Generalized Laplacian as Focus Measure

Muhammad Riaz¹, Seungjin Park², Muhammad Bilal Ahmad¹, Waqas Rasheed¹,
and Jongan Park¹

¹ School of Information & Communications Engineering, Chosun University,
501-759 South Korea

² Dept of Biomedical Engineering, Chonnam National University Hospital, Kwangju,
South Korea
japark@chosun.ac.kr

Abstract. Shape from focus (SFF) uses focus measure operator for depth measurement from a sequence of images. From the analysis of defocused image, it is observed that the focus measure operator should respond to high frequency variations of image intensity and produce maximum values when the image is perfectly focused. Therefore, an effective focus measure operator must be a high-pass filter. Laplacian is mostly used as focus measure operator in the previous SFF methods. In this paper, generalized Laplacian is used as focus measure operator for better 3D shape recovery of objects.

Keywords: Shape from focus, SFF, Laplace filter, 3D shape recovery.

1 Introduction

The well-known examples of passive techniques for 3D shape recovery from images include shape from focus (SFF). Shape From Focus (SFF) [1], [2] for 3D shape recovery is a search method which searches the camera parameters (lens position and/or focal length) that correspond to focusing the object. The basic idea of image focus is that the objects at different distances from a lens are focused at different distances.

Fig. 1 shows the basic image formation geometry. In SFF, the camera parameter setting, where the blur circle radius R is zero is used to determine the distance of the object. In Fig. 1, if the image detector (ID) is placed exactly at a distance v , sharp image P' of the point P is formed. Then the relationship between the object distance u , focal distance of the lens f , and the image distance v is given by the Gaussian lens law:

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

Once the best-focused camera parameter settings over every image point are determined, the 3D shape of the object can be easily computed. Note that a sensed image is in general quite different from the focused image of an object. The sensors

Fig. 1. Image formation of a 3D object

are usually planar image detectors such as CCD arrays; therefore, for curved objects only some parts of the image will be focused whereas other parts will be blurred.

In SFF, an unknown object is moved with respect to the imaging system and a sequence of images that correspond to different levels of object focus is obtained. The basic idea of image focus is that the objects at different distances from a lens are focused at different distances. The change in the level of focus is obtained by changing either the lens position or the focal length of the lens in the camera. A focus measure is computed in the small image regions of each of the image frame in the image sequence. The value of the focus measure increases as the image sharpness or contrast increases and it attains the maximum for the sharpest focused image. Thus the sharpest focused image regions can be detected and extracted. This facilitates auto-focusing of small image regions by adjusting the camera parameters (lens position and/or focal length) so that the focus measure attains its maximum value for that image region. Also, such focused image regions can be synthesized to obtain a large image where all image regions are in focus. Further, the distance or depth of object surface patches that correspond to the small image regions can be obtained from the knowledge of the lens position and the focal length that result in the sharpest focused images of the surface patches.

A lot of research has been done on the image focus analysis to automatically focus the imaging system [6], [7] or to obtain the sparse depth information from the observed scene [2], [3], [4], [8], [9]. Most previous research on Shape From Focus (SFF) concentrated on the developments and evaluations of different focus measures [1], [9]. From the analysis of defocused image [1], it is shown that the defocusing is a LFP, and hence, focus measure should respond to high frequency variations of image intensity and produce maximum values when the image is perfectly focused. Therefore, most of the focus measure in the literature [1], [9] somehow maximizes the high frequency variations in the images. The common focus measure in the literature

are; maximize high frequency energy in the power spectrum using FFT, variance of image gray levels, L1-norm of image gradient, L2-norm of image gradient, L1-norm of second derivatives of image, energy of Laplacian, Modified Laplacian [2], histogram entropy of the image, histogram of local variance, Sum-Modulus-Difference, etc. There are other focus measures based on moments, wavelet, DCT and median filters.

The traditional SFF (SFFTR) [2] uses modified Laplacian as focus measure operator. There are spikes in the 3D shape recovery using modified Laplacian. Laplacian and modified Laplacian operators are fixed and are not suitable in every situation [5]. In this paper, we have used generalized Laplacian as focus measure operator which can be tuned for the best 3D shape results.

