On Generating Seamless Mosaics with Large Depth of Field

Manoj Aggarwal and Narendra Ahuja
University of Illinois at Urbana-Champaign,
Beckman Institute,
405 N. Mathews Ave, Urbana IL 61801.
E-mail: {manoj,ahuja}@vision.ai.uiuc.edu
Internet: http://vision.ai.uiuc.edu/

Abstract

Imaging cameras have only finite depth of field and only those objects within that depth range are simultaneously in focus. The depth of field of a camera can be improved by mosaicing a sequence of images taken under different focal settings. In conventional mosaicing schemes, a focus measure is computed for every scene point across the image sequence and the point is selected from that image where the focus measure is highest. We have, however, proved in this paper that the focus measure is not the highest in the best focussed frame for a certain class of scene points. The incorrect selection of image frames for these points, causes visual artifacts to appear in the resulting mosaic. We have also proposed a method to isolate such scene points, and an algorithm to compose large depth of field mosaics without the undesirable artifacts.

1. Introduction

The problem of mosaicing a image from a sequence has received considerable attention in recent literature. It involves taking multiple images of a scene with varying imaging parameters (e.g., focus settings, viewpoints), and combining them into a single image. The combined image better captures the scene in some a priori desired ways, e.g., it may be a panoramic view of the scene showing the entire visual field covered piecemeal by the original images [8, 10, 12], or it may have better image quality such as large depth of field [4, 9], or it may have both of the above [5]. The combining process is often referred to as mosaicing. Image mosaicing involves selecting different regions from the sequence of images and pasting them together in one image. The regions are selected on the basis of special properties they posses. The two issues involved in this process are the correct selection and seamless pasting of the regions. In this paper, we will investigate these two issues for the case of mosaicing images to obtain large depth of field.

Images with large depth of field can be composed from a sequence of images taken at different focal settings [4, 5, 9]. The imaged scene consists of multiple regions and each region gets imaged at varying degrees of defocus. In the currently reported mosaicing techniques, for every scene point a focus measure is computed across the image sequence and it is then extracted from that image for which the focus measure is maximum. These extracted regions are then composed into a single mosaic. The resulting mosaic has large depth of field, but it has intensity discontinuities which appears as seams. A number of definitions for focus measure have been proposed in literature [1, 3, 6, 7, 11], all of which measure the high frequency content in a window around that point. This paper investigates, why seams can appear in mosaics generated by the above mentioned techniques and also proposes a joint solution for suppressing the seams, while preserving the large depth of field property.

The issues of seamless pasting and region selection are related. In many instances there are multiple images from which a region can be chosen without compromising the resulting depth of field. In particular, for uniform areas, which are away from textured regions, defocusing does not change their intensity distribution. Thus ideally any image can be chosen for extracting such regions. However, in practice due to lighting changes, vignetting etc., corresponding regions in the image sequence may not have exactly the same intensity. We define a 2-D array containing the sequence numbers of the image (or image numbers) used to extract the regions which are pasted in the mosaic. If the acquired images are sequenced in order of the focal settings then in regions of smooth depth variation, the image numbers must also vary smoothly. We have shown in this paper, that if the image-selection criteria is solely based on the highest focus measure, the smoothness property is not satisfied. Specifically, we have shown that for uniform regions focus mea-

0-7695-0750-6/00 $10.00

sure is highest in a image where nearby textured regions are defocussed. This behavior results from the finite size of the window function used to compute the focus measure. Consider a point A in a uniform region that is at least a units away from the nearest edge. If a rectangular window of size $2a$ is used to compute the focus measure, then its value will be very small if the nearest edge is in focus. But, as the nearest edge gets more defocussed, an intensity gradient start to appear near the edge, which then contributes to the focus measure computed at A. Thus, for scene points such as A, an image selection criterion based on the highest focus measure would choose an image in which that point is defocussed. Consequently, there will be a large disparity among the image numbers chosen for uniform regions and the nearby textured regions, which results in seams in the final mosaic. To avoid seams, we propose to isolate uniform regions from the textured regions and create the mosaic in two steps. In the first step, the focussed image selection based on the highest focus measure and pasting into a mosaic is performed *only* for the identified textured regions. In the second step, the points belonging to uniform regions are assigned image numbers that smoothly propagate the image numbers assigned to nearest textured regions. Since the textured regions are chosen from images with largest focus measure and uniform regions from images to ensure similarity with nearby textured regions, the final mosaic simultaneously has large depth of field and no visual seams.

We note that the focus analysis presented in this paper does not strictly apply to regions close to occluding edges, as the Gaussian blurring model is no longer valid [2]. The focus measure window near occluding edges, might span objects at multiple depths. The objects in such a window have different blur parameters and these parameters change independently as focal settings are changed. Consequently the focus measure for occluding edges are expected to behave differently from pattern edges. A more careful analysis, which incorporates the blurring model near occluding edges is a topic for further research.

In Section 2, we analyze the behavior of a representative focus criterion in textured, uniform regions and regions on the boundary of textured and uniform regions, across a sequence of images with varying degrees of defocus. The behavior of the focus measure is then used to derive conditions for isolating uniform regions from textured regions. For mathematical tractability, we have selected our focus measure to be a sum over a window of size $2a$ of the derivative of the image and have assumed a Gaussian blurring model to represent defocus. This measure is related to the Tenengrad's focus measure [6]. In Section 3, we present experimental results to substantiate our claims.

