

Depth and Focused Image Recovery from Defocused Images for Cameras Operating in Macro Mode

Xue Tu, Youn-sik Kang and Murali Subbarao

Dept. of Electrical and Computer Engineering
State University of New York at Stony Brook
Stony Brook, NY, USA 11794-2359

ABSTRACT

Depth From Defocus (DFD) is a depth recovery method that needs only two defocused images recorded with different camera settings. In practice, this technique is found to have good accuracy for cameras operating in Normal Mode. In this paper, we present new algorithms to extend the DFD method to cameras working in Macro Mode used for very close objects in a distance range of 5 cm to 20 cm. We adopted a new lens position setting suitable for Macro Mode to avoid serious blurring. We also developed a new calibration algorithm to normalize magnification of images captured with different lens positions. In some range intervals with high error sensitivity, we used an additional image to reduce the error caused by drastic change of lens settings. After finding the object depth, we used the corresponding blur parameter for computing the focused image through image restoration, which is termed as "soft-focusing". Experimental results on high-end digital camera show that the new algorithms significantly improve the accuracy of DFD in the Macro Mode. In terms of focusing accuracy, the RMS error is about 15 lens steps out of 1500 steps, which is around 1%.

Keywords: Depth from Defocus (DFD), Autofocusing, Macro Mode, Magnification Normalization, Image Restoration

1. INTRODUCTION

Depth-from-Defocus (DFD) technique is an elegant passive autofocus method. It only needs two or three images recorded with different camera settings to recover the depth of certain object by computing the degree of blurring.

There are generally two categories of DFD algorithms: statistical and deterministic. Statistical approaches like Maximum likelihood¹ and Markov Random field methods² require more image computing. Deterministic algorithms can be classified as frequency domain approaches^{3,4} and spatial domain approaches.⁵⁻⁷ The frequency domain approaches are generally computation expensive and yield lower depth-map density. While spatial domain approaches only use a small image region, thus require less computation and generate piecewise depth-map. Due to the inherent advantage of being local in nature, spatial domain approach is more suitable for real-time autofocus applications.

Subbarao proposed a Spatial-domain Convolution/Deconvolution Transform (S Transform)⁸ for n-dimensional signals for the case of arbitrary order polynomials. Surya and Subbarao⁵ utilized S Transform to estimate the blur parameter in spatial domain, named as STM. There are two basic variations of STM: STM1 changes the lens position and STM2 varies the aperture diameter.

In practice, STM1 has been found to have good accuracy for cameras operating in Normal Mode used for objects at a distance of 20 cm or more.⁹ When cameras work under macro-mode used for objects in a close distance between 5cm to 20cm, many difficulties arise, such as highly blurred images, significant magnification change for small lens motion, and specific settings in lens parameters. To extend STM1 to macro-mode, we adopt a new lens setting suitable for macro-mode to avoid serious blurring. We also develop new calibration algorithm to normalize magnification of images captured with different lens positions. In some range intervals

Further author information: (Send correspondence to Xue Tu)

E-mail: {txue, yskang, murali}@ece.sunysb.edu; Telephone: 1-631-632-9149; WWW: www.ece.sunysb.edu/ cvl

with high error sensitivity, we used an additional image to reduce the error caused by drastic change of lens settings. Experimental results on different test object data show that new algorithm significantly improve the accuracy of STM1 for macro mode. The RMS error is about 15 lens steps out of 1500 steps, which is around 1%. Apart from setting the lens at corresponding position to capture a focused image after finding the object depth, we also present an algorithm to compute the focused image through blur parameter, which is termed as "soft focusing".

2. STM OVERVIEW

The basic theory of STM is reviewed here to introduce relevant formulas and define terms for future discussion.

2.1 S Transform

A new Spatial-Domain Convolution\Deconvolution Transform (S Transform) was proposed for images and n-dimensional signals⁸ for the case of arbitrary order polynomials.

Let $f(x, y)$ be an image. Within a small region, we approximate it as a two variable cubic polynomial:

$$f(x, y) = \sum_{m=0}^3 \sum_{n=0}^{3-m} a_{m,n} x^m y^n, \quad (1)$$

where $a_{m,n}$ are the polynomial coefficients. This assumption of f can be made valid by applying a polynomial fitting least square smoothing filter to the image.

