Single-Image Depth Inference Based on Blur Cues

Jingwei Wang*, Hao Xu* and C.-C. Jay Kuo*
* University of Southern California, Los Angeles, California, USA
E-mail: wang1984wei@gmail.com, iamxuhao@gmail.com, cckuo@sipi.usc.edu

***Abstract*—With the rapid advancement of 3D visual technology, the technique of depth inference from a single image has received new attention. In this work, we present several single-image depth inference algorithms based on the blur degree in different regions in one image. We identify two major sources of image blur: camera defocus and atmospheric reflectance. The latter is also known as the haziness. We build models for these two scenarios with the depth information as a model parameter. Thus, we are able to infer the depth information from the observed image. Experimental results are conducted on a large variety of images to demonstrate that the robustness of the proposed depth inference method.**

I. Introduction

Due to the rapid progress in 3D video technology, the problem of depth inference from single or multiple images has received a lot of attention in recent years. Many depth-inference schemes have been proposed by exploiting different visual cues such as motion, stereo and focus. We examine the single-image depth inference problem by using the image blur cue in this paper, where image blur may be caused by camera defocus or atmospheric reflectance.

Many focus-based methods, *e.g.* [1], have been proposed to recover the depth information from one single image. Active illumination method [2] projected a sparse set of dots onto the image scene and estimate the depth information according to the geometry and illumination changes of projected dot patterns. The coded aperture method [3] introduced a modified camera aperture shape pattern to identify the camera's blurring scale, then the depth map can be estimated using blurring scale result. All aforementioned methods need to calibrate camera's aperture or modify camera's scene environment, which may not be an easy task.

To mitigate this requirement, several focus detection algorithms [4], [5] were proposed to recover the depth information without calibration or modification of camera lens and scene environment. For example, the defocus magnification method in [4] models the blur degree along edge points with the Gaussian kernel and propagates the blur degree from edges to the internal region of an object using a colorization approach. The work in [5] proposed an exemplar-based deblur strategy to estimate the blur degree on edge locations and propagates the depth information to other regions using an optimization procedure. However, the work in [4] suffers from severe estimation errors in the propagation stage while the method in [5] fails to provide an accurate edge blur degree. Therefore, both of them find limited applications in practice. It is of great interest to develop an improved focus detection algorithm to recover the depth information accurately.

In this work, we present an effective focus-detection algorithm with improved accuracy. Since it requires no calibration or control of camera lens, it can be used to infer the depth information from a single image captured by conventional cameras. Our focus-detection algorithm adopts a two-stage framework as done in [4], [5], yet it utilizes some techniques to improve its accuracy. First, in the blur estimation stage, it models the blur along edges of an image with Gaussian blur kernel or the atmospheric reflectance ratio. In the propagation stage, we formulate the blur degree propagation as an optimization problem that is similar to the classic soft matting problem [3]. There exists a closed-form solution of high accuracy. Thus, the blur degree can be accurately estimated on image edges and be propagated to the remaining parts effectively. As a result, the accuracy of the inferred depth can be significantly improved.

The rest of this work is organized as follows. Two blur models are examined in Sec. II. Experimental results are shown in Sec. III. Concluding remarks and future research directions are given in Sec. IV.

II. Depth Inference from Blur Models

We examine two blur models caused by camera de-focus and atmospheric reflectance, respectively, and show how to use the model to infer the depth information from a single image in this section.

A. Gaussian Model of Camera Defocus

The blur caused by the classic thin lens model [6] is shown in Fig. 1. When a point lighting source is placed at the focus distance, S, all rays from it will converge to a single point on an image plane, which is at distance f_S from lens, and the result is considered as a sharp point. On the other hand, rays from another lighting source point at distance $D \neq S$, it will be projected into multi points on the same image plane, and the result is a blurred spot, shown in rays denoted by orange lines in the figure. Thus, the blur effect is determined by the distance between the object and the lens.