This paper is organized as follows. Section 2 describes the image focus and defocus analysis and the traditional SFF method. Section 3 describes the generalized Laplacian and simulation results are shown in section 5.

2 Image Focus and Defocus Analysis

If the image detector (CCD array) coincides with the image plane (see Fig. 1) a clear or focused image $f(x,y)$ is sensed by the image detector. Note that a sensed image is in general quite different from the focused image of an object. The sensors are usually planar image detectors such as CCD arrays; therefore, for curved objects only some parts of the image will be focused whereas other parts will be blurred. The blurred image $h(x,y)$ usually modeled by the PSF of the camera system. In a small image region if the imaged object surface is approximately a plane normal to the optics axis, then the PSF is the same for all points on the plane. The defocused image $g(x,y)$ in the small image region on the image detector is given by the convolution of the focused image with the PSF of the camera system, as:

$$g(x, y) = h(x, y) \otimes f(x, y) \quad (2)$$

where the symbol $\otimes$ denotes convolution.

Now we consider the defocusing process in the frequency domain (ω). Let F , H and G be the Fourier Trans-forms of the functions f , h and g respectively. Then, we can express equ. (2) in the frequency domain by knowing the fact that the convolution in the spatial domain is the multiplication in the frequency domain, as:

$$G(w_1, w_2) = H(w_1, w_2) \cdot F(w_1, w_2) \quad (3)$$

The Gaussian PSF model is a very good model of the blur circle. So the PSF of the camera system can be given as:

$$h(x, y) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{x^2 + y^2}{2\sigma^2}\right) \quad (4)$$

The spread parameter σ is proportional to the blur radius R in Fig. 1. The Fourier Transform of PSF is OTF of the camera system and is given as:

$$H(w_1, w_2) = \exp\left(-\frac{w_1^2 + w_2^2}{2} \sigma^2\right) \quad (5)$$

We note that low frequencies are passed un-attenuated, while higher frequencies are reduced in amplitude, significantly so for frequencies above about $1/\sigma$. Now σ is a measure of the size of the original PSF; therefore, the larger the blur, the lower the frequencies that are attenuated. This is an example of the inverse relationship between scale changes in the spatial domain and corresponding scale changes in the frequency domain. In fact the product $R''\rho$ is constant, where R'' is the blur radius in the spatial domain, and ρ is the radius in its transform. Hence, defocusing is a low-pass filtering process where the bandwidth decreases with increase in defocusing.

A defocused image of an object can be obtained in three ways: by displacing the sensor with respect to the image plane, by moving the lens, or by moving the object with respect to the object plane. Moving the lens or sensor with respect to one another causes the following problems: (a) The magnification of the system varies, causing the image coordinates of focused points on the object to change. (b) The area on the sensor over which light energy is distributed varies, causing a variation in image brightness. However, object movement is easily realized in industrial and medical applications. This approach ensures that the points of the object are focused perfectly focused onto the image plane with the same magnification. In other words, as the object moves, the magnification of imaging system can be assumed to be constant for image areas that are perfectly focused.

To automatically measure the sharpness of focus in an image, we must formulate a metric or criterion of "sharpness". The essential idea underlying practical measures of focus quality is to respond high-frequency content in the image, and ideally, should produce maximum response when the image area is perfectly focused. From the analysis of defocused image, it is shown that the defocusing is a low-pass filtering, and hence, focus measure should respond to high frequency variations of image intensity and produce maximum values when the image is perfectly focused. Therefore, most of the focus measure in the literature somehow maximizes the high frequency variations in the images. Generally, the objective has been to find an operator that behaves in a stable and robust manner over a variety of images, including those of in-door and outdoor scenes. Such an approach is essential while developing automatically focusing systems that have to deal with general scenes.

An interesting observation can be made regarding the application of focus measure operators. Equation (2) relates a defocused image using the blurring function. Assume that a focus measure operator is applied by convolution to the defocused image. The result is a new image expressed as:

$$r(x, y) = o(x, y) \otimes g(x, y) = o(x, y) \otimes (h(x, y) \otimes f(x, y)) \quad (6)$$

Since convolution is linear and shift-invariant, we can rewrite the above expression as:

$$r(x, y) = h(x, y) \otimes (o(x, y) \otimes f(x, y)) \quad (7)$$

Therefore, applying a focus measure operator to a defocused image is equivalent to defocusing a new image obtained by convolving the focused image with the operator. The operator only selects the frequencies (high frequencies) in the focused image that will be attenuated due to defocusing. Since, defocusing is a low-pass filtering process, its effects on the image are more pronounced and detectable if the image has strong

high-frequency content. An effective focus measure operator, therefore, must high-pass filter the image.