Let $h(x, y)$ be a rotationally symmetric point spread function (PSF). The moments of the point spread function are defined by

$$h_{m,n} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x^m y^n h(x, y) dx dy. \quad (2)$$

The observed image $g(x, y)$ is the convolution of focused image $f(x, y)$ and PSF $h(x, y)$:

$$\begin{aligned} g(x, y) &= f(x, y) * h(x, y) \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x - \zeta, y - \eta) h(\zeta, \eta) d\zeta d\eta. \end{aligned} \quad (3)$$

Since $f(x, y)$ is a cubic polynomial, it can be expressed in a Taylor series as

$$f(x - \zeta, y - \eta) = \sum_{0 \leq m+n \leq 3} \frac{(-\zeta)^m}{m!} \frac{(-\eta)^n}{n!} f^{m,n}(x, y), \quad (4)$$

then $g(x, y)$ can be expressed as

$$g(x, y) = \sum_{0 \leq m+n \leq 3} \frac{(-1)^{m+n}}{m!n!} f^{m,n}(x, y) h_{m,n}, \quad (5)$$

Equation (5) expresses the convolution as a summation involving the derivatives of $f(x, y)$ and moments of $h(x, y)$. This corresponds to the forward S-Transform.

Applying the circular symmetric property of $h(x, y)$, we can rearrange Equation (5) as

$$f(x, y) = g(x, y) - \frac{h_{2,0}}{2} (f^{2,0}(x, y) + f^{0,2}(x, y)), \quad (6)$$

applying partial derivatives to the above equation on both sides, we obtain

$$f^{2,0}(x, y) = g^{2,0}(x, y), \quad (7)$$

and

$$f^{0,2}(x, y) = g^{0,2}(x, y). \quad (8)$$

By the definition of moments of h , we have $h_{2,0} = h_{0,2} = \sigma_h^2/2$, so the above deconvolution equation can be written as:

$$f(x, y) = g(x, y) - \frac{\sigma_h^2}{4} \nabla^2 g(x, y). \quad (9)$$

Equation (9) is termed as Inverse S-Transform. In the following section, we will describe the application of this formula to distance estimation of blurred images.

2.2 STM Autofocusing

A schematic diagram of a camera system is shown in Fig. 1. If an object point p is not focused, then a blur circle p'' is detected on the image detector plane. The radius of the blur circle can be calculated as:

$$R = \frac{Ds}{2} \left(\frac{1}{f} - \frac{1}{u} - \frac{1}{s} \right), \quad (10)$$

where f is the effective focal length, D is diameter of lens aperture, R is radius of blur circle. u , v and s are object distance, image distance and detector distance respectively.

We assume all the images have been normalized both in magnification and brightness, then normalized radius of blur circle can be expressed as

$$R' = \frac{Rs_0}{s} = \frac{Ds_0}{2} \left(\frac{1}{f} - \frac{1}{u} - \frac{1}{s} \right). \quad (11)$$

Since second moment of PSF is related to R' as $\sigma^h = kR'$ for some constant $k(k \approx \sqrt{2})$, we obtain

$$\sigma^h = mu^{-1} + c, \quad (12)$$

where

$$m = -\frac{kDs_0}{2} \quad \text{and} \quad c = -\frac{kDs_0}{2} \left(\frac{1}{f} - \frac{1}{s} \right) \quad (13)$$

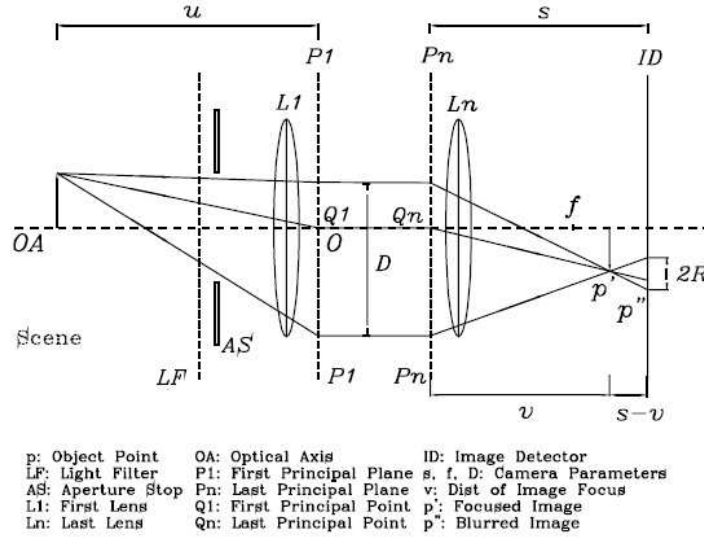


Figure 1. Schematic diagram of a camera system

Let $f(x, y)$ be the focused image of an object at distance u , $g_1(x, y)$ and $g_2(x, y)$ be two images of the object recorded for two different camera parameter settings $\mathbf{e}_1 = (s_1, f_1, D_1)$ and $\mathbf{e}_2 = (s_2, f_2, D_2)$, we obtain:

$$\begin{aligned}\sigma_1 &= m_1 u^{-1} + c_1, \\ \sigma_2 &= m_2 u^{-1} + c_2.\end{aligned}\tag{14}$$

