A Novel Depth Estimation Method Using Infocused and Defocused Images

Saeed Mahmoudpour and Manbae Kim

Abstract — *The blur amount of an image changes proportional to scene depth. Depth from Defocus (DFD) is an approach in which a depth map can be obtained using blur amount calculation. In this paper, a novel DFD method is proposed in which depth is measured using an infocused and a defocused image. Subbaro's algorithm [4] is used as a preliminary depth estimation method and two complimentary approaches are provided to overcome drawbacks in edge and smooth areas[1].*

Index Terms — Depth map, Image blur, Gradient, Saliency.

I. Introduction

Development of three dimensional (3D) video and image is growing rapidly in devices including 3DTV and mobile phones. In this intermediate period where the creation of 3D content is limited, the development of 2D-to-3D conversion algorithms is needed. A stereoscopic image can be constructed using an infocused 2D image and a scene depth map. Depth information shows relative distance of objects from camera [1]. Depth from Defocus (DFD) is one of 2D-to-3D conversion methods in which depth can be obtained based on blur amount of 2D images using a single camera. In conventional DFD methods, blur value is calculated using two defocused images [2][3]. However, it is well known that the quality of a reconstructed infocused image from defocused images is not satisfactory. Therefore, a novel DFD method is proposed in this paper that uses a pair of infocused and defocused images.

In the proposed method, Subbaro's approach [4] is used to achieve a basic depth map and two additional procedures are considered to improve the depth map quality of the basic method. Using both infocused and defocused images helps to provide useful depth cues to obtain an improved depth map. Furthermore, a proper stereoscopic image can be generated by an infocused image.

II. Proposed Method

A. Subbaro Algorithm

Subbaro [4] modeled an image as a two-variable cubic polynomial in a small (3 x 3 pixels) image neighborhood using spatial domain transform and proved that the relationship between two images g_1 and g_2 with blur values σ_1 and σ_2 can be represented as (1):

$$\sigma_1^2 - \sigma_2^2 = 4(g_1 - g_2)/\nabla^2 g = G \qquad (1)$$

where ∇^2 is a Laplacian operator and $\nabla^2 g$ is defined by (2):

$$\nabla^2 g = (\nabla^2 g_1 + \nabla^2 g_2)/2 \qquad (2)$$

Referring to (1), the blur amount for infocused image g_2 is zero (σ_2=0) and the blur value can be estimated in defocused image according to (3):

$$\sigma_1 = \sqrt{G} \qquad (3)$$

For two edge and non-edge pixels with identical depth, Laplacian value is much larger for edge pixel due to Laplacian edge sensitivity. Then according to (3), the blur calculated for edge pixel is much lower than non-edge pixel blur. Thus, the correct edge depth can not be measured accurately from Subbaro's method. To solve this, a new method is proposed in this paper to obtain a correct depth for edge regions.

B. Edge Blur Estimation

The methodology for the estimation of defocus blur in edge regions is described for 1-D signal first and then extended for 2D image. In Fig. 1, for co-located edge pixels in infocused and defocused images, the difference between gradient magnitudes of step edges is computed. The difference value varies proportional to blur changes.

The edge blur *e(x)* can be modeled as the convolution of an ideal edge *Au(x)+B,* with point spread function (PSF) where *u(x)* is a step function, *A* is an amplitude and *B* is an offset. The PSF for 1-D signal is expressed as a Gaussian kernel *g(x,σ)* (4) and a Gradient for an edge with blur value σ can be represented as (5):

$$g(x,\sigma) = \frac{1}{\sqrt{2\pi\sigma^2}}\exp(-\frac{x^2}{2\sigma^2}) \qquad (4)$$

$$\nabla e(x) = \nabla((Au(x)+B)\otimes g(x,\sigma)) = \frac{A}{\sqrt{2\pi\sigma^2}}\exp(-\frac{x^2}{2\sigma^2}) \qquad (5)$$

where $\nabla e(x)$ denotes the edge gradient at location *x*.

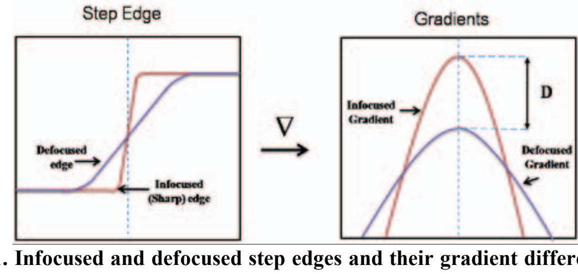

Fig. 1. Infocused and defocused step edges and their gradient difference. Vertical dash line indicates edge location.

Suppose that $\nabla e_i(x)$ and $\nabla e_d(x)$ are gradient magnitudes of infocused and defocused signals respectively. The maximum

[1] This work was supported in part by Small & Medium Business Administration (SMBA) under Technology Innovation Project (S2056930) and in part by the MKE (The Ministry of Knowledge Economy)/KEIT. [10041082, System and Semiconductor Application Promotion Project].

