Generating Omnifocus Images Using Graph Cuts and a New Focus Measure*

Ning Xu, Karhan Tan, Himanshu Arora and Narendra Ahuja
Beckman Institute & ECE Dept, University of Illinois, Urbana, IL, USA
{ningxu, tankh, harora1,ahuja}@vision.ai.uiuc.edu

Abstract

In this paper, we discuss how to generate omnifocus images from a sequence of different focal setting images. We first show that the existing focus measures would encounter difficulty when detecting which frame is most focused for pixels in the regions between intensity edges and uniform areas. Then we propose a new focus measure that could be used to handle this problem. In addition, after computing focus measures for every pixel in all images, we construct a three dimensional (3D) node-capacitated graph and apply a graph cut based optimization method to estimate a spatio-focus surface that minimizes the summation of the new focus measure values on this surface. An omnifocus image can be directly generated from this minimal spatio-focus surface. Experimental results with simulated and real scenes are provided.

1 Introduction

Imaging cameras usually have only a finite depth of field. In an image captured by a camera, only those scene points within the depth of field of the camera are focused, while other points are blurred. To obtain an omnifocus image of a scene, i.e., an image with every scene point in focus, usually we need to take a sequence of images from the same view point but under different focal settings. Different focal settings correspond to different depths of field. If the union of all depths of field covers the entire depth range of the scene, an omnifocus image could be generated by selecting different points from their most focused images and pasting them together in one image [9, 6].

In order to determine in which image a scene point is mostly focused, the usual practice is to define a function called focus measure on a small neighborhood centered at this scene point and compute its focus measure values in all images. The image having the best focus measure value captures the scene point the clearest. Three different categories of scene points can be distinguished based on the contents in their neighborhoods: 1) The scene point whose neighborhood contains intensity edges in its most focused image. 2) The scene point whose neighborhood is a uniform area in the most focused image, but contains intensity variations in some of its blurred images, arising from the nearby intensity edges. 3) The scene point whose neighborhoods in all images are uniform.

In computer vision literature, many different focus measures have been investigated and compared [4, 7, 8, 11, 12, 2]. However, as pointed out in [1], all these focus measures measure the high frequency content in a neighborhood around a pixel in an image. This works well for the first category of scene points, but might lead to significant misclassification for the second category of scene points. We will see how this kind of misclassification happens in Section 2. In this paper, we provide a new focus measure that is capable of handling both the first and the second categories of scene points.

For scene points of the third category, no focus measure can correctly distinguish their neighborhoods in all images since they are basically indistinguishable. Picking any of the images as the best focused image will suffice. However, if there is imaging noise and gradual lighting changes in the captured images, the focus measure will yield erratic results and the constructed omnifocus image will contain some artifacts in these uniform areas. To eliminate such artifacts caused by imaging noise and lighting changes, we assume spatial continuity of depth in such areas. This is a reasonable assumption as in most real-life cases, surface discontinuities are accompanied by image discontinuities. To find a surface that minimizes surface roughness, we apply a graph cut based optimization method to estimate a continuous surface in the 3D spatio-focus space (2D images with different focal settings). An artifact-free omnifocus image can be composed based on this estimated spatio-focus surface. The most related literature on using graph cut includes a maximum-flow formulation of N-camera stereo correspondence problem presented to optimally estimate the full disparity surface [10].

We analyze the behavior of three representative focus measures in Section 2 and propose our new focus measure in Section 3. Section 4 presents the use of graph cut for gen-

*The support of National Science Foundation under grant ECS 02-25523 is gratefully acknowledged.

0-7695-2128-2/04 $20.00 (C) 2004 IEEE

erating omnifocus images based on our new focus measure. Section 5 presents both synthetic and real data experiments. Section 6 presents concluding remarks.

2 Behavior of three representative existing focus measures

In this section we analyze the behavior of three representative existing focus measures in image areas having different degrees of defocus ranging from intensity edge to uniform area, containing all the three categories of scene points. Consider a 2D image with intensity distribution

$$f(x,y) = \begin{cases} 1 & x < 0, \\ 0 & x \geq 0, \end{cases} \tag{1}$$

and assume that the Point Spread Function (PSF) is a two dimensional Gaussian

$$h(x,y,\sigma) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}}, \tag{2}$$

where σ is determined by scene depth and imaging parameters. For different σ, the blurred version of image $f(x,y)$ is given by

$$g_\sigma(x,y) = f(x,y) * h(x,y,\sigma). \tag{3}$$

Three representative focus measures, as suggested in [11], are the variance, energy of image gradient, energy of Laplacian of the image, of an $N \times N$ window around a pixel, i.e.

$$M_1 = \frac{1}{N^2} \sum_x \sum_y (g_\sigma(x,y) - \mu)^2, \tag{4}$$

$$\text{where} \quad \mu = \frac{1}{N^2} \sum_x \sum_y g_\sigma(x,y), \tag{5}$$

$$M_2 = \sum_x \sum_y (g_x^2 + g_y^2), \tag{6}$$

$$M_3 = \sum_x \sum_y (g_{xx} + g_{yy})^2, \tag{7}$$

where g_x, g_y are image gradient and $g_{xx}+g_{yy}$ are Laplacian of the image.