One technique for passing the high spatial frequencies is to determine its second derivative, such as Laplacian, given as:

$$\nabla^2 I = \frac{\partial^2 I}{\partial x^2} + \frac{\partial^2 I}{\partial y^2} \quad (8)$$

The Laplacian masks of 4-neighbourhoods and 8- neighborhoods are given in Fig. 2.

0	-1	0
-1	4	-1
0	-1	0

4-neighbourhoods

-1	-1	-1
-1	8	-1
-1	-1	-1

8-neighbourhoods

Fig. 2. Laplacian masks

Laplacian is computed for each pixel of the given image window and the criterion function can be stated as:

$$\sum_x \sum_y \nabla^2 I(x, y) \quad \text{for } \nabla^2 I(x, y) \geq T \quad (9)$$

Nayar noted that in the case of the Laplacian the second derivatives in the x and y directions can have opposite signs and tend to cancel each other. He, therefore, proposed the modified Laplacian (ML) as:

$$\nabla_M^2 I = \left| \frac{\partial^2 I}{\partial x^2} \right| + \left| \frac{\partial^2 I}{\partial y^2} \right| \quad (10)$$

The discrete approximation to the Laplacian is usually a 3 x 3 operator. In order to accommodate for possible variations in the size of texture elements, Nayar computed the partial derivatives by using a variable spacing (step) between the pixels used to compute the derivatives. He proposed the discrete approximation of the ML as:

$$\nabla_{ML}^2 I(x, y) = |2I(x, y) - I(x - \text{step}, y) - I(x + \text{step}, y)| + |2I(x, y) - I(x, y - \text{step}) - I(x, y + \text{step})| \quad (11)$$

Finally, the depth map or the focus measure at a point (x,y) was computed as the sum of ML values, in a small window around (x,y), that are greater than a threshold value t:

$$F(x, y) = \sum_{i=x-N}^{i=x+N} \sum_{j=y-N}^{j=y+N} \nabla_{ML}^2 I(i, j) \quad \text{for } \nabla_{ML}^2 I(i, j) \geq T_1 \quad (12)$$

The parameter N determines the window size used to compute the focus measure. Nayar referred the above focus measure as the sum-modified-Laplacian (SML) or traditional SFF (SFFTR).

3 Generalized Laplacian as Focus Measure

For a given camera, the optimally accurate focus measure may change from one object to the other depending on their focused images. Therefore, selecting the optimal focus measure from a given set involves computing all focus measures in the set. In applications where computation needs to be minimized by computing only one focus measure, it is recommended to use simple and accurate focus measure filter for all conditions [5]. Laplacian has some desirable properties such as simplicity, rotational symmetry, elimination of unnecessary in-formation, and retaining of necessary information. Modified Laplacian [2] takes the absolute values of the second derivatives in the Laplacian in order to avoid the cancellation of second derivatives in the horizontal and vertical directions that have opposite signs.

In this paper, we tried to use tuned Laplacian [5] as focus measure operator. A 3x3 Laplacian (a) should be rotationally symmetric, and (b) should not respond to any DC component in image brightness. The structure of the Laplacian by considering the above conditions is shown in Fig. 3. The last condition is satisfied if the sum of all elements of the operator equals zero:

$$a + 4b + 4c = 0 \tag{13}$$

Fig. 3. (a) The 3x3 Laplacian kernel (b) Tuned Laplacian kernel with c = 0.4, b = -1 (c) The Fourier Transform of (b) when c = 0 and (d) when c = 0.4

If $b = -1$, then $a = 4(1-c)$. Now we have only one variable c . The problem is now to find c such that the operator's response should have sharp peaks. The frequency response of Laplacian for $c = 0$ and for $c = 0.4$ are shown in Fig. 3 (c) and (d). From Fig 3 (d), we see that the response of the tuned focus measure operator ($c = 0.4$) has much sharper peaks than the Laplacian ($c = 0$).