Rewriting Eqs. (14), σ_1 can then be expressed in terms of σ_2 as

$$\sigma_1 = \alpha \sigma_2 + \beta,\tag{15}$$

where

$$\alpha = \frac{m_1}{m_2} \quad \text{and} \quad \beta = c_1 - c_2 \frac{m_1}{m_2}.\tag{16}$$

Applying Inverse S Transform 9 on both observed images g_1 and g_2 , we obtain:

$$f = g_i - \frac{1}{4} \sigma_i^2 \nabla^2 g_i. \quad i = 1, 2\tag{17}$$

Equating the right sides of both equations, we have

$$g_1 - g_2 = \frac{1}{4} G \nabla^2 g,\tag{18}$$

where

$$G = \sigma_1^2 - \sigma_2^2 = \frac{4(g_1 - g_2)}{\nabla^2 g}.\tag{19}$$

Substituting for σ_1 in terms of σ_2 from Equation (15) and using the definition of G in Equation (19), we obtain

$$\sigma_2^2 (\alpha^2 - 1) + 2\alpha\beta\sigma_2 + \beta^2 = G,\tag{20}$$

where α and β are defined in Equation (16).

In STM1, the aperture diameter is not changed but lens position is moved during acquisition of two images g_1 and g_2 , (i.e., $f_1 \neq f_2$ and $s_1 \neq s_2$, but $D_1 = D_2$), then

$$\alpha = \frac{m_1}{m_2} = \frac{D_1}{D_2} = 1,\tag{21}$$

So the quadratic equation becomes a linear equation, and we get the solution directly:

$$\sigma_2 = \frac{G}{2\beta} - \frac{\beta}{2}.\tag{22}$$

3. MAGNIFICATION NORMALIZATION

As we have stated in Section 2.2, the observed images g_1 and g_2 should be normalized in terms of both brightness and magnification. In practical implementation, as magnification change is less than 3% for most of the cases, it is usually ignored in favor of computational speed. However, when a camera is working in macro-mode (for objects closer than 20cm) instead of the standard mode (for objects farther than 20 cm), such change (shown in Fig. 2) has to be taken into account as the shift between corresponding points could be up to 20 pixels. To ensure a correct comparison between two images, image normalization becomes a necessity.

The relation between two images' magnifications can be expressed as

$$\begin{bmatrix} x^{(2)} \\ y^{(2)} \\ z^{(2)} \end{bmatrix} = \begin{bmatrix} s & 0 & s_x \\ 0 & s & s_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x^{(1)} \\ y^{(1)} \\ z^{(1)} \end{bmatrix}\tag{23}$$

