DEPTH MAP ESTIMATION USING A ROBUST FOCUS MEASURE

Aamir Saeed Malik, Seong-O Shim, Tae-Sun Choi (Senior Member IEEE)

Department of Mechatronics, Gwangju Institute of Science and Technology,
1 Oryong-Dong, Buk-Gu, Gwangju, 500-712, Korea
(e-mail: aamir@gist.ac.kr, seongo@gist.ac.kr, tschoi@gist.ac.kr)

ABSTRACT

Accurate estimation of depth map leads to precise three-dimensional shape recovery. In this paper, we present a new focus measure for calculation of depth map. This new focus measure is based on an optical transfer function implemented in the frequency domain and it has shown robustness in the presence of noise as compared to the earlier focus measures. The results of the proposed focus measure have shown considerable improvement in the presence of noise with respect to other focus measures.

***Index Terms*—**Focus measure, noise, robust

1. INTRODUCTION

The objective of shape from focus is to find out the depth of every point of the object from the camera lens. The depth of every point is calculated by finding the best focused points, i.e., sharpest pixel values. Shape from focus (SFF) is one of the image processing techniques that uses depth map to recover 3D shape of the object.

In SFF, a sequence of images that correspond to different levels of object focus is obtained. A sharp image and the relative depth can be retrieved by collecting the best focused points in each image. The absolute depth of object surface patches can be calculated from the focal length and the position of lens that gave the sharpest image of the surface patches. The depth or best focus is obtained by using some focus measure.

So every SFF scheme relies on a Focus Measure operator and an approximation technique. Focus Measure operator plays a very important role for three dimensional shape recovery because it is the first step in calculation of the depth map. So a focus measure operator should provide a very good estimate of the depth map by showing robustness even in the presence of noise.

One factor to be kept in mind is that we have finite number of images in the image sequence. The information obtained from them does not represent actual object specification especially in the case of geometrically complex objects. The only way for obtaining accurate results from SFF techniques is estimating object specifications in the gap between images in the image sequence. Hence, the role of first depth map obtained by using some specific focus measure is increased manifolds.

This work was supported by the basic research project through a grant provided by the Gwangju Institute of Science and Technology in 2007.

2. RELATED WORK

The related work in SFF can easily be divided into two sections, namely, the focus measure operators and the approximation methods. Brief description of these sections is given below.

2.1. Focus Measure Operators

A Focus Measure operator is one that calculates the best focused point in the image. The focus measure calculates the best focused point by evaluating the sharpness of a pixel locally. Franz Stephan Helmli and Stefan Scherer [1] summarized the traditional focus measures while introducing three new focus measure operators.

The most commonly used focus measure is the Laplacian operator. Laplacian is obtained by adding second derivatives in the x and y directions. Modified Laplacian (ML) [2] is computed by adding squared 2nd derivates. In order to handle possible variations, Shree K. Nayar and Yasuo Nakagawa suggested a variable spacing (step) between the pixels used to compute derivatives. Since Laplacian yields bad results for weak textured images therefore to improve robustness for weak-texture images, Shree K. Nayar and Yasuo Nakagawa [2] presented focus measure at (x,y) as Sum of ML (SML) values in a local window (about 5x5).

Tenenbaum Focus Measure is gradient magnitude maximization method that measures the sum of squared responses of horizontal and vertical Sobel masks. Variance Focus Measure is based on the variance of gray-level which is higher than that in a blur image. Mean Method Focus Measure [1] depends on the ratio of mean grey value to the center grey value in the neighborhood. Curvature Focus Measure [1] exploits that the curvature in a sharp image is expected to be higher than that in a blur image. The surface is approximated using a quadratic equation and coefficients are calculated using a least squares approximation technique. Point Focus Measure [1] is approximated by a polynomial of degree four and its extremum is estimated.

2.2. Approximation Techniques

A more accurate depth range image can be obtained using some approximation reconstruction scheme on the results acquired from some focus measure operator. Joungil Yun and Tae-Sun Choi [3] summarized various approximation techniques. Some of them include Traditional SFF, Focused Image Surface etc.

Many approximation techniques have been reported in the literature lately. Generally, those approximation methods are based on machine learning tools, especially fuzzy logic, neural network and dynamic programming.

3. METHOD

This paper introduces a new focus measure for the estimation of depth map. This is done by finding out the best frame number for each pixel.

Since the focus measure calculates the sharpest pixels in the image hence their success depends on their ability to calculate the sharpness value of each pixel. Therefore, algorithms and techniques based on calculating sharpness and edges in an image automatically become potential candidates for the selection of focus measure.

