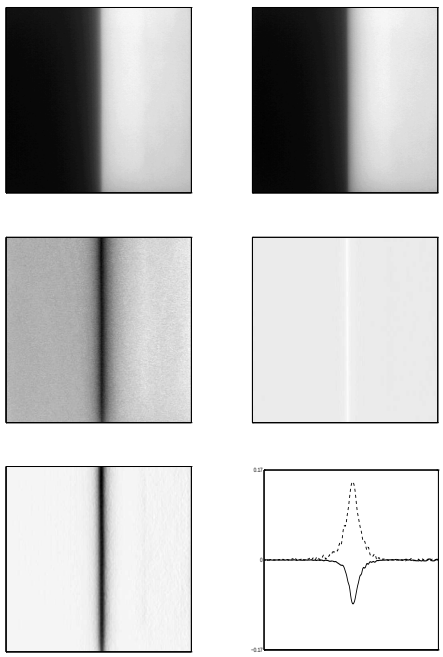

Fig. 5. Top: images obtained through masks $M_1(u,w)$ and $M_2(u,w)$. The scene is a frontal-parallel plane containing an intensity step edge placed 250 mm in front of the focal plane. Middle, left: image $I_v(x,y)$, the derivative with respect to viewing position (computed directly from $M_1(u,w)$ and $M_2(u,w)$). Middle, right: image $I_x(x,y)$, the derivative with respect to image position. Bottom, left: the image $\alpha(x,y)$. Bottom, right: slices through α-map shown at left. Also shown is a slice through another α-map for the same target placed 100mm further from the sensor (behind the focal plane).

is the computed "α-map", where α is monotonically related to the distance to the target. For comparison, the same target was translated by 100mm away from the sensor and the process repeated. Illustrated in Figure 5 are 1-D slices of the computed "α-maps". Notice that since the target moved through the focal plane, the α values are negated, that is, there is no ambiguity between objects equally spaced in front of and behind the focal plane.

6 Range from Aperture Size Derivatives

An interesting variant of the technique arises when considering a Gaussian mask, and its derivative with respect to σ:

$$G(u,w) = \frac{1}{\sigma^2} e^{-(u^2+w^2)/2\sigma^2}, \tag{30}$$

$$\begin{aligned} G_\sigma(u,w) &= \frac{\partial}{\partial\sigma} G(u,w) \\ &= -\frac{2}{\sigma^3} e^{-(u^2+w^2)/2\sigma^2} + \frac{(u^2+w^2)}{\sigma^5} e^{-(u^2+w^2)/2\sigma^2}. \end{aligned} \tag{31}$$

Let $I(x,y)$ and $I_\sigma(x,y)$ be the images obtained through the masks $G(u,w)$ and $G_\sigma(u,w)$, respectively. Using the same techniques as in Section 2, it can be shown that these two images obey the following constraint:

$$\begin{aligned} I_\sigma(x,y) &= \alpha^2\sigma\left[I_{xx}(x,y) + I_{yy}(x,y)\right], \\ &= \alpha^2\sigma\nabla^2 I(x,y), \end{aligned} \tag{32}$$

where $I_{xx}(x,y)$ and $I_{yy}(x,y)$ correspond to the horizontal and vertical second partial derivatives of $I(x,y)$, and ∇^2 is the Laplace operator. As before, α is inversely proportional to range, and is given by Equation (6). This formulation provides a differential algorithm for range-from-defocus. Unlike previous formulations (e.g., [Pentland 87]), this solution avoids the artifacts arising from the computation of local Fourier transforms.

7 Discussion

An optical mask placed in front of a lens-based imaging system produces an image which is a superposition of scaled and dilated copies of the mask function. The derivative of this image is related by a scale factor to a second image created with the derivative of the first optical mask. The scale factor is monotonically related to range. This simple observation has lead us to a *direct* differential method for estimating range from a single stationary camera. In particular, the derivative with respect to viewing position is computed directly: it is simply the image formed under the derivative mask.