S. Mahmoudpour and M. Kim are with Dept. of Computer & Communications Eng., Kangwon National University, Republic of Korea. (e-mail: {saeed, manbae}@kangwon.ac.kr).

978-1-4799-1291-9/14/$31.00 ©2014 IEEE

gradient amplitude occurs at $x=0$. Therefore, $\nabla e_i(0)$ and $\nabla e_d(0)$ are maximum gradient values. In general, $\sigma_d >> \sigma_i$ and σ_i is a small value near zero. Then, the gradient difference D between infocused and defocused signals is computed by (6).

$$D = \nabla e_i(0) - \nabla e_d(0) = \frac{A}{\sqrt{2\pi\sigma_i^2}} - \frac{A}{\sqrt{2\pi\sigma_d^2}} = \frac{A}{\sqrt{2\pi}}\left(\frac{\sigma_d - \sigma_i}{\sigma_i\sigma_d}\right) \quad (6)$$

Since σ_d is much larger than σ_i and σ_i is a small value near zero, the $\sigma_i\sigma_d$ is small compared to (σ_d - σ_i). Therefore, the gradient difference is proportional to blur changes. Then, we have:

$$\nabla e_i(0) - \nabla e_d(0) \approx \sigma_d - \sigma_i \quad (7)$$

The value σ_i is small and the defocused blur (σ_d) can be estimated from Gradient magnitude difference.

The calculation of blur for 2D image is similar to 1-D signal and Gradient magnitude can be calculated by using (8).

$$\|\nabla e(x,y)\| = \sqrt{\nabla e_x^2 + \nabla e_y^2} \quad (8)$$

where ∇e_x and ∇e_y are Gradients in x and y directions.

In edge regions, a depth map from Subbaro's method is merged with that of the edge blur method by (9).

$$G_{SE} = w_1 g_S + w_2 g_E \quad (9)$$

where G_{SE} is merged depth. g_S and g_E are depth values for Subbaro and edge blur methods respectively. Weighting coefficients are w_1=0.3, w_2=0.7.

Edge blur estimation method helps to overcome Subbaro's weakness on edges. However, a reliable blur estimation is not guaranteed in smooth areas. Gray value difference is an important factor in Subbaro's method and there is no prominent gray difference between infocused and defocused images in such homogeneous regions. Therefore, saliency map is applied to improve a depth map.

C. Saliency Map

Saliency map is the representation of an image that shows some objects saliently in comparison with its surrounding. Achanta et al [5] proposed a method that uses original image and the Gaussian blurred version of the original image to compute saliency. In the proposed DFD method, infocused and defocused images role as original and Gaussian blurred images for saliency algorithm and there is no need to create a Gaussian blurred version of the original image.

Infocused and defocused RGB images are converted to Lab color space and saliency map is computed using (10) :

$$Saliency = \sqrt{(L_{ave} - L_d)^2 + (a_{ave} - a_d)^2 + (b_{ave} - b_d)^2} \quad (10)$$

where L_d, a_d and b_d are Lab color space for the defocused image. L_{ave}, a_{ave} and b_{ave} are mean values of L, a, b for the infocused image. The final depth map is made by taking an average between saliency value and depth value derived from Subbaro and edge blur method (G_{SE}).

III. Experimental Results

Fig. 2 shows input images and results for three scenes. In scene (a), Subbaro's method extracts some acceptable depth information but it is not an ideal result. The final depth map is improved using combination of the edge blur method and saliency with basic depth. Subbaro in scene (b) fails to recover proper depth while two complimentary methods improve final depth map significantly. Depth extraction for scene (c) is challenging due to small object near lens and large smooth regions. Subbaro's result is not satisfactory and edge compare method helps to improve edge depth. Also, objects are not detected in saliency due to their size. Final depth shows improved result. However, some depth data is still missing.

Fig. 2. Blur data of three scenes in columns (a), (b) and (c). In each column, defocused and infocused images are inputs and results of each method and final depth images are shown.

IV. Conclusion

A depth estimation method is proposed using infocused and defocused images. In the proposed method, infocused image helps to simplify calculations for Subbaro's algorithm and edge blur method by assuming infocused blur near zero. Further, using both defocused and infocused images, saliency calculation is more efficient without generating Gaussian blurred image. Therefore, proposed method is suitable for both depth map estimation and practical 3D construction.

References

[1] L. Shen, Z. Zhang and Z. Liu, "Inter mode selection for depth map coding in 3D video," *IEEE Trans. Consumer Electron.,* vol. 58, no. 3, pp. 926-931, Aug. 2011.

[2] Y. Wei, Z. Dong and C. Wu, "Global depth from defocus with fixed camera parameters," *Proc. of IEEE ICMA,* pp. 1887 1892, 2009.

[3] V. P. Namboodiri, S. Chaudhuri, S. Hadop, "Regularized depth from defocus," *Proc. of IEEE ICIP.*, pp. 1520-1523, Oct. 2008.

[4] M. Subbaro, T. Wei and 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, 1995.

[5] R. Achanta and S. Susstrunk, "Saliency detection for content-aware image resizing," *Proc. IEEE ICIP,* pp. 1005-1008, 2009.