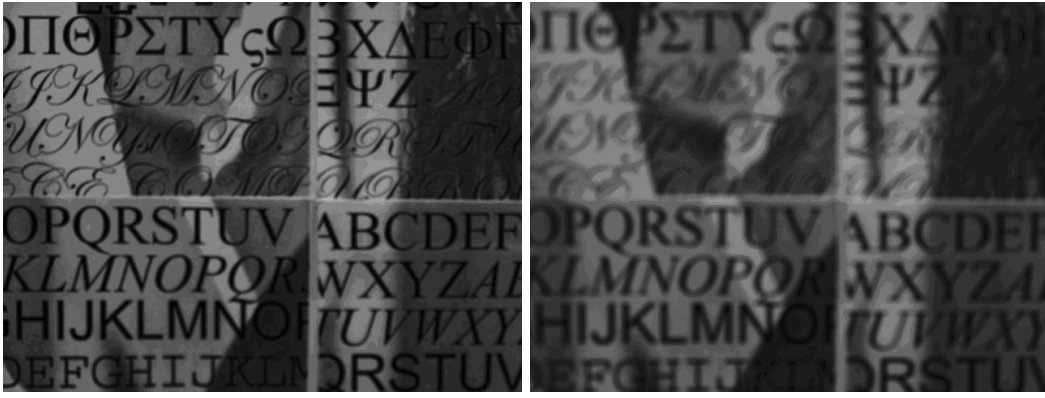


Figure 2. Images taken with different camera settings in macro mode

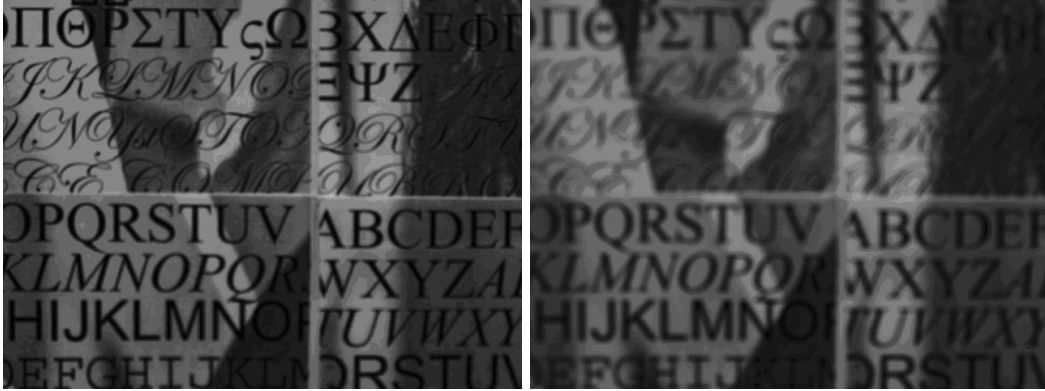


Figure 3. Images with Normalized Magnification

where $[x^{(1)}, y^{(1)}, z^{(1)}]^t$ and $[x^{(2)}, y^{(2)}, z^{(2)}]^t$ are homogeneous coordinates of pixels in observed images g_1 and g_2 .

To get the transformation matrix between magnification of two images, we need to calculate three coefficients s, s_x, s_y from calibration. One method is to take pictures of a simple pattern object (like black dots on white paper), and get two sets of corresponding pixel coordinates from pattern images g_1 and g_2 . Then we can form a group of over constrained equations shown below, which can be solved through a least square criterion.

$$\begin{bmatrix} x_1^{(1)} & 1 & 0 \\ y_1^{(1)} & 0 & 1 \\ x_2^{(1)} & 1 & 0 \\ y_2^{(1)} & 0 & 1 \\ x_3^{(1)} & 1 & 0 \\ y_3^{(1)} & 0 & 1 \\ \vdots & & \\ x_n^{(1)} & 1 & 0 \\ y_n^{(1)} & 0 & 1 \end{bmatrix} \begin{bmatrix} s \\ s_x \\ s_y \end{bmatrix} = \begin{bmatrix} x_1^{(2)} \\ y_1^{(2)} \\ x_2^{(2)} \\ y_2^{(2)} \\ x_3^{(2)} \\ y_3^{(2)} \\ \vdots \\ x_4^{(2)} \\ y_4^{(2)} \end{bmatrix}$$

If we write it as $Ax = b$, then x could be solved as:

$$x = (A^t A)^{-1} A^t b \quad (24)$$

After getting the transformation matrix, the next step is to warp the target image (e.g., g_1 , usually the one with broader view field). As forward warping usually maps integer pixel coordinates to sub pixel coordinates, here a backward warping scheme is applied. Integer pixel coordinates in transformed image are projected back

to the original image, usually resulting in sub pixel position. Then its gray scale value will be computed by linear interpolation. The result of normalized images after warping are shown in Fig. 3.

4. NEW LENS SETTINGS

Besides obvious magnification change, another problem in macro-mode DFD is that images are extremely blurred. In this case, the estimated Laplacian of observed image ∇g is noisy and unreliable (SNR will be low as Laplacian magnitude is reduced by high levels of blur). We have to change the camera settings to make the images more focused.

Another problem is relation between σ and distance u does not quite follow Equation (14) due to change of lens system. We have to calibrate the camera to obtain a new lookup table specifically for macro-mode autofocusing. The lookup table has two entries: step number and blur parameter σ . Step number is a measure of lens position, which is linear with respect to reciprocal of focused object distance $1/u$. We use Depth-from-Focus(DFE)¹⁰ method to get the step number corresponding to a certain object distance. It is a procedure that acquires many images and search the best camera parameter according to some focus measure. It needs much more computation than DFD but usually considered to have higher accuracy.

We put objects at several certain distances and compute the corresponding blur parameter σ and step number. After establishing the lookup table, we will use linear interpolation to get the appropriate step number for any blur parameter σ calculated from Equation (22), thus recover the depth of object.

Due to some special change of lens position when camera works under macro-mode, mapping from σ to step number is not linear or monotonic at some certain sensitive region (shown in Fig. 4), which makes the looking up task difficult. Under such circumstances, we take another picture with new camera setting and use a new lookup table to get the step number as the new lookup table's sensitive region differs from the old one. This combining strategy will effectively improve the accuracy as each of their sensitive regions are avoided. Result of this 3 image DFD approach is shown in Section 6.

