A New Approach to 3D Shape Recovery of Local Planar Surface Patches from Shift-Variant Blurred Images

Xue Tu, Murali Subbarao, Youn-Sik Kang
{*tuxue, murali, yskang*}*@ece.sunysb.edu, Dept. of Electrical and Computer Engg.,*
State University of New York, Stony Brook, NY 11794-2350

Abstract

A new approach is presented for 3D shape recovery of local planar surface patches from two shift/space-variant defocused images. It is based on a recently proposed technique for inverting the shift-variant blurring in a camera system. It is completely localized, applicable to general point spread functions, and facilitates fine-grain parallel implementation and computationally efficient recovery of 3D shape. The shape of a visible object at each pixel is specified by the distance and slopes of a local planar surface patch that approximates the object surface at that pixel. The theory underlying the new approach and experimental results are presented. Results of both simulation and real experiments indicate that the new approach is effective and useful in practical applications.

1. Introduction

Image defocus analysis has been investigated by many researchers for 3D shape recovery. In the case of object surfaces normal to the direction of view, blurring is modeled by a shift-invariant/convolution operation. This case has been addressed through both frequency domain and spatial domain techniques ([7], [9], [8], [6], [10]). Recently, the problem of 3D shape recovery from space/shift-variant defocused images has been investigated by several researchers. This problem arises when the surface of an object in the scene is not normal to the direction of view, but slanted and possibly curved. It also arises in the presence of various optical aberrations away from the center of the field of view. Convolution based methods can be used in this case assuming a blockwise/piecewise shift-invariance approximation, but the results would be inaccurate.

Favaro and Soatto [1] formulate the image blurring process as an inner product in Hilbert space and solve the inverse problem using regularized functional singular value decomposition. Rajagopalan and Chaudhury [3] model the focused image and scene related blur parameter as Markov Random fields and recover them through Maximum a Posteriori criteria. Favaro et al. [2] convert shift variant integral equation to inhomogeneous diffusion equation and solve shape related diffusion coefficients through gradient flow approach. The new method presented here is an extension of the S Transform approach [6] for shift-invariant (convolution) blurring. It is based on the recently developed Rao Transform (RT) [4, 5] for image restoration and linear integral equations. In the new method, the computations are completely localized (i.e. solution at a pixel is computed using image data in a small neighborhood around that same pixel), efficient, and suitable for fine-grain parallel implementation. The effectiveness of the new approach is demonstrated with both simulation and real experiments. Results of the experiments show that the new method is useful in practical applications.

2. Shift-variant Blurring

In this section we consider image blurring in a thin lens camera, and derive an interesting result that the blur circle radius R of a planar surface patch on an object is a linear function of image coordinates. This result provides the crucial relation between the shift-variant blur parameter σ proportional to R and the 3D shape parameters of the planar surface patch. Using this relation, we obtain the 3D shape parameters given the blur parameters at each pixel.

Figure 1 shows a schematic diagram of image formation by a thin lens L with diameter D, focal length f, object point P at distance u, focused image p' at distance v, image detector ID at distance s, and blur circle with radius R. Using the lens formula $1/f = 1/u + 1/v$ we can derive

$$R = \frac{s-v}{v}.\frac{D}{2} = s\frac{D}{2}\left(\frac{1}{f} - \frac{1}{u} - \frac{1}{s}\right) \qquad (1)$$

978-1-4244-2175-6/08/$25.00 ©2008 IEEE

In order to compare images captured with different parameters s, image magnification is normalized with $s = s_0$. In this case, the normalized radius R' is given by

$$R' = (D\ s_0/2)(1/f - 1/u - 1/s) \tag{2}$$

Using a world coordinate system with its origin at the

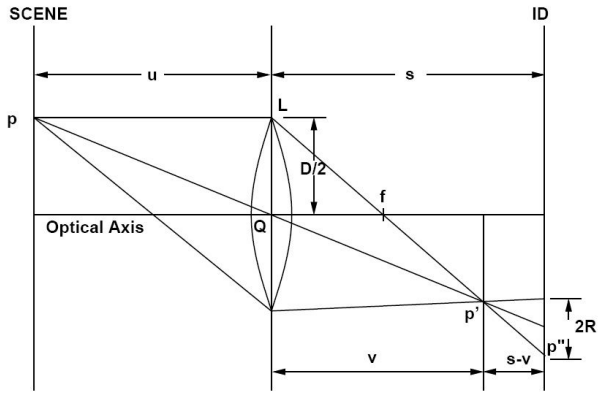

Figure 1. Camera Model

optical center Q and Z-axis along the optical axis, the image p'' of the object point P at (X, Y, Z) has the image coordinates (x, y) given by perspective projection relations:

$$x = s_0 X/Z \qquad \text{and} \qquad y = s_0 Y/Z. \tag{3}$$

The equation of the tangent plane of the object surface at point P can be expressed as

$$Z(X, Y) = Z_0 + Z_X X + Z_Y Y \tag{4}$$

where Z_0 is the distance of the plane along the optical axis, and Z_X and Z_Y are slopes along the X and Y axes respectively. The three parameters Z_0, Z_X, and Z_Y, are the 3D shape parameters to be determined. Using $u = Z(X, Y)$, Eq. (2), Eq. (3), and Eq. (4), we can derive

$$R'(x, y) = R_0 + R_x x + R_y y \tag{5}$$

where

$$\begin{aligned} &R_0 = (Ds_0/2)\ (1/f - 1/Z_0 - 1/s) \\ \text{and}\quad &R_x = DZ_X/(2Z_0), \quad R_y = D\ Z_Y/(2Z_0) \end{aligned} \tag{6}$$

Note that the blur circle radius R' is a linear function of the image coordinates (x, y) for a planar surface patch. The blur spread parameter σ (square root of the second central moment of PSF) for real camera systems is known to be roughly proportional to this radius [6], and therefore this σ can also be approximated as a linear function of the image coordinates. One popular model of point spread function (PSF) corresponding to this blur circle radius is the cylindrical function:

$$h(x, y, u, v) = \begin{cases} 1/(\pi R'^2(x, y)) & \text{for } u^2 + v^2 \le R'^2 \\ 0 & \text{otherwise} \end{cases}$$

The blurred image $g(x, y)$ is related to the focused image $f(x, y)$ through the Rao Transform (RT) [4, 5] relation:

$$g(x, y) = \int_{x-b}^{x-a} \int_{y-d}^{y-c} h(x - u, y - v, u, v) \times f(x - u, y - v)\, du\, dv. \tag{7}$$

The above RT based blurring model is valid for general PSFs, and not restricted to just cylindrical PSFs. Inverting this transform for a general shift-variant PSF (SV-PSF) is presented in the next section.

3. Inverting Shift-Variant Blur

The shift-variant blurring in Eq. (7) can be inverted to solve for the focused image $f(x, y)$ using the RT theory [4, 5]. A summary of the inversion results are presented below. We use superscript (m, n) to denote the m^{th} partial derivative with respect x and n^{th} partial derivative with respect to y, and the subscript p, q to denote $(p, q)^{th}$ moments with respect to u, v, so $h_{p,q}^{(m,n)}$ is defined as:

$$h_{p,q}^{(m,n)} = \int_{x-b}^{x-a} \int_{y-d}^{y-c} u^p v^q h^{(m,n)}(x, y, u, v)\, du\, dv$$

for $m, n, p, q = 0, 1, 2, \ldots$.

Using the above notation, the truncated Taylor series expansion of $f(x - u, y - v)$ around (x, y) up to order N and $h(x-u, y-v, u, v)$ around the point (x, y, u, v) up to order M will be used below. For example, we express

$$h(x - u, y - v, u, v) \approx \sum_{m=0}^{M} a_m \sum_{j=0}^{m} C_j^m u^{m-j} v^j h^{(m-j,j)} \tag{8}$$

where C_i^n and C_j^m denote the binomial coefficients and $a_m = (-1)^m/m!$. Substituting the truncated Taylor-series expansions of h and f into the RT in Eq.(7) and simplifying, we get

$$g(x, y) \approx \sum_{n=0}^{N} \sum_{i=0}^{n} S_{n,i} f^{(n-i,i)} \tag{9}$$

where

$$S_{n,i} = a_n C_i^n \sum_{m=0}^{M} a_m \sum_{j=0}^{m} C_j^m h_{m+n-i-j,i+j}^{(m-j,j)} \quad (10)$$

We can now write expressions for the various partial derivatives of order (p, q) of g with respect to x, y, as

$$g^{(p,q)} \approx \sum_{n=0}^{N} \sum_{i=0}^{n} \frac{\partial^p}{\partial x^p} \frac{\partial^q}{\partial y^q} [S_{n,i} f^{(n-i,i)}] T(n+p+q)$$

where

$$T(n+p+q) = \begin{cases} 1 & \text{if } n+p+q \leq N \\ 0 & \text{otherwise} \end{cases} \quad (11)$$

assures that terms with derivatives of f of order greater than N are set to zero, for $p + q = 0, 1, 2, \ldots, N$.

