

REAL METROLOGY BY USING DEPTH MAP INFORMATION

E. Ardizzone^a, S.Battiato^b, A.Capra^c, S.Curti^c,

^aUniversity of Palermo - DINFO, Viale delle scienze, 90128 Palermo, Italy

^bUniversity of Catania, Dipartimento di Matematica ed Informatica, 95121 Catania, Italy

^cSTMicronics - AST Catania Lab, Stradale Primosole 50, 95121 Catania, Italy

ABSTRACT

Usually in an image no real information about the scene's depth (in terms of absolute distance) is available. In this paper, a method that extracts real depth measures is developed. This approach starts considering a region located in the center of the depth map. This region can be positioned, interactively, in any part of the depth map in order to measure the real distance of every object inside the scene. The histogram local maxima of this region are determined. Among these values the biggest, that represents the gray-level of the most considerable object, is chosen. This gray-level is used in an exponential mapping function that converts, using the input camera settings, the depth map gray-levels into real measures. Experiments over a large dataset of images show good performances in terms of accuracy and reliability.

Keywords: Depth map, Depth From Defocus, Metrology, Mapping Function.

INTRODUCTION

To recover the depth information of a general scene, starting from an image acquired with a digital camera, several approaches has been proposed^{1 2 3 4 5 6 7 8 9 10}. Depth from Defocus (DFD) provides a depth map by using two different images of the same scene captured with different camera settings. Subbarao^{1 2} uses this technique to evaluate the distances among the objects of the acquired scene obtaining a camera system with a speedy auto-focus. Darrell³ evaluates the depth information from the sharpness analysis of the scene varying the focus position. Simon⁴ evaluates the local depth information from both in-focus and out-of-focus edges information. Schneider⁵ compares the out-of-focus degree of two images of the same scene acquired with different apertures. Ziou⁶ uses a technique based on the local decomposition of the image using the Hermite transform. Rayala⁸ approaches the DFD as system identification: the out-of-focus process is modeled by a linear system using an appropriate transfer function. Watanabe-Nayar⁹ proposes a class of operators, textures invariant, that produce a dense and accurate depth map.

Using these described DFD methods It's possible to obtain, starting from different setting of the camera acquisition system, only a relative depth map. Therefore, even if the two images are obtained with different focal distances, $d1$ and $d2$, and the objects of the scene are inside the distance range $[d1, d2]$, the depth map does not provide the precise position of the objects inside of the scene. Such depth maps show the depth relations among the objects inside the scene (close relation of the object with respect to the acquisition device, size relations of the objects, etc). Although the results obtained from these DFD methods are already encouraged, for some applications to know the relative depth map could not be enough (the automatic guide of the vehicles, the exact three-dimensional reconstruction of a scene, etc).

Our work is a novel technique to recover real depth information by the depth map of a real scene. The depth map is a gray-level image representing the depth of each object inside the source image¹¹. In particular, the depth map is obtained by Watanabe-Nayar method⁹. Starting from the depth map, the local maxima of the histogram applied to a central region are collected. The biggest one value (that represents the gray-level of the most considerable object in the working region) is chosen. This gray-level is used by a proper function that maps the depth map gray-levels into real measures.

This function tries to express the depth perception in mathematical law. It also depends on the difference (Δ) between the focal distances of the two images that generate, using the Watanabe method, the depth map. The mapping function approximates a linear function when Δ is small while it becomes exponential when this value increases.

Moreover an algorithm able to select significant depth map values has been implemented. These values are used to calculate the real depth distances of all the object of the image. The results show an error range below the 10%.

The paper is organized as follows. Section 1 describes the proposed methods used to extract the real distance from the relative depth map. Experimental results are given in Section 2. Section 3 concludes the paper.