When there is no blur, an edge of this image located at $x = 0$ can be modeled by an ideal step function as

$$f(x) = Au(x), \tag{1}$$

where $u(x)$ is the step function and A is the amplitude of the edge. If there exists a blur caused by the camera defocus, the image of the object at distance D in Fig. 1 shall be a spot

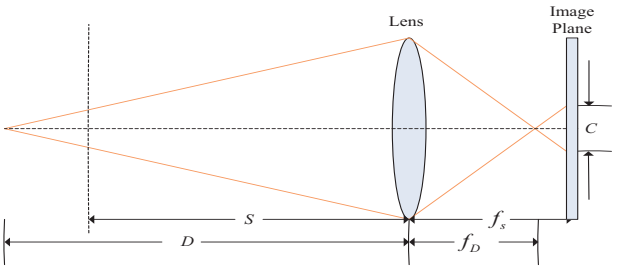


Fig. 1. (a) The lens model [6].

rather than a point. This spot is called the circle of confusion (CoC) [6], which can be used to characterize the blur due to the imperfect focus. Specifically, the diameter of CoC, denoted by c in Fig. 1, can describe the amount of defocus. The relationship between the diameter of CoC, c, and the distance, D, is shown in Fig. 1. Thus, we can infer the depth value from the diameter of CoC. It is worthwhile to point out that c is a "non-linear" monotonically increasing function of D [4]. The diameter of CoC is often determined based on blur estimation. The defocus blur, $i(x)$, can be modeled as a convolution of ideal step function $f(x)$ with a point spread function (PSF) [4]. Specifically, the PSF is usually modeled by a zero-mean Gaussian function $g(x, y, \sigma)$,

$$i(x) = f(x) * g(x, y, \sigma), \tag{2}$$

where the standard deviation σ is proportional to the diameter of the CoC c. Mathematically, we have

$$\sigma = kc,$$

where k is a constant parameter. Thus, the standard deviation σ of PSF can be used as a measure of the diameter of CoC.

Based on the above discussion, we adopt the following focal blur kernel in our work:

$$g(x, y, \sigma_b) = \frac{1}{2\pi\sigma_b^2} \exp\left(-(x^2 + y^2)/2\sigma_b^2\right), \tag{3}$$

which is a zero-mean Gaussian function with standard deviation σ_b. The parameter, σ_b, is called the *blur degree* since it is proportional to the diameter of CoC and can be used to characterize the defocus degree.

Another challenge is the measure of the value of σ_b from a given image. To estimate the blur degree, we study a blurred edge of an input image along the y-axis with amplitude A and blur parameter σ_b, which is indicated by the red curve in Fig. 2. In this way, the blur degree σ_b can be measured by the distance where this curve changes from low to high. Since it is difficult to estimate σ_b directly, we model the edge response to the second derivative filter as proposed by Elder and Zucker in [4]. Mathematically, we have

$$r_2^x(x, y, \sigma_2) = Au(x) * g_2^x\left(x, y, \sigma_b^2 + \sigma_2^2\right) \tag{4}$$

$$= \frac{-Ax}{\sqrt{2\pi}(\sigma_b^2 + \sigma_2^2)^{3/2}} \exp(-x^2/2\left(\sigma_b^2 + \sigma_2^2\right)) \tag{5}$$

where $u(x)$ is a step function, σ_2 is the scale of the second derivative operator and A can be derived from the local extrema within each window around an edge pixel.

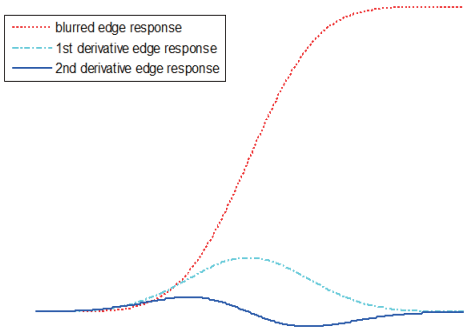


Fig. 2. The blurred edge response model.