The above equation for $g^{(p,q)}$ for $p, q = 0, 1, 2, \ldots, N$, and $0 \leq p + q \leq N$ constitute $(N + 1)(N + 2)/2$ equations in as many unknowns $f^{(p,q)}$. The system of equations for $g^{(p,q)}$ can be expressed in a vector-matrix form as

$$\begin{bmatrix} g^{(0,0)} \\ g^{(1,0)} \\ \vdots \\ g^{(0,N)} \end{bmatrix} = \begin{bmatrix} k_{00} & k_{01} & \cdots & \cdots \\ k_{10} & k_{11} & \cdots & \cdots \\ \vdots & \vdots & \ddots & \vdots \\ & \cdots & & \end{bmatrix} \begin{bmatrix} f^{(0,0)} \\ f^{(1,0)} \\ \vdots \\ f^{(0,N)} \end{bmatrix} \quad (12)$$

or

$$\mathbf{g}_{x,y} = \mathbf{K}_{x,y}\, \mathbf{f}_{x,y} \quad (13)$$

where the subscripts (x, y) make explicit the dependence of the vectors/matrix on (x, y). This matrix equation can be solved to obtain $f^{(p,q)}$, by inverting the kernel matrix $\mathbf{K}_{x,y}$. The solution can be written in the form

$$\mathbf{f}_{x,y} = \mathbf{K}'_{x,y}\, \mathbf{g}_{x,y} \quad (14)$$

where $\mathbf{K}'_{x,y} = \mathbf{K}^{-1}_{x,y}$.

4. Shape Recovery

RT (Eq. (13)) and inverse RT (Eq. (14)) provide relations between the shift-variant blurred image $g(x, y)$ and the focused image $f(x, y)$ in terms of the parameters of the Shift-Variant PSF (SV-PSF) $h(x, y, u, v)$ specified by $\mathbf{K}_{x,y}$. Inverse RT (Eq. (14)) can be used to solve for the shape parameters (Z_0, Z_X, Z_Y) using two shift-variant blurred images $g_1(x, y)$ and $g_2(x, y)$ of the object recorded with different camera parameter settings, e.g. $s = s_1$ and $s = s_2$. From Eq. (14), we have

$$\mathbf{f} = \mathbf{K}'_1\, \mathbf{g}_1 \qquad \text{and} \qquad \mathbf{f} = \mathbf{K}'_2\, \mathbf{g}_2. \quad (15)$$

Equating the left hand sides, we obtain

$$\mathbf{K}'_1\, \mathbf{g}_1 = \mathbf{K}'_2\, \mathbf{g}_2. \quad (16)$$

Given the camera parameters D_i, s_i, and f_i, associated with the blurred images g_i for $i = 1, 2$, the matrix $\mathbf{K}'_i$ can be expressed in terms of the shape parameters Z_0, Z_X and Z_Y. The theory here can be extended to include higher order shape parameters such as surface curvatures (e.g. Z_{XX}, Z_{YY} and Z_{XY}) but in practice such extension does not seem to be useful because of the limited gray level and spatial resolution of image data and noise. Curved surfaces can be adequately approximated by planar polygons (e.g. triangles) on local tangent planes at each pixel. Therefore solving Eq. (16) at each pixel (x, y) will provide a solution for 3D shape of the whole object in the scene.

5. Experiments

In practice image derivatives are very noise sensitive. Therefore it is necessary to truncate the Taylor Series expansion of the image function f. Experimental results indicated that $N = 2$ and $M = 1$ would provide good results. Also we assume the SV-PSF $h(x, y, u, v)$ to be rotationally symmetric. In this case, inverse RT in Eq. (14) provides the following simplified expression for the focused image $f(x, y) = f^{(0,0)}$:

$$\begin{aligned} f^{(0,0)} = & g^{(0,0)} + 2h_{2,0}^{(1,0)} h_{0,2}^{(0,1)} g^{(1,1)} \\ & - h_{2,0}^{(1,0)} g^{(1,0)} - h_{0,2}^{(0,1)} g^{(0,1)} \\ & + \left(\frac{3}{2} (h_{2,0}^{(1,0)})^2 + \frac{1}{2} h_{0,2}^{(0,1)} h_{2,0}^{(0,1)} - \frac{1}{2} h_{2,0} \right) g^{(2,0)} \\ & + \left(\frac{1}{2} h_{2,0}^{(1,0)} h_{0,2}^{(1,0)} + \frac{3}{2} (h_{0,2}^{(0,1)})^2 - \frac{1}{2} h_{0,2} \right) g^{(0,2)} \end{aligned}$$