1. THE PROPOSED TECHNIQUE

The main steps of the proposed technique can be resumed as:

1. Relative depth map generation;
2. Scaling function evaluation;
3. Real depth estimation.

1.1. Depth map generation

In order to generate the depth map, the Watanabe-Nayar DFD method⁹ has been used. This method, starting from two images of the same scene acquired with different camera settings, generates a relative depth map of the objects inside the scene. In particular the two images are acquired with only different focal distance, then the generated depth map will give the depth information for the objects included in the range between the two-selected focal distances.

After the two images have been acquired (Figure 1), these images are filtered with a class of broadband operators that, when used together, provide invariance to the scene's texture. This class of operators is efficient and produces accurate and dense depth maps. Since the operators are broadband, a small number of them are sufficient for depth estimation of scenes with complex textural properties. This class of operators depends on the camera parameter settings. In this way a gray level depth map is generated¹¹ as shown in Figure 2 (pixel level 0 corresponds to a closest object, level 255 to a farthest away one).



Figure 1: Original images used for depth map generation. (a) with the object in-focus and (b) with the background in-focus .



Figure 2: The gray level depth map.

1.2. Scaling function

Starting from the relative depth map obtained with the Watanabe-Nayar method⁹ such information has been converted in a gray level image as shown in Figure 2; the proposed technique estimates the real distance between the selected object in the scene and the acquisition device.

The gray level depth map has the value inside the range [0, 255] and each gray level corresponds to a distance of the lens. If for the acquired images the focal distances of the objects inside the scene are $d1$ and $d2$, the range [0, 255] corresponds, as the real distance, to the range $[d1, d2]$. Unfortunately, the relationship between the gray level range and the real depth values is not linear when the difference $d2-d1$ is more than one meter. In order to solve this problem has been used the following not linear scaling function:

$$y = \frac{\Delta}{x_{\max}} \cdot \frac{1}{e^{c(\Delta)}} \cdot xe^{\left(c(\Delta) \cdot \frac{x}{x_{\max}} \right)} \quad (1)$$

$$c(\Delta) = e^{\frac{\Delta}{\beta}} - 1 \quad (2)$$

$$\beta = \frac{\Delta_{\max}}{\log_e L} \quad (3)$$

where Δ is a value inside the range $[0, \Delta_{\max}]$, $\Delta_{\max} = d1 - d2$ and $c(\Delta)$ is inside the range $[0, L-1]$.

It is possible to change the value of L in order to modify in opportune way the curve (1) and to correct eventual imperfections in the scaling law. If the values of Δ , Δ_{\max} and $x_{\max}(=255$ in our examples) have been fixed and the gray value x , corresponding to a value of the relative depth map, is replaced in (1), the value y , expressed in millimeters, represents the real depth. The Figure 3 shows the plot of the function (1) with $L=4$. Note that for little Δ values, the slope of the function (1) is almost linear, while it will become more similar to an exponential function increasing the Δ value.