Instead of estimating the distance in which the blurring edge response changes from low to high, we measure the distance between the two second derivative extrema of opposite signs in edge's second derivative response as shown in Fig. 2. That is, we evaluate the distance, along the edge's gradient direction by solving a least square fitting problem with the pixel responses model [4]. Then, the size of blur degree σ_b can be evaluated using Eq. (5). Finally, the blur degree on edges of an image can be estimated accordingly.

B. Atmospheric Reflectance Model

Another cause of image blur is atmospheric reflectance. One example is shown in Fig. 3(b), where the Fuji mountain is far away from the viewpoint and it has a clear edge boundary but of low contrast. In this case, the focus degree estimation using only the Gaussian blur kernel is not effective [4] for a very large value of the distance, D. However, the low contrast can be used as a clue for the blur degree estimation based on the following atmospheric reflectance model [7] for an input image I:

$$\tilde{t}(x) = 1 - \omega \min_{c\in\{r,g,b\}} \left(\min_{y\in\Omega(x)} \left(\frac{I^c(y)}{A^c}\right) \right), \tag{6}$$

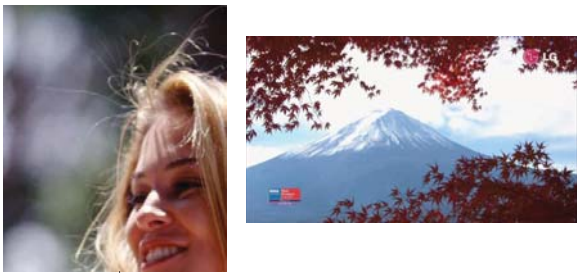
where $\Omega(x)$ is a local patch centered at x, and A is the estimated atmospheric light of each channel. ω is set as 0.9 here in our application. Once $t(x)$ is computed from the above equation, we can use it to infer the depth information d via

$$t(x) = e^{-\beta d(x)}, \tag{7}$$

where β is a constant to be estimated.

III. Experimental Results

In this section, we will show the effectiveness of the proposed algorithm in depth inference. Two other depth recovery algorithms based on focus-detection were implemented for performance comparison. They are the defocus magnification method [4] and the exemplar-based deblur model [5]. All experiments were carried out on two image databases: the



(a) (b)

Fig. 3. Two defocus models: (a) the Gaussian blur model and (b) the atmospheric reflectance model.

DOF-PRO image database [8], which includes mostly camera-focused images, and the Make3D range image database [9], which contains low contrast images with the effect of atmospheric reflectance. In our experiment, 10 images with different scenes were chosen from each database as test examples. They are shown in Figs. 4 and 5.

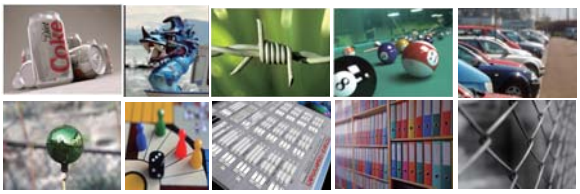


Fig. 4. The test image data for the Gaussian blur.

.

Fig. 5. The test image data for the low-contrast atmospheric reflectance blur.

The "ground-truth" depth maps were obtained using the laser scan render that detects the depth field with laser scanning and can achieve very accurate depth information. We applied three different depth-inference algorithms to all test images. The corresponding test environment (i.e., screen size, monitor type, etc.) is listed in Table I.

TABLE I
SUBJECTIVE EVALUATION

Display	17 inch Dell LCD
Number of Observers	10
Evaluation Method	PSPC

We use the following two methods to measure their performance.

1) Preference Scaling Paired Comparison (PSPC)
The method demands viewers/users to compare two extracted depth maps from different algorithms at each time, and distinguishes them with rating scores (i.e., 4 - best quality; 3 - satisfied quality; 2 - acceptable quality; 1 - unacceptable quality). For example, the average scores for total 20 test images are shown in Table II. Clearly, the proposed algorithm receives the highest score in depth map quality as compared with the two other benchmarking algorithms.