The 4-neighbourhood kernel in Fig. 2 is obtained by $c = 0$, $b = -1$, and 8-neighbourhood kernel in Fig. 2 is obtained by $c = -1$, $b = -1$.

4 Simulation Results

We analyze and compare the results of 3D shape recovery from image sequences using the SFFTR with modified Laplacian and generalized Laplacian. Experiments were conducted on three different types of objects to show the performance of the new operator. The first object is a simulated cone whose images were generated using camera simulation software. A sequence of 97 images of the simulated cone was generated corresponding to 97 lens positions. The size of each image was 360×360 . The second object is a real cone whose images were taken using a CCD camera system. The real cone object was made of hard-board with black and white stripes drawn on the surface so that a dense texture of ring patterns is viewed in images. All image frames in the image sequences taken for experiments have 256 gray levels.

(a) At lens step 15

(b) At lens step 40

(c) At lens step 70

Fig. 4. Images of simulated cone at different lens steps

(a) At lens step 20

(b) At lens step 40

(c) At lens step 90

Fig. 5. Images of real cone at different lens steps

Figs. 4 and 5 show the image frames recorded at different lens position controlled by the motor. In each of these frames, only one part of the image is focused, whereas the other parts are blurred to varying degrees.

We apply Modified Laplacian and the Generalized Laplacian as fo-cus measure operator using SFFTR method on the simulated and real cone images. The improvements in the results (Fig. 6) on simulated cone are not very prominent except a slight sharpness in the peak. However, on real cone, we see in Fig. 7 (a) that there are some erroneous peaks using Modified Laplacian which are removed as shown in Fig. 7 (b) using generalized Laplacian.

Fig. 6. (a) 3D shape recovery of the Simulated cone using SFFTR with Modified Laplacian as Focus Measure Operator (b) with Tuned Laplacian as Focus Measure operator with $b = -0.8$, $c = 0.45$

Fig. 7. (a) 3D shape recovery of the Real cone using SFFTR with Modified Laplacian as Focus Measure Operator (b) with Tuned Laplacian as Focus Measure operator with $b = -1$, $c = 0.4$

5 Conclusions

In this paper, we have proposed a generalized Laplacian method as focus measure operator for shape from focus. Some improvements in the 3D shape recovery results are obtained. It is also noticed through simulation that erroneous peaks can be reduced

by using modified Laplacian, as discussed in the previous section. Further investigation is in process for generalized focus measure operator in-stead of fixed operators.

Acknowledgement

This research was supported by the second BK 21 program of the Korean Government.

References

1. Krotkov, E.: Focusing. *International Journal of Computer Vision* 1, 223–237 (1987)
2. Nayar, S.K., Nakagawa, Y.: Shape from focus. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 16(8) (August 1994)
3. Subbarao, M., Choi, T.-S.: Accurate recovery of three dimensional shape from im-age focus. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 17(3) (March 1995)
4. Nayar, S.K., Watanabe, M., Noguchi, M.: Real-time focus range sensor. In: *Proc. of Intl. Conf. on Computer Vision*, pp. 995–1001 (June 1995)
5. Subbarao, M., Tyan, J.K.: Selecting the Optimal Focus Measure for Autofocusing and Depth-from-Focus. *IEEE Trans. Pattern Analysis and Machine Intelligence* 20(8), 864–870 (1998)
6. Schlag, J.F., Sanderson, A.C., Neumann, C.P., Wimberly, F.C.: Implementation of Automatic Focusing Algorithms for a Computer Vision System with Camera Control. *Carnegie Mel-ion University, CMU-RI-TR-83-14* (August 1983)
7. Tenenbaum, J.M.: Accommodation in Computer Vision. Ph.D. dissertation, Stanford University (1970)
8. Hiura, S., Matsuyama, T.: Depth Measurement by the Multi-Focus Camera. In: *Proc. IEEE Int. Conf. Computer Vision and Pattern Recognition*, June 1998, pp. 953–959 (1998)
9. Jarvis, R.A.: A Perspective on Range Finding Techniques for Computer Vision. *IEEE Trans. Pattern Analysis and Machine Intelligence* 5(2) (March 1983)