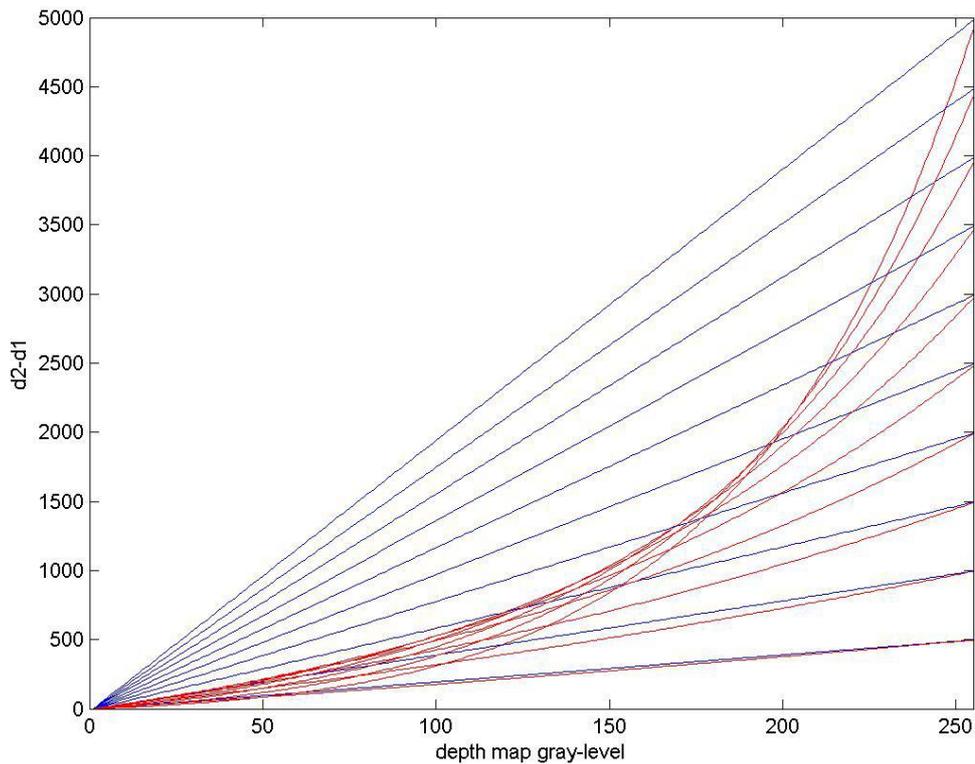


Figure 3: Plot of the function (1) with $L=4$.

1.3. Real depth estimation

The scaling function (1) requires a right estimation of the range $d2-d1$. To correctly evaluate this range the real depth is estimated considering only a region of the depth map. This region corresponds to the regions normally used in AutoFocus techniques. Usually this ensures that the most significant scene information is inside this region. The size of this selected region of the image has been fixed as $1/16$ of the image size and normally is placed in the center of the image but could be positioned everywhere in the image.

To correctly apply (1) is required that the most significant objects in the region are selected. To do this the two biggest local maxima carried out from the histogram of the gray depth map levels of the region are selected. Even if the depth gray level associated with the background is not selected using the histogram, the (1) can be applied correctly because normally the background depth value is near the $d2$ measure.

The Figure 4 shows the histogram obtained when the selection region is placed in the center of the image.

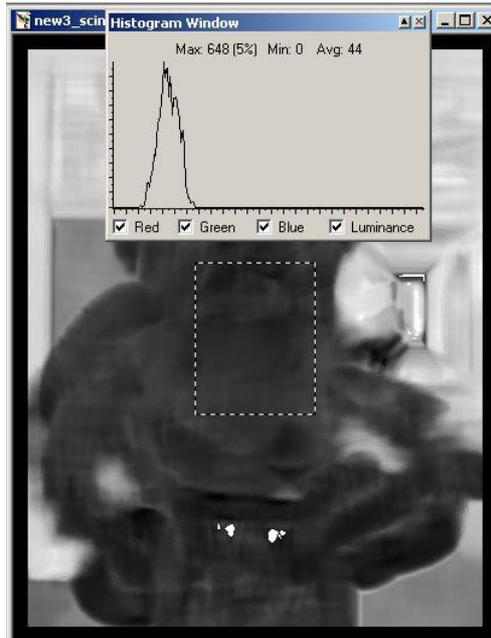


Figure 4: Histogram for the selected region of the gray level depth map.

2. EXPERIMENTAL RESULTS

The input images have been acquired with the Olympus Camedia C-2500L, the manual mode has been set in order to fix the chosen aperture and focal distances.

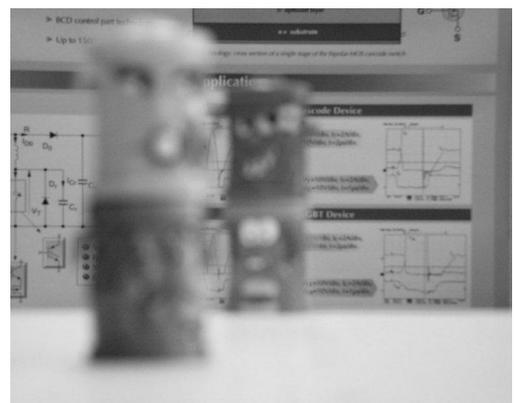
For all results, it is presented:

1. the original acquired images with the two focal distances $d1$ and $d2$;
2. the depth map with the selected region and the related histogram;
3. the distances, computed by the equation (1), related to the most considerable objects in the region;
4. the real distances corresponding to the object inside the region and related error percentage.