TABLE II
EVALUATION USING THE PSPC METHOD.

method	Relative depth error result
deblur model	1.25
defocus magnification	2.5
Proposed algorithm	3.75

2) Average Relative Depth Error (ARDE) [10]
The extracted depth-maps are compared with the ground-truth ones and the average relative depth error is computed. To make them comparable, we have to scale depth maps obtained from both methods (the inferred depth map and laser-scanned depth map) to the same range, which is set to [0,1] in this work. After than, the ARDE at pixels can be calculated to quantify the error of the extracted depth map. We show the ARDE results of all test images in Table III. Clearly, the proposed depth inference algorithm outperforms the other two benchmarking algorithms.

TABLE III
EVALUATION USING THE ARDE METHOD.

method	Relative error of extracted depth map
deblur model	0.902
defocus magnification	0.699
Proposed algorithm	0.428

In addition to the above two performance measures, we provide several extracted depth maps for visual comparison. To save space, only three images are selected from each image database, and their depth maps extracted from various methods are shown in Fig. 6 (for the DOF-PRO image database) and Fig. 7 (for the Make3D range image database). Clearly, the depth maps obtained by the proposed algorithm have a better match, as shown in the fifth row of Figs. 6 and 7, with the "ground-truth" ones, as shown in the second row of the corresponding figures.

IV. CONCLUSION AND FUTURE WORK

Two major sources of image blur; namely, camera defocus and atmospheric reflectance, were studied and modeled with the depth information as a model parameter. Then, an algorithm was proposed to extract the depth information from the observed image. Experimental results were conducted on a large variety of images to demonstrate that the superior

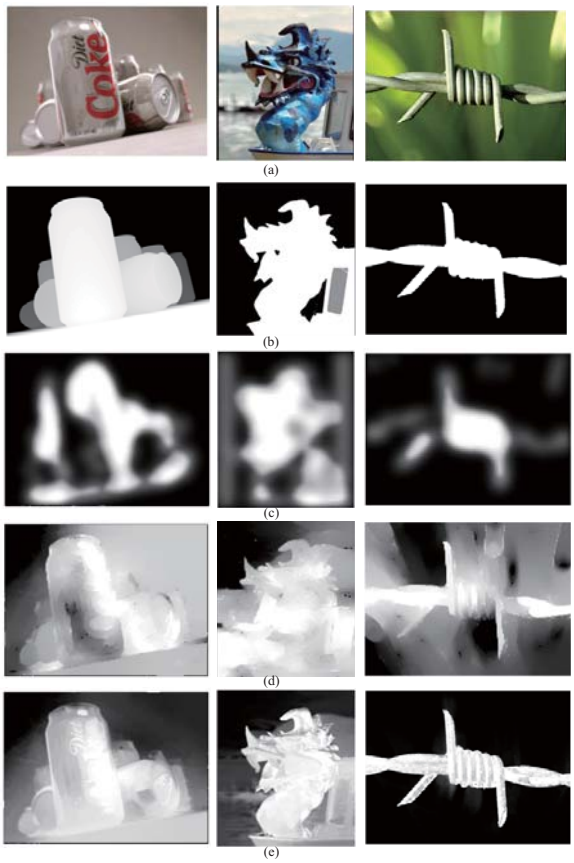


Fig. 6. Visual comparison of extracted depth maps due to camera defocus: (a) the original image, (b) the ground-truth depth map, (c) the depth map result based on the deblur model in [5], (d) the depth map result based on the defocus magnification algorithm in [4], and (e) the depth map result of the proposed algorithm.

performance of the proposed depth inference algorithm over known prior art. We would like to extend the idea in this work, and obtain robust depth inference algorithms for a more generic class of images.

REFERENCES

[1] S. W. Hasinoff, and K. N. Kutulakos, "Confocal stereo," *International Journal of Computer Vision,* vol. 81, pp. 82-104, 2009.

[2] F. Moreno-Noguer, P. N. Belhumeur, and S. K. Nayar, "Active refocusing of images and videos," *ACM Transaction on Graphics,* vol. 26, pp. 67-75, 2007.