We introduce a new focus measure based on bipolar incoherent image processing and we call it Optical Focus Measure and denote it as FM_O. Ting-Chung Poon and Partha P. Banerjee [4] have discussed bipolar incoherent image processing in detail. Generally, there are severe limitations of incoherent processing with standard incoherent systems in that the Optical Transfer Function achievable is the autocorrelation of the pupil function. Or equivalently the Point Spread Function is real non-negative. Among the acousto-optic heterodyning image processing, a number of novel techniques have been devised to implement bipolar point spread functions in incoherent systems. These techniques are usually referred to as bipolar incoherent image processing in the literature.

The sharpness of pixel values in the image is found by convolving the spectrum of the intensity image with the transfer function which in our case is Optical Transfer Function (OTF). The computed image $[i_c(x, y)]$ is given as:

$$i_c(x, y) = \mathrm{Re}\,[|\Gamma_0(x, y)|^2 * h_\Omega(x, y)\}] \qquad (1)$$

where '*' indicates convolution and:

$$|\Gamma_0(x, y)|^2 = \text{Spectrum of the Intensity Image}$$
$$h_\Omega(x, y) = \text{Transfer Function}$$

Transfer function is basically the OTF which is calculated in frequency domain using either Fourier or Cosine transform. Transfer function $h_\Omega(x, y)$ is given as:

$$h_\Omega(x, y) = F^{-1}\{OTF_\Omega(k_x, k_y)\}$$

where:

$$OTF_\Omega(k_x, k_y) = \text{Optical Transfer Function}$$
$$k_x, k_y = \text{Spatial frequencies}$$

So finally we can write the computed image as:

$$i_c(x, y) = \mathrm{Re}\,[F^{-1}\{F\{|\Gamma_0(x, y)|^2\}\, OTF_\Omega(k_x, k_y)\}] \qquad (2)$$

where F is for Fourier Transform and F^{-1} is for Inverse Fourier Transform. The OTF itself is calculated as:

$$OTF_\Omega(k_x,k_y)=\iint p_1(x',y')p_2(x'+fk_x/k_0,y'+fk_y/k_0)dx'dy' \qquad (3)$$

where f is the focal length of the lenses and k_0 is the wave number of light. The OTF is the cross correlation of the two pupils (p_1 and p_2) in the incoherent optical system [4]. Hence, the point spread function becomes bipolar.

In equation (3) above, p_1 is a difference of Gaussian aperture function and p_2 is a small pin hole aperture. p_1 is given as [4]:

$$p_1 = \exp[-a_1(x^2 + y^2)]-\exp[-a_2(x^2 + y^2)]$$

where a_1 and a_2 are constants. p_2 is given as [4]:

$$p_2 = \delta(x,y)$$

For implementation purposes, equation (3) can be rewritten as:

$$OTF_\Omega(k_x,k_y)=\exp[-\sigma_1(k_x^2+k_y^2)]-\exp[-\sigma_2(k_x^2+k_y^2)] \qquad (4)$$

where:

$$\sigma_1 = a_1\,(f/k_0)^2$$
$$\sigma_2 = a_2\,(f/k_0)^2$$

The above equation shows that OTF is basically a band-pass filter with gradual cut-off frequency. Therefore, this filtering provides sharpness at pixel points in an image. The filtering operation depends upon σ_1 and σ_2. Sharp focus measure is obtained by adjusting these two parameters. The high frequency component of an image area is determined by processing in the Fourier domain and analyzing the frequency distribution. Fourier transform used to be computationally expensive but with high speed personal computers available today, this complexity has decreased exponentially and it is not a matter of concern anymore. The processing in the frequency domain is particularly useful for noise reduction as the noise frequencies are easily filtered out. Fig 1 shows the filter with $\sigma_1 = 0.01$ and $\sigma_2 = 0.1$ and the corresponding Fourier spectrum of the "Test Image" (Test image shown in fig 2).

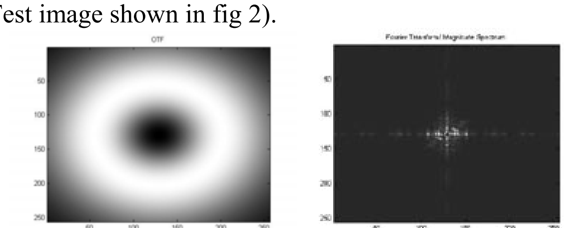


Fig 1: Filter design & Fourier spectrum of the Image

The next step is to select the best focused point in the sequence of images. Equation 4 is used to compute the focus measure at a point (i,j) in a small window around (i,j) and the value at (i,j) is replaced by the sum of computed values of all pixels in that window. This operation is similar to that used for Sum of Modified Laplacian [2]. We have used optimum window size [5] for our experiments.