In particular, for the images in Figure 1 the original image have been acquired with $d1=500\text{ mm}$ (foreground focused) and $d2=5000\text{ mm}$ (background focused). The two real distances obtained after the selection of the local maxima by the histogram in Figure 4 are $d_{max1}=556\text{ mm}$ (monkey) and $d_{max2}=581\text{ mm}$ (monkey) with an estimated error percentage equal to $e(d_{max1}) = 7.3\%$, $e(d_{max2}) = 3.16\%$ respect to the real distance $d=600\text{ mm}$ of the monkey.



(a)



(b)

Figure 5: Original image with the object in-focus (a) and with the background in-focus (b).

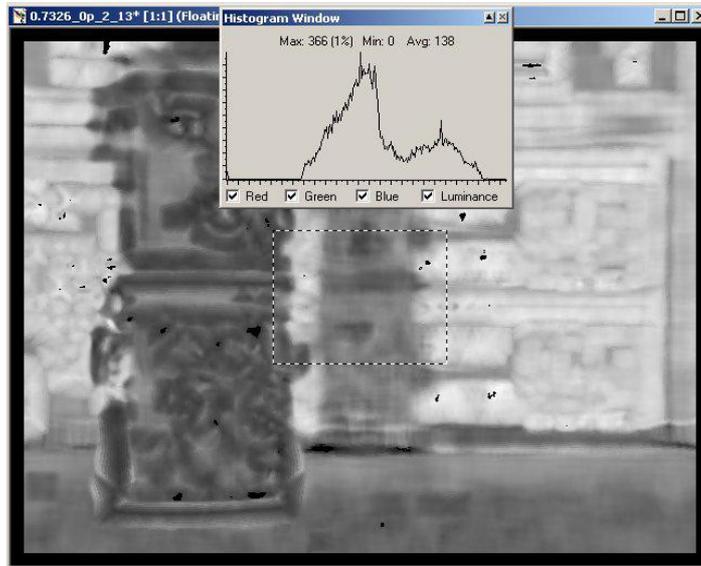


Figure 6: Histogram of the gray level depth map for the selected region.

For the example in Figure 5 and Figure 6:

- Focal distances: $d1 = 300 \text{ mm}$, $d2 = 5000 \text{ mm}$;
- Distances computed by the equation (1): $dmax1 = 759 \text{ mm}$ and $dmax2 = 880 \text{ mm}$;
- Real distances: $d(\text{central object}) = 800 \text{ mm}$;
- Error Percentage: $e(dmax1) = 5.125\%$ and $e(dmax2) = 10\%$.

For the images shown in Figure 7 and

Figure 8:

- Focal distances: $d1 = 300 \text{ mm}$, $d2 = 500 \text{ mm}$;
- Distances computed by the Equation (1): $dmax1 = 391 \text{ mm}$ (central object), $dmax2 = 442 \text{ mm}$ (background);
- Real distances: $d(\text{central object}) = 350 \text{ mm}$; $d(\text{background}) = 500 \text{ mm}$;
- Error Percentage: $e(max1) = 11.7\%$, $e(max2) = 11\%$

Preliminary results confirm the effectiveness of the proposed methodology. Using a database of 100 pairs of images, acquired at different setting and resolution, the estimated error percentage is almost constant. Future works will include also a better performance evaluation of the overall technique.



(a)



(b)

Figure 7: Original image with the object in-focus (a) and with the background in-focus (b).

