A Focus-Measurement Based 3D Surface Reconstruction System for Dimensional Metrology

S.M.Vignesh\textsuperscript{1*}, R. Senthilnathan\textsuperscript{1}

\textsuperscript{1}Department of Mechatronics Engineering, SRM Institute of Science and Technology, SRM Nagar, Kattankulathur, 603203, Kanchipuram, Chennai TN, India.

E-mail: vigneshs1@srmist.edu.in

Abstract. Image focus is an important parameter which depends on various other parameters of an imaging system. One among them is the distance of the object plane from the lens. This parameter is referred as the depth in computer vision literatures. Since depth affects the focus level of pixels in an image, focus measurement can be used to interpret the depth of points of an object. The paper presents the details of the novel indigenously developed automated measurement system which can perform high-precision movements based on feedback control. The measurement system is inspired by classical systems used in techniques such as shape from focus. This system fundamentally being a mechatronics system integrates a high-precision multi-stage translation mechanism which is actuated with a servo motor with sensors such as camera, linear encoder, optical limit switches. The hardware interface is developed for automated measurement through a USB data acquisition system. The paper presents the methodology of surface reconstruction and detailed description of the various challenges encountered in the integration. The surface reconstruction is performed for various millimetres and sub millimetres level engineered specimens. The paper presents the future direction of research and also various potential application possibilities for such measurement setup.

Keywords: metrology, surface reconstruction, shape from focus, depth from defocus.

1. Introduction

In modern manufacturing industries, non-contact dimensional measurement techniques plays a vital role. The use of non-contact measurement techniques preserves the surface quality of the part and stability of the measurements which in turn improves the final value of the products.

In many industrial applications, machine vision systems are becoming increasingly relevant, such as automatic inspection, robotic manipulations etc. All these activities are carried out in a 3-dimensional environment, which in turn requires the reliable information of the 3D depth of the scene. For some cases sparse 3D depth information is all require to understand the scene, whereas for other cases it requires dense 3D depth map is required. Sometimes the accurate shape of the surface of an object in the scene may be required instead of estimating the surface's mere range value. As a result, the computer vision and robotics community has developed a great deal of research interest in the acquisition of depth information. Currently available vision based range finder can be classified into
two broad categories, active and passive techniques. In active technique special lighting devices illuminates the scene, whereas in passive technique the ambient light provides lighting for the scene.

1.1. Shape from focus
Shape from focus (SFF) technique is one of the passive method for depth estimation, which requires stack of images from an object captured at different focal setting of the imaging acquisition device. As SFF is the passive approach, there is no physical contact between the surface of the target and the imaging system. In addition to this fact, SFF uses single camera, this over-comes the correspondence problem usually occurs in the stereo vision imaging device. The local variation of the focus in the image is the cue for depth.

According to the thin lens camera model the focus in the camera is adjusted by changing the distance, \( v \), between the lens plane and the sensor plane. During image formation only points a particular distance will be in focus this depends on the focal setting of the lens. The relationship between focal length, distance from the image and distance from the target is given by the thin lens equation:

\[
\frac{1}{f} = \frac{1}{u} + \frac{1}{v}
\]

Since each point in the image plane has the corresponding point in the object plane, the maximum focus will only happen at a particular distance.

![Figure 1. Formation of image.](image)

The degree of focus is the function of lens setting and the depth of the scene. Hence the focus of an image is cue to estimate the depth with the known lens setting. Shape from focus (SFF) technique is used to interpret the depth of a 3-dimensional object. For the SFF technique a sequences or stack of images are required which are taken at different focal distance from a reference plane. These stack of images are captured by fixing the optical elements fixed and by translating the object at different stages. And the translation of different stages are recorded as the focal distances for the stack of images. Then the focus measure operator is applied for all the pixels along the stack of images. The focus measure values obtained at each pixel along the focal stack are interpolated by the Gaussian curve to fit the peak value of the focus measure. After obtaining the peak values of each pixel the depth of each pixel is interpolated by Gaussian 3 point approximation.

1.2. Shape from defocus
When an object point is not in focus, a blurred circle would be formed in the image whose radius is defined by a blur parameter \( \sigma \) is the image of the point, given by.

\[
\sigma = \rho r v \left( \frac{1}{f} - \frac{1}{v} - \frac{1}{D} \right)
\]

Where \( f \) is focal length, \( D \) is the distance between the point of the object and the lens, \( v \) is distance between the lens and image detector, \( r \) is the aperture radius of the lens, and \( \rho \) is a camera constant [6].
Shape from defocus (SFD) is another technique to interpret the depth of a 3-dimensional object. SFD requires two defocused images of an object captured at different lens settings. The relation between two defocused images is given by:

$$\sigma_1 = \alpha \sigma_2 + \beta$$

Where $\alpha$ and $\beta$ are known parameters depending in camera settings. The relative blur between two defocus pictures is evaluated by assuming a suitable optical transfer function model. Over the local region depth is assumed to be constant [6]. It is possible to calculate the depth of the object corresponding to the local area by understanding the lens setting and the blur parameters.

