Optics & Laser Technology 45 (2013) 21–31

Contents lists available at SciVerse ScienceDirect

Optics & Laser Technology

journal homepage: www.elsevier.com/locate/optlastec

Adaptive window selection for 3D shape recovery from image focus

Ikhyun Lee [a], Muhammad Tariq Mahmood [b], Tae-Sun Choi [a,*]

[a] School of Information and Mechatronics, Gwangju Institute of Science and Technology, 123 Cheomdan gwagiro, Buk-Gu, Gwangju 500-712, Republic of Korea
[b] School of Computer Science and Engineering, Korea University of Technology and Education, 1600 Chungjeolno, Byeogchunmyun, Cheonan, Chungnam 330-708, Republic of Korea

ARTICLE INFO

Article history:
Received 3 May 2012
Received in revised form
16 July 2012
Accepted 2 August 2012
Available online 21 August 2012

Keywords:
Shape from focus
Adaptive window selection
3D shape reconstruction

ABSTRACT

In shape from focus (SFF) techniques, focus measure plays an important role. Conventionally, two fixed windows at two stages have been used to compute and enhance the focus measures, respectively. The window sizes largely affect the accuracy of the recovered depth map. A smaller window is unable to compute focus quality in textureless and smooth areas in the images. Whereas, a large window may over smooth or distort the object shape. Moreover, the use of any fixed window may not provide optimal results. In this paper, instead of using a fixed window at each stage, an adaptive window is proposed to compute and enhance the focus measure. At first stage, to compute the focus measure, a criterion based on dispersion of image intensities determines the size of the mask. Similarly, at second stage, to enhance the focus measure, a criterion based on dispersion of initial focus values estimates the window size. The proposed scheme is tested using image sequences of simulated and real objects. The comparative analysis has demonstrated the effectiveness of the proposed adaptive windows.

1. Introduction

Shape form focus (SFF) is one of the passive optical methods that uses image focus as a cue for 3D shape estimation of an object. In first step, in this technique, the image sequence of an object is acquired by using a single camera with different focus levels. Multiple focus levels can be obtained by changing the focal length of the camera lens or by changing the distances between lens and the object as shown in Fig. 1(a). In first step, a focus measure, that computes focus quality for each pixel in the image sequence, is applied. The focus measure varies for each frame and gains maximum value at best focus plane. The depth for each point is computed by measuring the distance between initial position of the plane and the translated plane at maximum focus level as shown in Fig. 1(b). In other words, an initial depth map is obtained by maximizing the focus measure along the optical axis. In order to further enhance the 3D shape, initial results are refined using some approximation or machine learning techniques [1,2]. The performance of these depth estimation techniques, generally, relies on the accuracy of the initial focus volume and corresponding initial depth map thus, in SFF techniques, accurate focus measurements play important role [3,4].

In literature, many focus measures in spatial and frequency domains have been proposed [2,5–7]. In these methods, usually, the initial focus volume is computed in two stages using two fixed windows. At first stage, a focus measure is applied using a fixed mask usually of size 3×3. However, using fixed window, these focus measures are unable to compute focus value accurately. Particularly, textureless and smooth areas in the images have a tendency to the noisy focus measurements. At the second stage, to enhance the initial focus volume, the most common approach is to aggregate the focus values within a small window [4,6,8–10]. This summation is similar to averaging or linear filtering of the image focus volume. Malik and Choi [9] analyzed the effect of window size and concluded that the larger window produces blurring in the depth map (over-smoothing of object shape). They suggested the use of a smaller window for accurate depth map. Thelen [4] discussed the importance of the window size and suggested an algorithm for the second stage that chooses the effective window size from several neighborhood sizes based on a confidence criterion. However, it is difficult to select an appropriate window size. A smaller window is unable to suppress the noise completely while a larger window causes over smoothness of the object shape and more likely to remove the edges. Consequently, an erroneous depth map is obtained. Recently, Aydin and Akgul [3] suggested an adoptive weighted window that adjusts the weights using the information from all-in-focus image. All above mentioned techniques consider the second stage only, i.e., the enhancement of the focus volume. For the first stage, the initial focus measurements are computed using a small fixed mask.

In this paper, we address the window selection problem and suggest the use of adaptive windows for both stages to compute and enhance the focus measurements, respectively. At each stage, the window size is selected dynamically through an iterative

* Corresponding author. Tel.: +82 62 715 2392.
E-mail addresses: ihlee@gist.ac.kr (I. Lee), tariq@kut.ac.kr (M. Tariq Mahmood), tschoi@gist.ac.kr (T.-S. Choi).