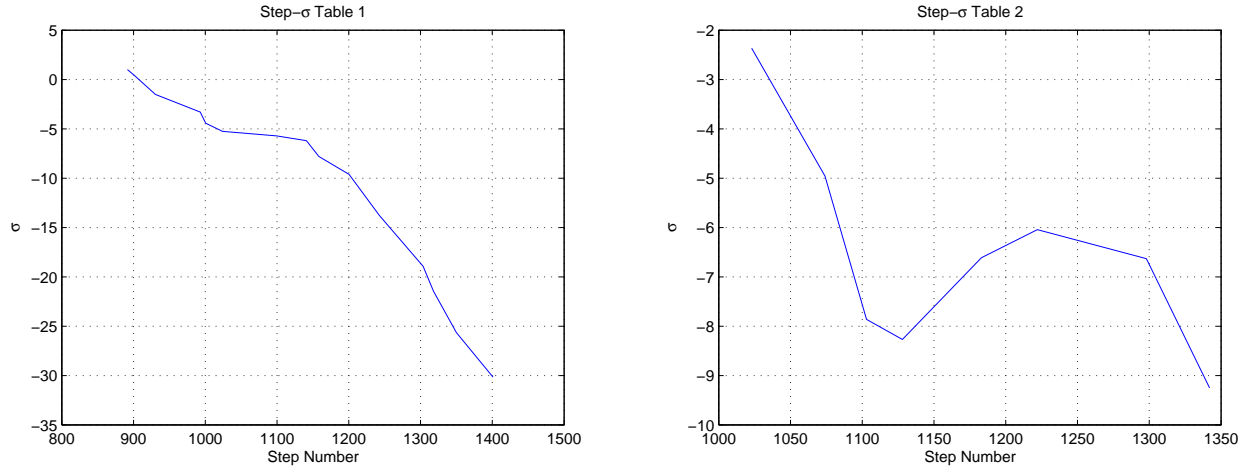


Figure 4. Step- σ Table for Different Camera Settings

5. SOFT FOCUSING

After getting the blur parameter σ from Equation (22), instead of getting the corresponding step number through lookup table and set lens to focus position to get high quality image, we can also apply Equation (9) to get focused image through image restoration, which we term as Soft Focusing. In DFD algorithm, the blurring process is modeled as convolution of focused image and point spread function, which implicitly assumes that the object is flat and perpendicular to camera's optical axis. when object does not have a regular shape, we can split it into small pieces and perform DFD on each one. Though it will introduce some artifacts, yet it provides

another method for image restoration. When object is flat and parallel to lens, soft focusing will provide very good results (shown in Fig. 11 and Fig. 13). Another important advantage of soft focusing is that it only needs very few computations. Compared with some complex restoration algorithms, like iterative optimization or SVD, soft focusing only needs some basic arithmetical calculations, which will make it suitable for hardware implementation.

6. EXPERIMENTAL RESULTS

We implement our experiment on Olympus E-1 SLR camera. The camera is controlled by a host computer(Pentium 4 2.4GHz) from a USB port. The lens' focus motor ranges from step 0 to step 1500. Step 0 corresponds to focusing a distant object and step 1500 corresponds to focusing a nearby object.

The performance of modified DFD algorithms for macro-mode is evaluated with real experiments using different objects shown in Fig. 5. Each object is placed at eight different distances in the range of 5cm to 25cm at roughly 2cm intervals. The distance and the corresponding steps are listed in Table 1. The steps were obtained using the DFF algorithm. Two images were taken with camera set to step 800 and step 1100. If the focus step calculated by DFD is between 1000 and 1150, then a third picture is taken with camera set to step 1250 and use the new lookup table to retrieve the depth. The focusing window was set to 96×96 located at the center of the scene. Before performing DFD, all the images are normalized with respect to magnification and brightness and smoothed by a Gaussian filter. The image Laplacians were threshold to weed out low contrast pixels with low SNR.

Position	1	2	3	4	5	6	7	8	9
Distance(mm)	250	225	200	180	150	120	100	80	55
Step	786	806	893	902	993	1099	1158	1217	1401

Table 1. Distance-Step Table

First we will show the RMS error of 2 Image DFD and 3 Image DFD for objects at different distances. All the objects here are flat and parallel to lens system, so computing depth from central focus window is sufficient. For 2 Image DFD, RMS errors are quite small for most regions, except for the certain range ($u = 10$ cm to 12.5 cm or focus step position around 1100), which is the sensitive region shown in Fig.4. If we combine the result from the third image, we find that the results are largely improved. Most RMS errors are smaller than 15 steps out of 1500 steps, the biggest error is just 22 steps, the average error is about 1%.

Then we present the result on inclined flat object (shown in Fig. 8). We split the image into small pieces and compute depth separately by setting focus window onto each block. The magnification normalization procedure plays an important role here as magnification can change up to 20 pixels near image borders. Shape Reconstruction result with magnification normalization is largely improved (shown in Fig. 9), especially around the edges, the surface becomes much smoother if magnification is normalized before performing DFD.