[3] A. Levin, R. Fergus, F. Durand, and W. T. Freeman, "Image and depth from a conventional camera with a coded aperture," *ACM Transaction on Graphics,* vol. 18, pp. 1186-1198, 2007.

[4] B. Soonmin, and D. Fredo, "Defocus Magnification," *Computer Graphics Forum,* vol. 26, pp. 571-579, 2007.

[5] Y. Tai, H. Tang, M. S. Brown, and S. Lin, "Detail Recovery for Single-Image Defocus Blur," *IPSJ Transactions on Computer Vision and Applications,* vol. 1, pp. 95-104, 2009.

[6] S. Zhuo, and T. Sim, "On the Recovery of Depth from a Single Defocused Image," *International Conference on Computer Analysis of Images and Patterns(CAIP),* vol. 1, pp. 889-897, 2009.

[7] K. He, J. Sun, and X. Tang, "Single Image Haze Removal Using Dark Channel Prior," *IEEE Transaction Pattern Analysis and Machine Intelligence,* vol. 33, pp. 2341-2353, 2011.

[8] DOFPRO database, http://www.dofpro.com/cgigallery.htm

[9] Make3D Range Image Database, http://make3d.cs.cornell.edu/data.html

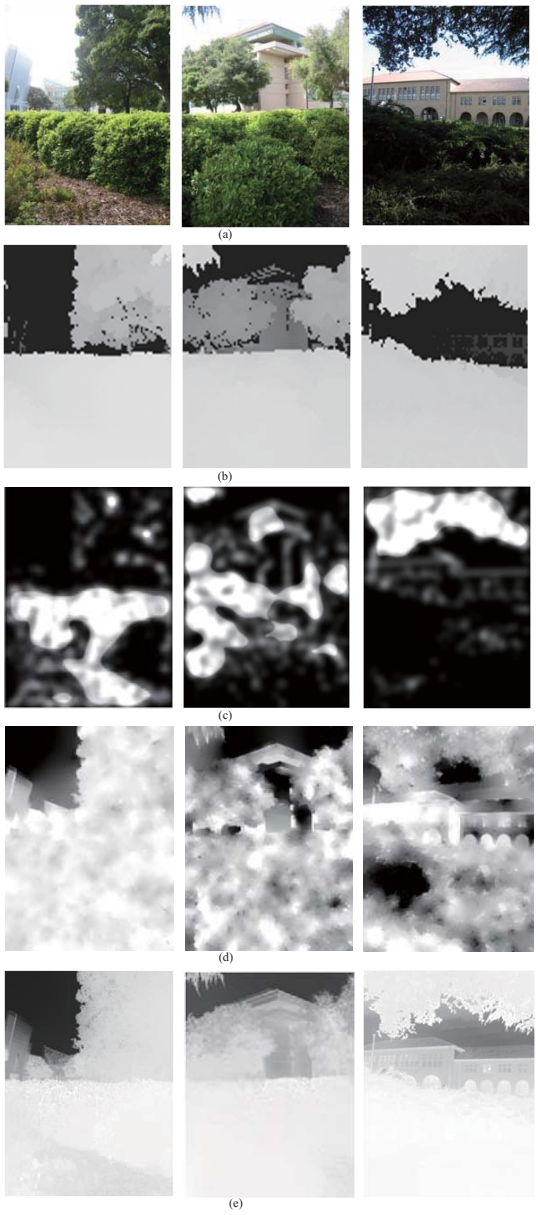


Fig. 7. Visual comparison of extracted depth maps due to atmospheric refletance: (a) the original image, (b) the ground-truth depth map, (c) the depth map result based on the deblur model in [5], (d) the depth map result based on the defocus magnification algorithm in [4], and (e) the depth map result of the proposed algorithm.

[10] A. Saxena, M. Sun, and A. Y. Ng, "Make3D: Learning 3-D Scene Structure from a Single Still Image," *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI),* vol. 31, pp. 824-840, 2009.