$M_i, i = 1,2,3$, is aimed at maximizing the focus measure in the most focused image. However, if we plot M_i as a function of σ for different values of x, it is shown that not every curve in the plot is maximized at $\sigma = 0$, where the image is most focused. In Fig.1, we plot the curves for $x = 1,2,3,4,5$ for each focus measure with $N = 5$ and σ varying from 0 to 6. The focus measure values are scaled for illustration. As we can see in this figure, the focus measure M_i is maximized at $\sigma = 0$ only when $x = 1$ and $x = 2$. For other x values, M_i is maximized at $\sigma > 0$, where the image is blurred.

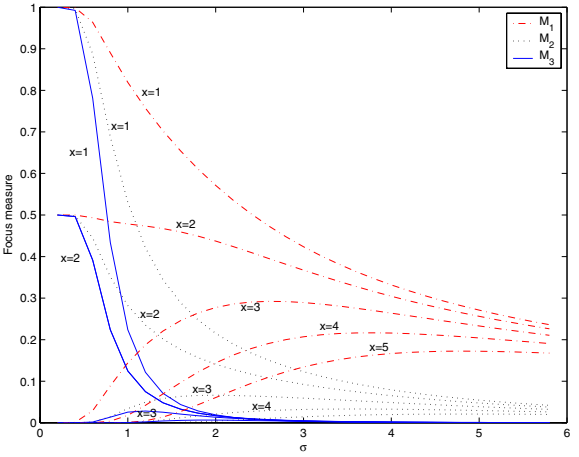


Figure 1. The plot of focus measures $M_i, i = 1,2,3$ for points at different distance, x, from the step edge, with $N = 5$ and $0 \leq \sigma \leq 6$. The focus measure values are scaled for illustration.

The reason why this happens is that high frequency contents outside the uniform local window are blurred into this window when they are out of focus. Therefore, the local window may have a greater high frequency content when it is blurred than when it is focused.

3 A new focus measure

When using the existing focus measures such as $M_i, i = 1,2,3$ to detect which image is most focused for a scene point of the second category, significant misclassification might happen. To handle this problem, we propose a new focus measure M_0. Different from existing focus measures, M_0 takes a 3D volume data ($N \times N \times$ number of images) into account by defining two values V_{min} and V_{max} as

$$V_{\min} = \min_{\sigma;(x,y)\in D} g_\sigma(x,y), \tag{8}$$

$$\text{and} \quad V_{\max} = \max_{\sigma;(x,y)\in D} g_\sigma(x,y), \tag{9}$$

where D is the $N \times N$ local window. The new focus measure M_0 is proposed to measure the distribution of gray levels between these two values, i.e.,

$$M_0 = \frac{1}{N^2} \sum_x \sum_y (\min(g_\sigma(x,y) - V_{min}, V_{max} - g_\sigma(x,y)))^2. \tag{10}$$

The reason why we select this focus measure is based on the fact that the blurring will never increase the gray level range in an image, but only flatten the distribution within the gray level range. Therefore, assuming there are two intensity modes within the local neighborhood, the distribution in a better focused image will be more concentrated

around these two intensity modes. We approximate the center of these two modes using the minimum and maximum intensity values in order to reduce the computation. When a scene point is best focused in an image, M_0 will be minimized (This is different from the traditional maximizing focus measures). We also plot M_0 as a function of σ for different values of x in Fig. 2. As we can see, M_0 is minimized at $\sigma = 0$ for all x values.

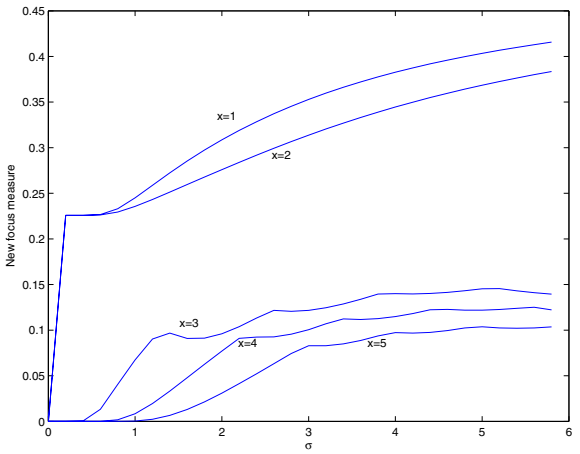


Figure 2. The plot of new focus measure M_2 for points at different distance, x, from the step edge, with $N = 5$ and $0 \leq \sigma \leq 6$.