At last, we will present some image restoration result done by soft focusing. Still, objects are placed at certain distance and two pictures are taken with camera set to step 800 and step 1100 (See Fig. 10 and Fig. 12). Then blur parameter σ is calculated by Equation (22). Then apply Equation (9) on each pixel to get the corresponding focused images (shown in Fig. 11 and Fig. 13). The restoration results are satisfactory.

7. CONCLUSION

A modified DFD algorithm is presented here for cameras working under macro-mode. Before performing DFD, we first normalize the magnification of both images to ensure the correct comparison between corresponding pixels, and results have been largely improved especially for areas far away from the center. We also use new lens settings and lookup table specifically for macro-mode, under some certain circumstances, a third image is taken and depth is retrieved through a second lookup table. Such 3 image DFD scheme is proved to be very accurate, the RMS error is around 15 steps out of 1500 steps, about 1% on average. We also present a soft focusing approach for image restoration from two defocused images, which is a fast and effective scheme.



Figure 5. Test Objects

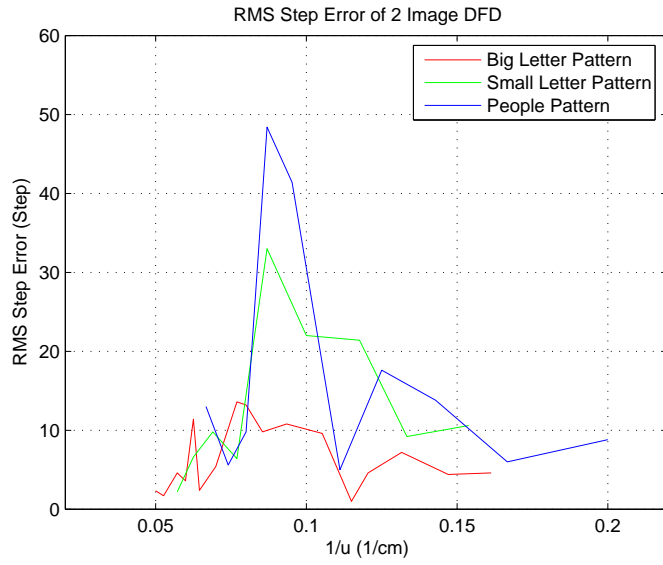


Figure 6. RMS Error of 2 Images DFD

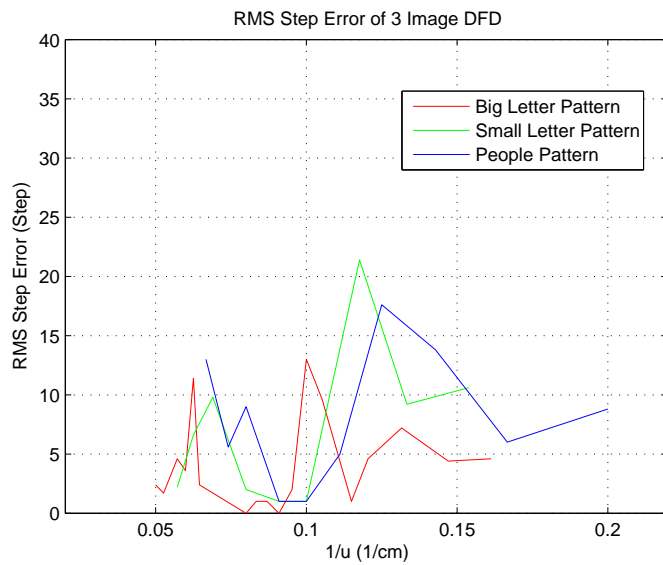


Figure 7. RMS Error of 3 Images DFD

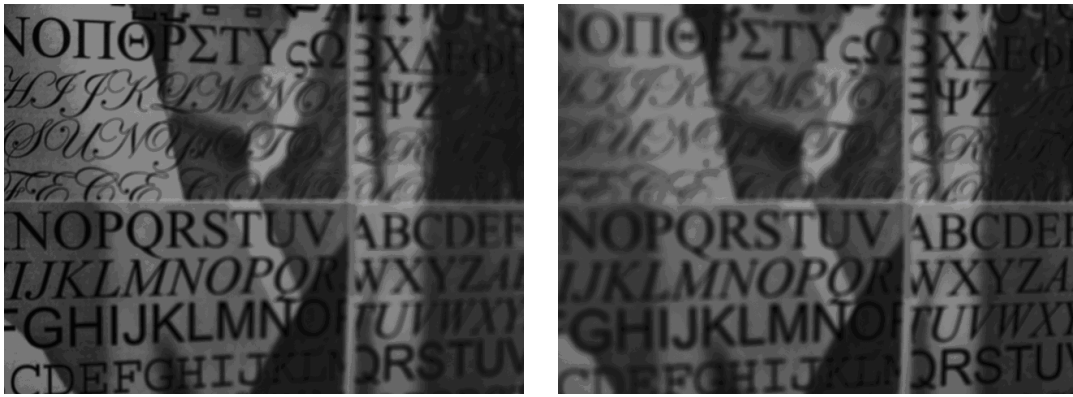
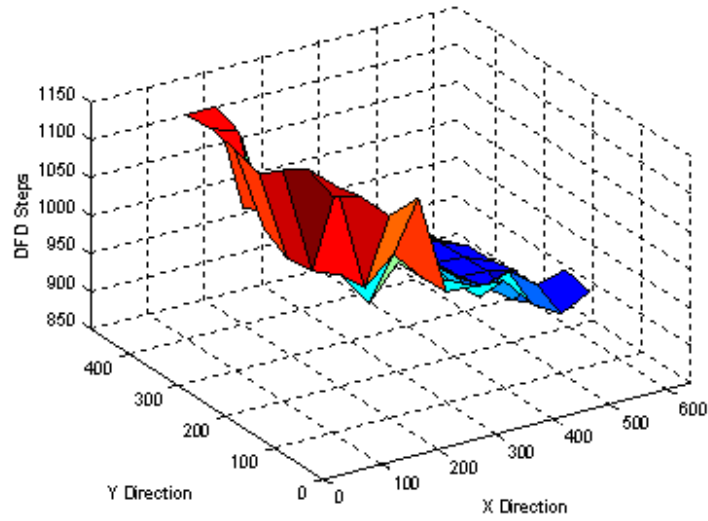