In [1], Peter Kinnel et al. used a 3D metrology vision system to automatically detect the location of the target object. Victor Alonsoa et al. in [2], expressed the application of machine vision system for metrology in industry 4.0.

This paper focuses on development of an automated image capturing system with a translation mechanism required to capture stack of image from an object at various focal distances. The linear translation mechanism is coupled with a linear potentiometer for collecting the feed-back of the distance travelled by the translation stage. The hardware system is connected to the computing system through the USB data acquisition system.

2. Experimental Setup

The experimental set-up consists of linear translation table attached with high-precision multi-stage translation mechanism which is actuated with a servo motor and optical limit switches. And to measure the translation distance a linear potentiometer is attached with the linear translation table. For image acquisition USB camera attached with tele-centric lens is used. The hardware setup is interface with the computer through USB data acquisition system.

![Image Acquisition System](image.jpg)

**Figure 2.** Image Acquisition System.
2.1. Specifications
All images are captured by Contras tech VT-EX500CS USB 2.0 camera with the 2592 X 1944 image resolution and 5fps frame rate. The object is focused through a TC23036 bi-tele centric lens 0.243x magnification, working distance of 102.5 mm, and field depth of 11 mm. The distance travelled by the translation table is measured by the linear potentiometer with useful electrical stroke 100 mm, repeatability 0.01 mm, and displacement speed 3 m/s. Reversing the direction of the table movement is achieved by the optical limit switches placed at the extreme ends of the translation mechanism. The various inputs and outputs of motor for the translation table, linear potentiometer, and optical limit switch are connected to the system with the help of NI USB-6009, 8 inputs, 14-bit data acquisition device.

3. Motion Control and Image Acquisition

3.1. Characteristics of the translation
The rectilinear motion of the translation table is achieved by the worm wheel and belt drive connected with the servo motor. The characteristic of this rectilinear motion is analyzed for the better understanding of the system. In order to analyze, the translation stage is allowed to translate multiple times and the corresponding voltage change in the linear potentiometer which is coupled with the stage is recorded through the data acquisition device. The graphical representation of the voltage change is shown in the following figure. 3 (a) and (b). From the graph it is observed that during the start of the table motion both from bottom as well as top there is no significant change in voltage in the potentiometer. This happens because of the initial friction force which resisted the motion of linear potentiometer coupled with the translation table.

![Figure 3. (a) Bottom to Top Motion (b) Top to Bottom Motion.](image)

3.2. Correction strategy
If the selection of object height is lesser than the maximum sweep of the translation table. Then along the region where the motion of translation table is restricted due to friction force may not be considered for image acquisition. Moreover all the points of image acquired in this region will not be in focus.

3.3. Feedback control
The remaining linear region of the graph shows the repeatability of the linear potentiometer. From this it is observed that voltage corresponding to the distance is not consistent for the repeated cycles. So the various distance of the translation table has to be controlled by the different voltages for each trail.
3.4. Image Acquisition

Image acquisition of the object takes place at various translational stages. The object is placed over the translation stage and allowed to move along z-axis towards the lens. The distance \( d \), between the fixed reference plane and the translation stage is noted. Depending on the optical parameter, working distance, the focused plane will be fixed at a specific distance from the lens. The distance between the focused plane and the reference plane is also fixed. As the object reaches the focused plane, all points lying on it will be perfectly focused while other points will be blurred. Images of the object will be captured at regular interval of distances.

\[ I_d (x, y) = h (x, y) * I_f (x, y) \]

\[ r (x, y) = o (x, y) * (h (x, y) * I_f (x, y)) \]

This can be rewritten as:

\[ r (x, y) = h (x, y) * (o (x, y) * I_f (x, y)) \]

Since convolution is linear and shift-invariant. From the above expression it is obvious, 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 focus operator should be a high pass filter which

3.5. Image Stack Preparation

As the object moves towards the focused plane, images are captured at different translation stage. For each captured image corresponding distance from the reference plane is noted. This distance value helps to stack the image in sequence. Each point in the image will be perfectly focused on any one of the image along image stack. The point which is at the best focus along the image stack has to be estimated. The estimated value is known as the focus measure of the point in an image.

3.6. Focus Measurement

In order to estimate the focus measure, a focus operator is applied for each image along the stack of images. The focus measure operator \( o (x, y) \) is applied to the defocused image \( I_d (x, y) \), the result will be a new image \( r (x, y) \) that may be expressed as:

\[ I_d (x, y) = h (x, y) * I_f (x, y) \]

\[ r (x, y) = o (x, y) * (h (x, y) * I_f (x, y)) \]

This can be rewritten as:

\[ r (x, y) = h (x, y) * (o (x, y) * I_f (x, y)) \]

Since convolution is linear and shift-invariant. From the above expression it is obvious, 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 focus operator should be a high pass filter which
highlights the high frequency variation along the image. Several operators are proposed for this focus measurement. Among which the most widely used operators are the Modified Laplacian, Gray-Level variance, Modified Gray Level Variance, etc. In this work Modified Gray Level Variance is used for focus measurement.