0030-3992/$ - see front matter © 2012 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.optlastec.2012.08.003

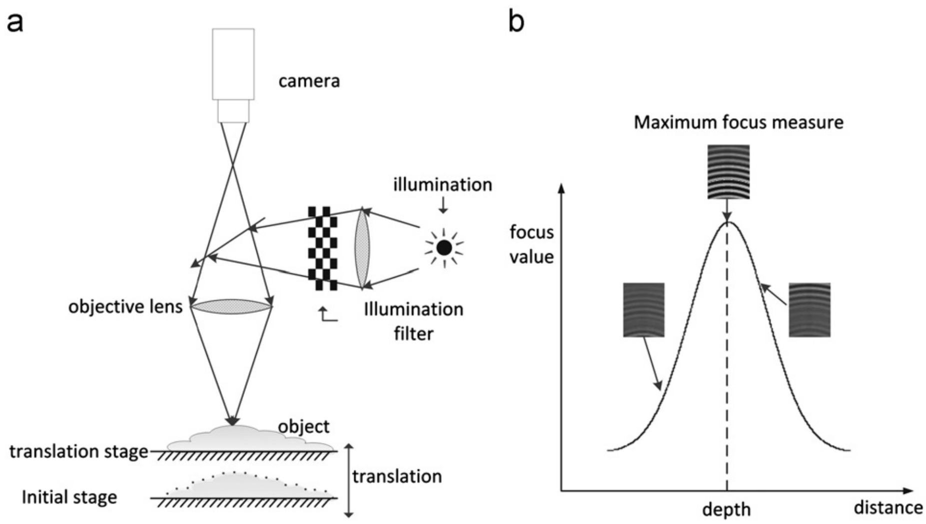

Fig. 1. Shape from focus and depth estimation.

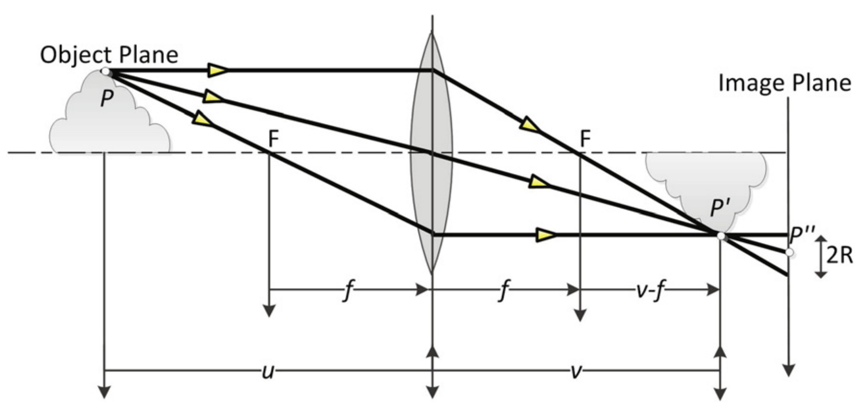

Fig. 2. Image formation.

process using a criterion. At first stage, a criterion based on dispersion of image intensities determines the size of window. Smaller disparity indicates the lower texture while the higher disparity means high variation in image intensities. At second stage, dispersion in focus values is used to obtain an appropriate window size. The effectiveness of the proposed scheme is validated using image sequences of synthetic and real objects.

In the remainder of the paper, Section 2 discusses the SFF in detail and its related work. The proposed adaptive window selection for both stages is explained in Section 3. Section 4 presents the experimental results and comparative analysis. Finally, Section 5 concludes this study.

2. Related work

2.1. Basic image formation

Fig. 2 shows the basic geometry of image formation of focused and defocused objects through the convex lens. An object of unknown depth is translated in the optical direction in fixed finite steps with respect to a real aperture camera. A point p on the surface of the object becomes focused gradually and at one stage, it will be in sharp focus and its focused image is obtained at p' on the image plane. This sharp focus stage provides information about the depth of this point. On the other hand, a blurred image is obtained at p'' of a defocused object point. By considering the circular aperture, the blurred image is also a circle of diameter $2R$. Due to effects of diffraction, polychromatic illumination, lens aberrations, etc. it is a circular blob with the brightness falling of gradually at the border. Therefore, this blurring effect is usually modeled by 2D Gaussian function also known as point spread function (PSF). So, a sensed image is the convolution of the actual image $I(x,y)$ and a Gaussian PSF. At every step, an image is captured and a stack of visual observations $I_Z(x,y)$, consisting of Z images, each of size $X \times Y$, is obtained. Due to the limited depth-of-field of the camera lens and the 3D nature of the object, the captured images are space-variantly blurred such that some parts of the object come into focus in each frame. We are interested in determining the particular distances of all object points from the camera for which they are well focused on the image plane.

2.2. Focus measure

Focus measure is applied for each pixel in every frame to measure the focus quality and plays a very important role in 3D shape recovery as it is a crucial step in the calculation of the depth map. One type of famous measure methods is based on derivatives. The Tenengrad [11] focus measure (TEN) is a gradient magnitude maximization method that measures the sum of the squared responses of the horizontal and vertical Sobel masks. Another derivative based medhod is using Laplacian. For applying