4 Omnifocus image via graph cut

Given an image sequence, $I(x, y, i)$, $i = 1, 2, ..., n$, of a scene from the same viewpoint with different focal settings, a straightforward approach to generate an omnifocus image $O(x, y)$ of the scene would be to first compute a focus measure value $F(x, y, i)$ for each image point (x, y, i) using focus measure M_0, and then obtain a spatio-focus surface $S(x, y)$, such that

$$S(x, y) = \arg\min_i F(x, y, i), \tag{11}$$

and assign each pixel (x, y) of the omnifocus mosaic $O(x, y)$ as

$$O(x, y) = I(x, y, S(x, y)). \tag{12}$$

However, the resulting omnifocus image $O(x, y)$ will contain intensity discontinuities in uniform regions appearing as artifacts. This is mainly because we estimate $S(x, y)$ pixel by pixel and there are imaging noise and lighting changes in the captured images. For a scene point of the third category, there is no effect of blurring edge in any of the images. SNR is very low in all images and the focus measure will therefore give erratic estimate of $S(x, y)$ for this point. Therefore, the estimated spatio-focus surface $S(x, y)$ will be discontinuous in the uniform regions. The discontinuities in $S(x, y)$ together with imaging noise and lighting condition changes will cause artifacts in the resulting omnifocus image $O(x, y)$.

In order to generate an artifact-free omnifocus image, we propose to use a graph cut based minimization approach to estimate a continuous spatio-focus surface $S(x, y)$ that minimizes the summation of the focus measure values along the surface. We construct a three dimensional node weighted graph $G = (V, E)$ as in Fig. 3. Each pixel $I(j, k, i)$is represented as a node $v(j, k, i)$ with a non-negative capacity $c(v) = F(j, k, i)$, where $1 \leq j \leq w$, $1 \leq k \leq h$, $1 \leq i \leq n$, and $w \times h$ is the image size. Each node is connected to its six neighboring nodes. In this graph, we also add a source node s and a sink node t, connecting to the nodes corresponding to pixels in the first image and last image, respectively.

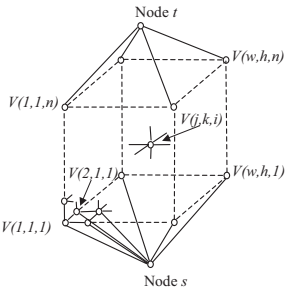


Figure 3. 3D node weighted graph.

An $s - t$ node cut in such a node capacitated graph is a node set $C \subset V$, and by removing C from G, s and t are disconnected. The capacity $c(C)$ of a node cut C is summation of the capacities of all nodes in $C \subset V$, i.e. $c(C) = \sum_{v \in C} c(v)$. An $s - t$ minimum node cut yields a minimum capacity of all possible node cuts that separate s and t. Discussion on how to transform a node capacitated cut problem into an edge capacitated graph cut problem can be found in [5], and the edge capacitated graph cut algorithms can be found in [3].

Therefore, by using node cut algorithm, we could obtain a node cut C corresponding to a continuous and optimal spatio-focus surface $S(x, y)$. Since there might be more than one value for each pixel (x, y) on surface $S(x, y)$, when we compute omnifocus image $O(x, y)$, we modify the equation (12) as

$$O(x, y) = I(x, y, m(x, y)), \tag{13}$$

$$\text{where} \qquad m(x, y) = \min(S(x, y)). \tag{14}$$

The continuous and optimal spatio-focus surface $S(x, y)$ will yield a better omnifocus image. The continuity will reduce the image artifacts and the optimal solution will reduce the errors caused by imaging noise and lighting condition changes.

5 Experimental Results

5.1 Synthetic images

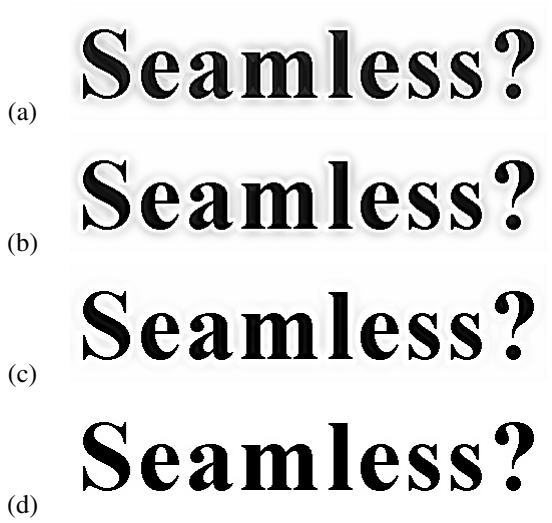


Figure 4. Omnifocus image generated by (a) M_1, (b) M_2, (c)M_3, and (d) M_0.