Therefore, the final focus measure is given as FM_O which stands for Optical Focus Measure and is calculated as:

$$FM_O\,(i,j) = \sum_{x=i-N}^{i+N} \sum_{y=j-N}^{j+N} i_c\,(x,y)$$

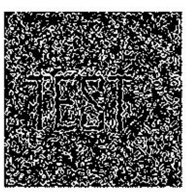
4. RESULTS & DISCUSSION

The proposed focus measure is tested with various types of images including a "Test" image, a sequence of 97 simulated cone images and a sequence of 97 real cone images. The resolution of the simulated and real cone images is 360x360 pixels. The real cone is taken from the CCD camera system. The real cone object was made of hardboard with black and white stripes drawn on the surface. The results are compared with 3 other operators, i.e., SML, Tenenbaum and Gray Level Variance (GLV).

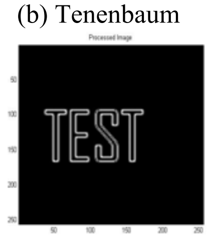
Fig 2(a) shows the "Test" image with Gaussian noise with zero mean and variance of 0.5 added to the image. We can see from Fig 2(b), (c) and (d) that results for both the SML and Tenenbaum have deteriorated while the proposed optical focus measure (FM_O) still shows very good result.

(a) Test Image

(b) Tenenbaum



(c) SML



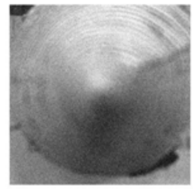
(d) FM_O

Fig 2: Applying focus measure operators on test image

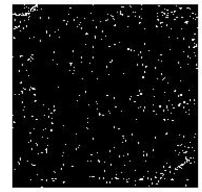
Fig 3 shows real cone image with Gaussian noise (mean=0, variance=0.005) added and the corresponding processed images with Tenenbaum, SML and the FM_O.

Hence, as clear from the figures, the performance of Tenenbaum and SML degrades when noise is added to the images. However, FM_O performs satisfactorily well. In real time applications, various type of noise like Rayleigh, exponential, uniform, shot, speckle, Gaussian etc may occur. Therefore, a robust focus measure is required to deal with noisy situations.

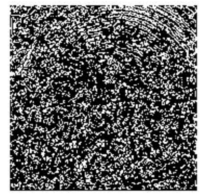
As for depth map calculation, consider sequence of 97 simulated cone images. Fig 4(a) & (b) show two of the frames for the simulated cone. Fig 4(c) & (d) show the depth map using SML & FM_O without noise addition and Fig 4(e) & (f) show the depth map using SML & FM_O with Gaussian noise added.



(a) Gaussian noise added



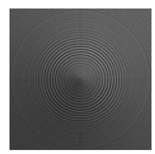
(b) Tenenbaum



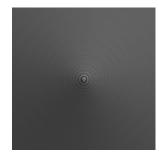
(c) SML

(d) FM_O

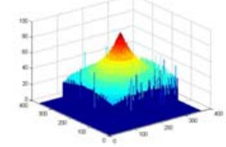
Fig 3: Applying various focus measures on real cone image



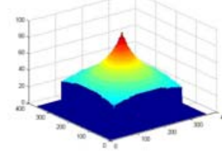
(a) Frame 50



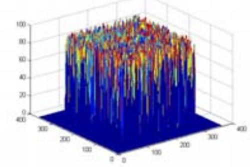
(b) Frame 90



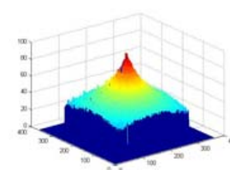
(c) SML



(d) FM_O



(e) SML

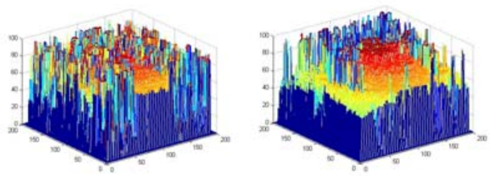


(f) FM_O

Fig 4: Depth maps for the simulated cone object

It can be seen from the figures that the 3D depth map obtained using FM_O is much smoother as compared to that of SML when there is no noise added to the images. Also, in the presence of Gaussian noise, the depth map obtained using FM_O is very clear but that of SML has degraded significantly. The degradation of SML is due to the fact that the noise has enhanced individual pixel values resulting in spikes in the depth map. On the other hand, the result for FM_O is excellent in comparison to SML.

Now consider the sequence of 97 images of real cone. Fig 5(a) & (b) show the depth maps for real cone images with Gaussian noise added. As can be seen from the figures, SML results deteriorate significantly while FM_O result is degraded but still recognizable. These results can be processed further by applying median or wiener filter etc. However, this type of post-processing does not improve the result because the images are already processed with the noise and the noise values are now the inherent property of each of the pixel values.