Similar (but shorter) algebraic expressions will be obtained for the various derivatives of $f(x, y)$.

Savitzky-Golay filters are used to compute image derivatives $g_i^{(m,n)}(x, y)$ after smoothing the images. As second order derivatives are highly noise sensitive, especially when the image contrast is low, we use data at only those pixels where the image derivative magnitudes were above a preset threshold. For each pixel with reliable data in a small image block, we equate the right sides of the above equations and form a set of non-linear equations with unknowns Z_0, Z_X and Z_Y. Using data at more than 3 nearby pixels will result in an over-constrained set of non-linear equations. These were solved iteratively in a least square sense to solve for the unknowns.

Both simulation and real experiments on simple objects were carried-out. In simulation experiments, a

Figure 2. Simulated blur images

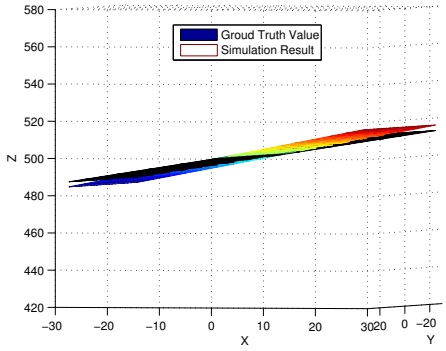

Figure 3. Depth map of Recovered scene and Ground truth

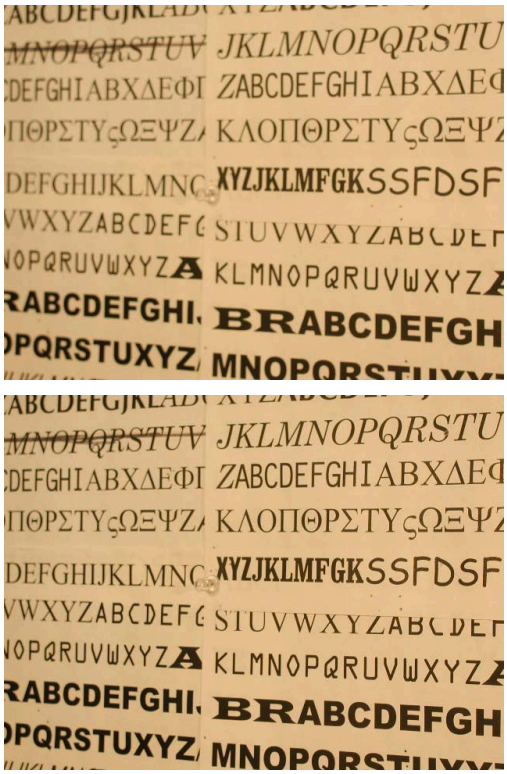

Figure 4. Images taken with different camera settings

slanted planar object with a chessboard pattern was assumed, and two shift-variant blurred images were generated using corresponding SV-PSFs (shown in Figure 2). Then an exhaustive search of the shape parameter space was carried-out to find a solution that minimized an error measure. Solutions in small overlapping image blocks were averaged to obtain the final results. The depth map is presented in Figure 2 along with its ground truth for comparison. The average and maximum error in depth were around 1% and 3% respectively.

In real experiments a slanted textured poster on a plane was recorded with a calibrated digital camera at different focus settings (See Figure 4). Then, using the same approach as in the simulation experiments, the depth map of the object was computed. The result is shown in Figure 5. Images of size 640×480 were processed in blocks of size 30×30 on a 1.6GHz laptop computer with unoptimized code. The processing time was about 5 minutes, most of which was spent on searching for the final solution. This step could be improved substantially.

Results of our experiments show that the new approach presented here is effective and efficient for computing the 3D shape of a planar object using shift-variant blurred images. Unlike approaches based on piecewise constant distance approximation (local convolution or shift-invariant blurring) the results of the new method does not exhibit blocking artifacts in the shape.