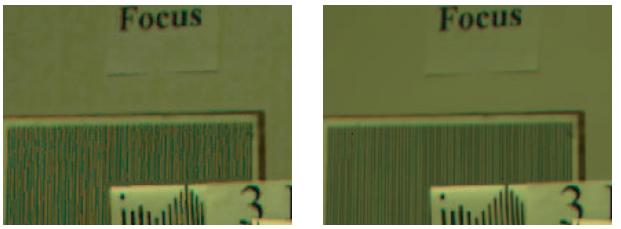
We simulate an image sequence with different amount of gaussian blurring from a focused image containing many points belonging to the second and third categories. Omnifocus images are computed by using focus measures M_1, M_2, M_3 and M_0, respectively. Fig. 4 shows our comparison results. Since there is no imaging noise and changing of lighting conditions, all omnifocus images are generated using the simple method indicated by Eqn. 12. The root mean square (RMS) errors of the generated omnifocus images compared with the original image are $33.67, 23.16, 10.08$ and 0.65, respectively. It could be seen that the omnifocus images in Fig. 4 (a-c) contain some artifacts around the intensity edges. We also conduct experiments on images containing much less scene points of the second and third categories, the RMS errors are small for all focus measures while M_3 performs the best.

5.2 Real data experiment

We capture a sequence of real images with different focal settings. The scene contains two objects at distances of 3 feet and 4 feet, respectively, and a wall at about 7 feet. Fig. 5 (a) shows the omnifocus image generated using the simple approach indicated by Eqn. 12 and Fig. 5 (b) shows the omnifocus image generated using optimization, in each case using focus measure M_0. It can be seen that in the uniform region, there are many artifacts in Fig. 5(a) while much less artifacts can be perceived in Fig. 5(b).

6 Conclusions

In this paper, we presented a method of generating artifact-free omnifocus images from a sequence of different focal setting images captured from the same viewpoint. We first proposed a new focus measure M_0 that could handle the known problem of existing frequency based focus measures. Then, we formulated a global optimization problem and solved it with a graph cut based optimization method to obtain a continuous and optimal spatio-focus surface and generate artifact-free omnifocus image based on it. The future work will focus how to obtain an accurate range map based on the spatio-focus surface computed by graph cut.



(a) (b)

Figure 5. The omnifocus image generated using (a) the approach indicated by Eqn. 12, (b) the proposed graph cut approach.

References

[1] M. Aggarwal and N. Ahuja. On generating seamless mosaics with large depth of field. In *International Conference on Pattern Recognition*, volume 1, pages 588–591, 2000.

[2] N. Asada, H. Fujiwara, and T. Matsuyama. Edge and depth from focus. *International Journal of Computer Vision*, 26(2):153–163, 1998.

[3] T. Cormen, C. Leiserson, and R. Rivest. *Introduction to Algorithms*. McGraw–Hill Companies, 1990.

[4] T. Darrel and K. Wohn. Pyramid based depth from focus. In *Conference on Computer Vision and Pattern Recognition*, pages 504–509, 1988.

[5] L. Ford and D. Fulkerson. *Flows in Networks*. Princeton University Press, 1962.

[6] K. Kaneda, S. Ishida, A. Ishida, and E. Nakamae. Image processing and synthesis for extended depth of field of optical microscopes. *The Visual Computer*, 8:351–360, June 1992.

[7] E. P. Krotkov. *Active Compputer Vision by Cooperative Focus and Stero*. Springer-Verlag, 1989.

[8] S. K. Nayar. Shape from focus system. In *Conference on Computer Vision and Pattern Recognition*, pages 302–308, 1992.

[9] R. Pieper and A. Korpel. Image processing for extended depth of field. *Applied Optics*, 22(10):1449–1453, May 1983.

[10] S. Roy and I. Cox. A maximum-flow formulation of the n-camera stereo correspondence problem. *IEEE Proc. of Int. Conference on Computer Vision*, pages 492–499, 1998.

[11] M. Subbarao, T. Choi, and A. Nikzad. Focussing techniques. *Optical Engineering*, 32(11):2824–2836, 1993.

[12] Y. Xiong and S. Shafer. Depth from focusing and defocusing. In *DARPA93 Image Understanding Workshop*, pages 967–976, 1993.