(a) SML (b) FM_O

Fig 5: Depth maps for the real cone object

Another solution to tackle the noise is pre-processing the images with some type of filtering. We used Wiener filter. It filters an intensity image that has been degraded by constant power additive noise. Since we already know that, in our case, this additive noise is Gaussian noise, therefore, we use this information for implementing this filter. However, we found that there is little improvement in the results of focus measures after the usage of Wiener filter. Still, the best performance is shown by FM_O.

We used Mean Square Error (MSE), Root Mean Square Error (RMSE) and correlation to compare the results of SML, GLV and the proposed FM_O. The MSE is a distortion metric while RMSE is simply the square-root of MSE. Correlation provides a measure of similarity of two images. For better graphic representation, we have used the normalized values for MSE and RMSE.

We found that the MSE and RMSE values are lowest for the proposed focus measure. Also, we found that the correlation coefficient of FM_O is highest among all the three focus measures. Table 1 shows the comparison results for the above mentioned metric measures for the various focus measures in the presence of Gaussian noise. However, we found that the results of SML and Tenenbaum are very similar in nature. So, we ignore Tenenbaum and we show the results for the three focus measures, namely, SML, GLV and FM_O. Figure 6 shows the corresponding visual representation of table 1.

	MSE(Norm)	RMSE(Norm)	Correlation
SML	0.5244	0.7209	0.5901
GLV	0.3900	0.6157	0.6379
FMo	0.2883	0.4739	0.7633

Table 1: Metric measures for Gaussian Noise

It can be seen from the above table that the lowest correlation value is for SML followed by GLV and then FM_O. So in terms of correlation, best results are shown by FM_O followed by GLV and then SML. Similarly, for MSE and RMSE, the lower values mean the result is better. Again, it can be observed from table 1 that the lowest MSE and RMSE values are depicted by the proposed focus measure FM_O followed by GLV and then SML.

Another important factor is the selection of Parameter values for FM_O. The two sigma parameters (eq. 4) define the frequency that is effectively blocked. The selection of these two parameters directly depends on the frequency characteristics of the type of noise. Results shown in this paper use $\sigma_1 = 0.01$ & $\sigma_2 = 0.1$.

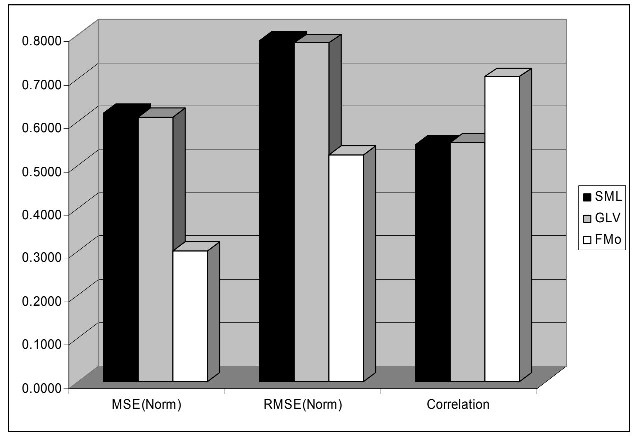


Fig 6: Comparison of Focus Measures

5. CONCLUSION

A new focus measure is introduced in this paper. This focus measure is based on an optical transfer function which has band-pass characteristics. We tested and compared this focus measure using simulated cone images and real cone images [6]. The results show that this new focus measure tends to perform better than the traditional focus measures (SML, Tenenbaum, GLV) in the presence of noise.

6. REFERENCES

[1] Franz Stephan Helmli and Stefan Scherer, "Adaptive Shape from Focus with an Error Estimation in Light Microscopy", *2nd Int'l Symposium on Image and Signal Processing and Analysis (ISPA01),* June 2001.

[2] Shree K. Nayar and Yasuo Nakagawa, "Shape from focus", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 16, No. 8, pp 824-831, August 1994.

[3] Joungil Yun and Tae-Sun Choi, "Accurate 3-D Shape Recovery using Curved Window Focus Measure", *IEEE ICIP*, vol. 3, pp 910-914, Oct 1999.

[4] Ting-Chung Poon, Partha P. Banerjee., *Contemporary optical image processing*, 1st ed., Elsevier Science Ltd., New York, 2001.

[5] Aamir S. Malik and Tae-Sun Choi, "Consideration of illumination effects and optimization of window size for accurate calculation of depth map for 3D shape recovery", Pattern Recognition, Vol. 40, No. 1, pp 154-170, Jan 2007.

[6] Murali Subbarao, and Tae-Sun Choi, "Accurate recovery of three dimensional shape from image focus", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 17, No. 3, pp 266-274, March 1995.