Two assumptions have been made in our solution to this problem. Both of these assumptions are made (although often not explicitly) in nearly every structure from stereo or motion algorithm. The first assumption is that the light emanating from each point in the scene is constant across the lens (i.e., the brightness constancy assumption). Note that this assumption will typically be violated at occlusion boundaries, because the light emanating from a partially occluded point will hit only a portion of the lens. One potential solution to this problem is to expand the function describing the light emanating from a point in a Taylor series. The coefficients of these terms may be estimated by collecting additional measurements (i.e., images) with higher-order derivative masks.

The second assumption is that of locally frontal-parallel surface orientation. This assumption was necessary in order to solve for α_p given the two image measurements described by Equations (16) and (17). Solving without this assumption is a nonlinear optimization problem (since the α_p appears inside the argument of $M(\cdot)$), which should be amenable to an iterative solution.

The accuracy of our technique has not yet been tested empirically, but we expect it to be comparable to other single-lens techniques (e.g., [Pentland 87, Adelson 92, Jones 93, Dowski 94]). As with most ranging techniques, accuracy behaves according to the rules of triangulation. In particular, errors will be proportional to the square of the range, and inversely proportional to both the focal

length and baseline.[5] We have verified these relationships via simple simulations. As in many other range-imaging systems, the accuracy may be improved with the use of structured illumination.

A counterintuitive aspect of our technique is that it relies on the defocus of the image. In particular, a perfectly focused image corresponds to $\alpha = 0$, leading to a singularity in Equation (9). In practice, this may be alleviated by focusing the camera at infinity (i.e., $d = f$), thus ensuring that points at distances within the operating range of the algorithm will be sufficiently blurred.

Finally, since the calculations required for recovering range are simple (convolution followed by arithmetic combination of a pair of images), we believe it will be appropriate for real-time implementation. A computer-controlled liquid-crystal lattice could be used in place of a fixed optical attenuation mask, and could be switched back and forth between the two masks at video frame rates.

References

[Adelson 92] Adelson, E.H. and Wang, J.Y.A. Single Lens Stereo with a Plenoptic Camera. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14(2):99–106, 1992.

[Dowski 94] Dowski, E.R. and Cathey, W.T. Single-Lens Single-Image Incoherent Passive-Ranging Systems. *Applied Optics*, 33(29):6762–6773, 1994.

[Horn 86] Horn, B.K.P. *Robot Vision*. MIT Press, Cambridge, MA, 1986.

[Jones 93] Jones, D.G. and Lamb, D.G. *Analyzing the Visual Echo: Passive 3-D Imaging with a Multiple Aperture Camera*. Technical Report CIM-93-3, Department of Electrical Engineering, McGill University, 1993.

[Krotkov 87] Krotkov, E. Focusing. *International Journal of Computer Vision*, 1:223–237, 1987.

[Lucas 81] Lucas, B.D. and Kanade, T. An iterative image registration technique with an application to stereo vision. In *Proceedings of the 7th International Joint Conference on Artificial Intelligence*, pages 674–679, Vancouver, 1981.

[Nayar 95] Nayar, S.K., Watanabe, M., and Noguchi, M. Real-Time Focus Range Sensor. In *Proceedings of the International Conference on Computer Vision*, pages 995–1001, Cambridge, MA, 1995.

[Pentland 87] Pentland, A.P. A New Sense for Depth of Field. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 9(4):523–531, 1987.

[Simoncelli 94] Simoncelli, E P. Design of Multi-dimensional Derivative Filters. In *First International Conference on Image Processing*, Austin, Texas, November 1994.

[Subbarao 88] Subbarao, M. Parallel Depth Recovery by Changing Camera Parameters. In *Proceedings of the International Conference on Computer Vision*, pages 149–155, 1988.

[Xiong 93] Xiong, Y. and Shafer, S. Depth from Focusing and Defocusing. In *Proc. of the DARPA Image Understanding Workshop*, pages 967–976, 1993.

This article was processed using the LaTeX macro package with ECCV'96 style

[5] Effective baseline in our system depends on the mask function, and is proportional to the lens diameter.