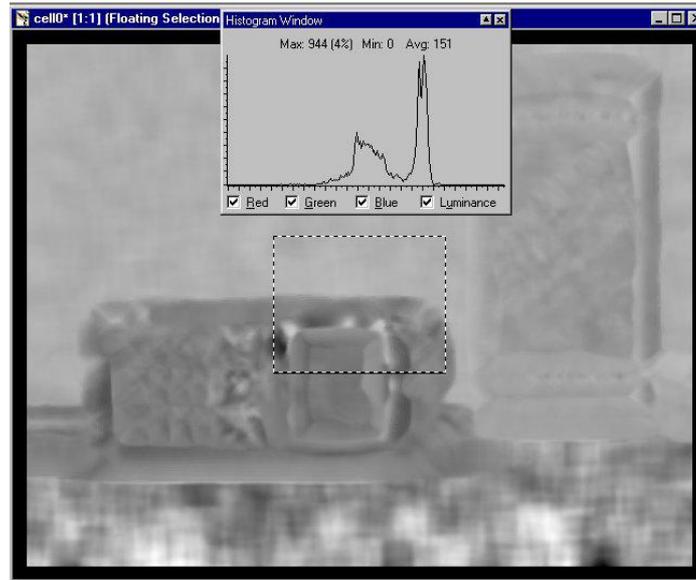


Figure 8: Histogram for the selected region of the gray level depth map.

3. CONCLUSIONS

A model that estimates the real distance between the objects in the acquired image has been discussed. This approach estimates the real distance starting from a gray level depth map obtained using opportune models^{9 11}. The robustness of this method has been tested with a percentage error that is always below the 10%. The percentage error has been evaluated starting from both estimated and real measured distance of the interested object with respect to the acquisition device.

4. REFERENCES

1. G. Surya, M. Subbarao, "Depth from Defocus by Changing Camera Aperture: A Spatial Domain Approach", In Proceedings of CVPR, IEEE Computer Society Conference, 1993.
2. M.Subbarao, G.Surya. "Focused Image Recovery from two Defocused Images Recorded with Different Camera Settings", IEEE *Trans. Image Processing*, **Vol. 4**, No. 12, pp. 1613-1628, Dec. 1995.
3. T. Darrell, K. Wohn. "Pyramid Based Depth From Focus". *Computer Vision and Pattern Recognition*, In Proceedings of CVPR, *Computer Vision and Pattern Recognition*, pp.504-509, June 1988.
4. C. Simon, F. Bicking. "Estimation of Depth on Thick Edges from Sharp and Blurred Images", IEEE *Instrumentation and Measurement Technology Conference*, Anchorage, USA, May 2002.
5. G. Schneider, B. Heit. "Monocular Depth Perception by Evaluation of the Blur in Defocused Images", In Proceedings of IEEE ICIP, *Intl. Conf. on Image Processing*, **Vol. 2**, pp.116-119, November 1994.
6. D.Ziou, S. Wang. "Depth from defocus using the Hermite Transform", In Proceedings of IEEE ICIP, *Intl. Conf. on Image Processing*, **Vol. 2**, pp. 958-962, October 1998.
7. P. Favaro, S. Soatto. "Shape and Radiance Estimation from the Information Divergence of Blurred Images", In Proceedings of ECCV, *European Conference on Computer Vision*, **Vol. 1**, pp.755-768, 2000.
8. J. Rayala, S. Gupta, S.K. Mullick, "Estimation of Depth From Defocus as Polynomial System Identification", In IEE Proceedings of *Vis. Image Signal Process.*, **Vol. 148**, No. 5, October 2001.

9. S.K. Nayar, M. Watanabe, "Rational Filters for Passive Depth from Defocus", *International Journal of Computer Vision*, **Vol. 23**, No.3, pp. 203-225, 1998.
10. A.P. Pentland. "A New Sense for Depth of Field", *IEEE Trans. Pattern Anal. Machine Intelligence*, **Vol. 9**, No. 4 pp.523-531, July 1987.
11. S. Curti, D. Sirtori, F. Vella, "3D Effect Generation from Monocular View", In *IEEE Proceeding of 3DPVT'02, First International Symposium on 3D Data Processing Visualization and Transmission*, pp. 550-553, 2002.