This method can be extended from planar to more complex 3D shapes. For example, curvature parameters (second-order partial derivatives Z_{XX}, Z_{XY}, and Z_{YY}) can be included in Eq. (4) by approximating the surface patch as

$$Z(X,Y) = Z_0 + Z_X X + Z_Y Y + Z_{XX} X^2 + Z_{XY} XY + Z_{YY} Y^2. \quad (17)$$

In this case, with two images captured at two different camera parameter settings, image derivatives at six or more (non-degenerate) pixels will need to be used to determine the six shape parameters. Six simultaneous equations corresponding to Eq. (16), one at each pixel, is solved. As an alternative, image derivatives at a single pixel can be used by using seven defocused images as follows. A set of simultaneous equations correponding to Eq. (16) of the form

$$\mathbf{K}'_i \, \mathbf{g}_i = \mathbf{K}'_{i+1} \, \mathbf{g}_{i+1} \quad (18)$$

could be solved for $i = 1, 2, \cdots, 6$ by processing 7

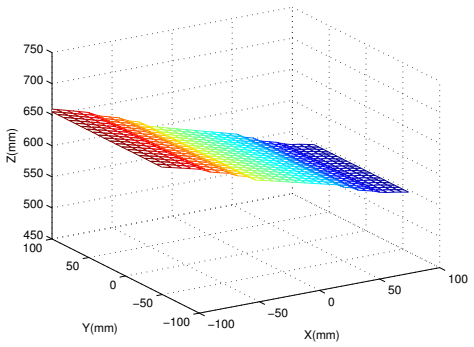

Figure 5. Depth Map of Inclined Plane

images g_i recorded with 7 different (non-degenerate) camera parameter settings. Note that, in addition to the six shape parameters, the focused image at a pixel is the seventh parameter to be determined. In the case of a planar surface patch, there is an alternative to the method presented earlier based on two defocused images at three or more pixels. Four images captured at four different camera parameter settings and image derivatives at a single pixel can be used to obtain three equations from Eq. (18). They can be solved to determine the three shape parameters. These approaches are under investigation.

6. Conclusion

Theory and experimental results are presented for a novel approach to recover 3D shapes of objects from two shift-variant blurred images. Experimental results indicate that the method is effective and efficient. In addition to shape, the new method can be used to compute the focused image of the object from its blurred images [5]. The method here will be improved further using regularization and better iterative techniques for solving non-linear system of equations. The method here represents a basic theoretical and computational advance in 3D scene recovery from shift-variant blurred images.

References

[1] P. Favaro and S. Soatto. A geometric approach to shape from defocus. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 27(3):406 – 417, March 2005.

[2] P. Favaro, S. Soatto, M. Burger, and S. Osher. Shape from defocus via diffusion. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 30(3):518 – 531, May 2008.

[3] A. N. Rajagopalan and S. Chaudhuri. Optimal recovery of depth from defocused images using an MRF model. In *ICCV*, pages 1047–1052, 1998.

[4] M. Subbarao. U.S. Patent Application No. 11/235724, Filed on 9/26/05, U.S. Patent Application No. 11/242191, Filed on 10/03/2005, and U.S. Patent Application No. 11/450024, 6/10/2006. *Rao Transforms...* , U.S. Copyright No. TX 6-195-821, June 1, 2005. See http://www.integralresearch.net.

[5] M. Subbarao, Y. Kang, S. Dutta, and X. Tu. Localized and computationally efficient approach to shift-variant image deblurring. In *IEEE International Conference on Image Processing*, 2008.

[6] M. Subbarao and G. Surya. Depth from defocus: A spatial domain approach. *International Journal of Computer Vision*, 13(3):271–294, 1994.

[7] M. Watanabe and S. K. Nayar. Rational filters for passive depth from defocus. *International Journal of Computer Vision*, 27(3):203–225, 1998.

[8] T. Wei. Three-dimensional machine vision using image defocus. Ph.D. Thesis, Dept. of Electrical Eng., SUNY at Stony Brook, 1994.

[9] Y. Xiong and S. Shafer. Depth from focusing and defocusing. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 68–73, 1993.

[10] D. Ziou and F. Deschenes. Depth from defocus estimation in spatial domain. *Computer Vision and Image Understanding: CVIU*, 81(2):143–165, 2001.