3.7. Gaussian Interpolation and Depth Estimation

Once the focus measure of all the pixels using any one of the focus operator is determined, next is to estimate the depth of each point of the object. Gaussian interpolation is the widely used method for this task. For each pixel the image with highest focus measure is evaluated and depth corresponding to the pixel is interpolated using Gaussian function around the position. In the figure, different focus measure of the pixel at a coordinate (i, j) are shown along the image stack. For the Gaussian 3 point interpolation maximum focus measure value $F_m$, previous value $F_{m-1}$ and the next value $F_{m+1}$ are used. The depth of the pixel is assumed to be the peak value of the Gaussian.

![Figure 5. Gaussian 3 Point Interpolation.](image)

3.8. Filtering

After obtaining the depth map using shape from focus it would be noticed that some regions will be inaccurate and corrupted by noise. To order to filter this several smoothening technique are available like median filter, bilateral filter, etc.

3.9. Confidence level estimate

Depending on the optical configuration and image content the quality of the depth map obtained will be affected. The response of focus measure operation on the image stack is also based in the two effects. These factors determines the accuracy of the depth map obtained. So confidence level estimate has to detect the variation due to optical configuration and imaging condition. In addition to that it should be able to detect the focal variation when the focal operator fails to detect.

4. Results and Discussion

The image stack of the 3D object is acquired from the image acquisition device as mentioned in section 2. And the stack of images are sequenced according to the focal configuration of the optics in the image acquisition system as in figure. Then the focal measure operator is applied to the stack of images. In this work Modified Gray Level Variance focus operator is used for the focus measurement. Then by using the Gaussian 3 point interpolation the peak values of the focus corresponding are estimated. Figure 7. Shows the depth plot of the 3 object using shape from focus technique.
5. Conclusions
The depth map plot can be used in various industrial applications surface measurement, surface reconstruction, reverse engineering etc. By improving the calibration of the image acquisition system the accuracy of the depth map with close tolerance can be achieved. In this work the 3D object with good surface texture was chosen for the experiment to show the better visualisation of surface variation. Whereas the experiment can be extended for the smooth surfaces. In case of smooth surface the focus measure operator need to be modified accordingly to capture the focal variation of the image stack. The same technique can be extended for various environment by designing an image acquisition according to the requirement. Accordingly the suitable focus measure operator need to be developed for capturing the focal difference. In case for industrial application the environment for the measuring device will be fixed, this makes the measurement system for robust application for different surface characteristics.
References

[1] Peter Kinnell, Tom Rymer, John Hodgson, Laura Justham, Mike Jackson, Autonomous metrology for robot mounted 3D vision systems.

[2] Victor Alonsoa, Angel Dacal-Nietoa, Luis Barretob,c, António IdEduardo Riveroa, Industry 4.0 implications in machine vision metrology: an overview.

[3] S.K. Nayar and Y. Nakagawa. Shape from Focus. IEEE Transactions on Pattern Analysis and Machine Intelligence 16(8):824-831, 1994.

[4] S. Pertuz, D. Puig, M. A. Garcia. Reliability measure for shape-from focus. Image and Vision Computing 31:725–734, 2013.

[5] S. Pertuz, D. Puig, M. A. Garcia. Analysis of focus measure operators for shape-from-focus. Pattern Recognition. 46(5): 1415-1432, 2013.

[6] Subhasis Chaudhuri, A. N. Rajagopalan. "Depth From Defocus: A Real Aperture Imaging Approach", Springer Nature, 1999.

[7] RR Sahay, AN Rajagopalan. Dealing with parallax in shape-from-focus. IEEE Transactions on Image Processing 20 (2), 558-569.

[8] Aamir Saeed Malik, Tae-Sun Choi. Anovel algorithm for estimation of depth map using image focus for 3D shape recovery in the presence of noise. Pattern Recognition 41 (2008) 2200 – 2225.

[9] Murali Subbarao, Tae Choi, Arman Nikzad. Focusing techniques. Optical Engineering 32(11), 2824—2836.

[10] Usman Ali, Muhammad Tariq Mahmood. 3D Shape Recovery by Aggregating 3DWavelet Transform-Based Image Focus Volumes Through 3DWeighted Least Squares. Journal of Mathematical Imaging and Vision 10851-019-00918-8.

[11] Zhiqiang Ma, Dongjoon Kim, Yeong-Gil Shin. Shape-from-focus reconstruction using nonlocal matting Laplacian prior followed by MRF-based refinement. Pattern Recognition 103 (2020) 107302.

[12] Usman Ali, Ik Hyun Lee, Muhammad Tariq Mahmood. Guided image filtering in shape-from-focus: A comparative analysis. Pattern Recognition 111 (2021) 107670.

[13] Sung-An Lee and Byung-Geun Lee. Accurate 3D Surface Reconstruction for Smart Farming Application with an Inexpensive Shape from Focus System. Journal of Sensors Volume 2020, Article ID 1859512, 7 pages.

[14] Murali Subbarao, Gopal Surya. Depth from Defocus: A Spatial Domain Approach. International Journal of Computer Vision, 13, 3, 271-294.