Figure 8. Images of Inclined Flat Poster

DFD Result without Magnification Correction



DFD Result with Magnification Correction

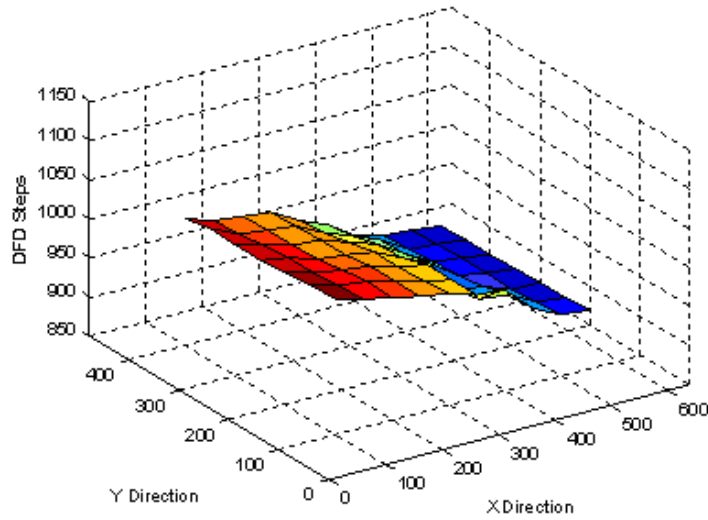


Figure 9. Shape of Inclined Poster from Two Schemes

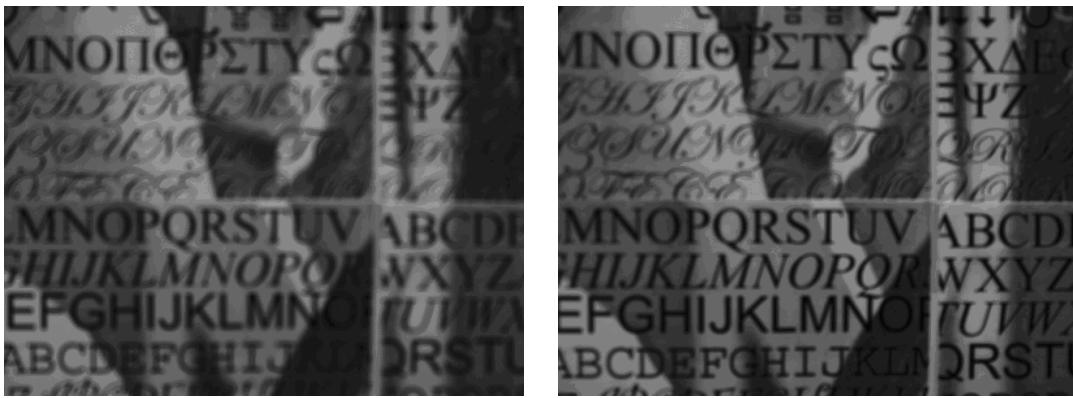


Figure 10. Letter Images taken at Step 800 and Step 1100

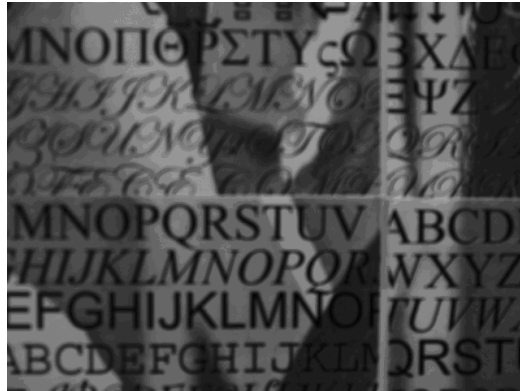


Figure 11. Restoration Result by Soft Focusing

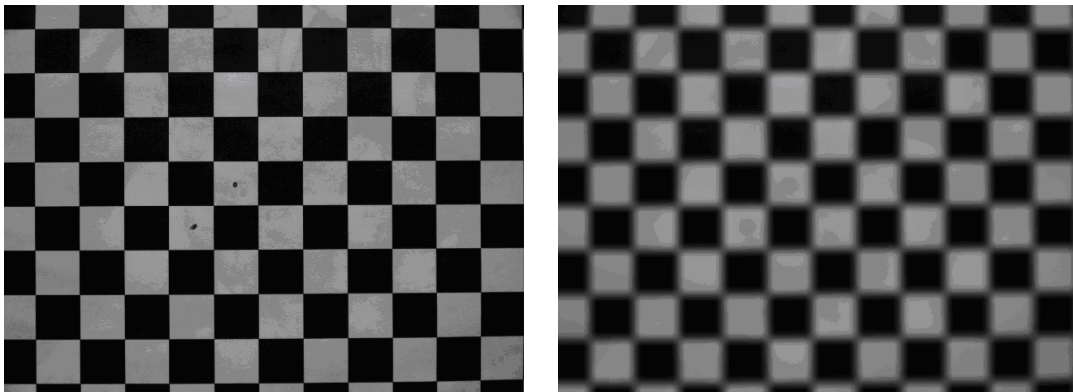


Figure 12. Checkerboard Pattern taken at Step 800 and Step 1100

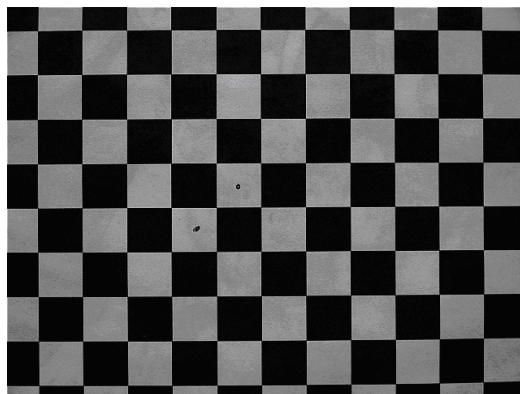


Figure 13. Restoration Result by Soft Focusing

REFERENCES

1. A. N. Rajagopalan and S. Chaudhuri, "A recursive algorithm for maximum likelihood-based identification of blur from multiple observations," 1998.
2. A. N. Rajagopalan and S. Chaudhuri, "Optimal recovery of depth from defocused images using an MRF model," in *ICCV*, pp. 1047–1052, 1998.
3. Y. Xiong and S. Shafer, "Depth from focusing and defocusing," in *DARPA93*, pp. 967–, 1993.
4. M. Watanabe and S. Nayar, "Rational filters for passive depth from defocus," Tech. Rep. CUCS-035-95, Dept. of Computer Science, Columbia University, Sept. 1995.
5. M. Subbarao and G. Surya, "Depth from defocus: A spatial domain approach," *IJCV* **13**, pp. 271–294, December 1994.
6. D. Ziou and F. Deschenes, "Depth from defocus estimation in spatial domain," *Computer Vision and Image Understanding: CVIU* **81**(2), pp. 143–165, 2001.
7. P. Favaro, A. Mennucci, and S. Soatto, "Observing shape from defocused images," *International Journal of Computer Vision* **52**(1), pp. 25–43, 2003.
8. M. Subbarao, "Spatial-domain convolution/deconvolution transform," Tech. Rep. 91.07.03, Computer Vision Laboratory, Dept. of Electrical Engineering, State University of New York, 1991.
9. T. Xian and M. Subbarao, "Performance evaluation of different depth from defocus (dfd) techniques," in *Proceedings of SPIE, Two- and Three-Dimensional Methods for Inspection and Metrology III*, **6000**, 2005.
10. M. Subbarao and J.-K. Tyan, "Selecting the optimal focus measure for autofocusing and depth-from-focus," *IEEE Transactions on Pattern Analysis and Machine Intelligence* **20**(8), pp. 864–870, 1998.