2. Behavior of a focus criterion across image sequence

Consider a 2-D imaging geometry and a object placed perpendicular to the optical axis, with intensity distribution

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

Assuming Gaussian blurring model, the resulting image $g_\sigma(x)$ is given by

$$g_\sigma(x) = f(x) * h(x;\sigma) \tag{2}$$

where σ is the blur parameter determined by the distances of the object and the sensor from the lens; $h(x;\sigma)$ is the Gaussian kernel with parameter σ given by

$$h(x;\sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp(-\frac{x^2}{2\sigma^2}) \tag{3}$$

The blurred image $g_\sigma(x)$ can be expressed as

$$g_\sigma(x) = \begin{cases} 0.5 & x = 0 \\ \int_{-x}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}} \exp(-\frac{x^2}{2\sigma^2})dx & x < 0 \\ \int_{x}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}} \exp(-\frac{x^2}{2\sigma^2})dx & x > 0 \end{cases} \tag{4}$$

The derivative of $g_\sigma(x)$ is given by

$$g'_\sigma(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp(-\frac{x^2}{2\sigma^2}) \tag{5}$$

Let the size of the averaging window be $2a$, then the focus measure $fm(x,\sigma)$ can be computed by integrating the derivative $g'_\sigma(x)$ over a window of size $2a$.

$$fm(x,\sigma) = \begin{cases} 0.5erf\left(\frac{x+a}{\sqrt{2\sigma^2}}\right) - 0.5erf\left(\frac{x-a}{\sqrt{2\sigma^2}}\right) & x \geq a \\ 0.5erf\left(\frac{x+a}{\sqrt{2\sigma^2}}\right) + 0.5erf\left(\frac{x-a}{\sqrt{2\sigma^2}}\right) & x < a \end{cases} \tag{6}$$

This function is plotted as a function of σ, for different values of x, in Fig. 1. We refer to such a plot as the spatio-temporal plot for focus measure [1], where temporal refers to the variation in σ across the image sequence. The figure shows that for $x < a$, $fm(x,\sigma)$ is unimodal and attains its maximum at $\sigma = 0$, while for $x > a$, the plot has two maxima and one minimum and the minimum is attained at $\sigma = 0$. This behavior implies that for all points close to the edge such that $|x| < a$, the computed focus measure would be the largest for the case when the image is best focussed, however for points $|x| > a$, the computed focus measure would be largest when the image is defocussed ! This suggests that the output of the focus criterion is correct only in regions close to edges and for regions away from edges the focus criterion would actually point

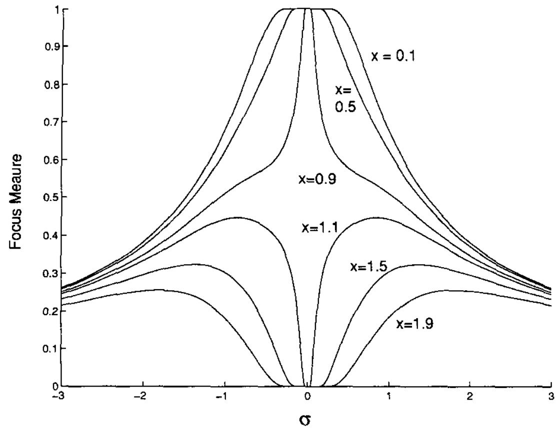


Figure 1. The spatio-temporal plot of focus measure for point at different distances x from the step edge. The window size $a = 1$

to a defocussed image. Since the corresponding regions in images with large differences in imaging parameters can be quite different, choosing the defocussed image for a region $|x| > a$ would generate a seam in the final mosaic. Thus, in order to prevent seams, the uniform regions for which the focus criterion chooses an incorrect image, need to be isolated, so that the correct image numbers can be chosen using the method proposed in Section 1. The selective treatment given to uniform and textured regions ensures that the resulting image has large depth of field while simultaneously being seamless.

It may seem that uniform regions can be easily identified by evaluating a local derivative measure around every point in the scene. This would indeed be the case if we have a focussed image of the scene, however, we have a sequence of images of the scene at different focal settings. In fact, each of the images have spatially varying blur, thus, using this simple derivative measure would in general yield significant misclassification. For example, some textured regions may yield a low derivative measure in a image where it is defocussed, on the other hand, uniform regions close to edge boundary will have a large local derivative, when the image is blurred as shown in Fig. 1. We propose a classification scheme based on the spatio-temporal behavior of the focus measure for uniform and textured regions. A number of criteria can be derived from the spatio-temporal plot in Fig. 1 to determine uniform regions. The spatio-temporal plot for uniform regions has 3 variations (2 maxima, 1 minima) while regions close to the edge have only one variation. Experimentally, for real images it has been found that in general for uniform regions the spatio-temporal plot is quite erratic due to camera noise, while for textured regions there are usually few (typically 1-2) dominant variations. In addition, for uniform regions, the focus measure is usually small for most images in the sequence. The uniform regions can then identified to be all those points whose focus measure is erratic with σ and is small for most values of σ.

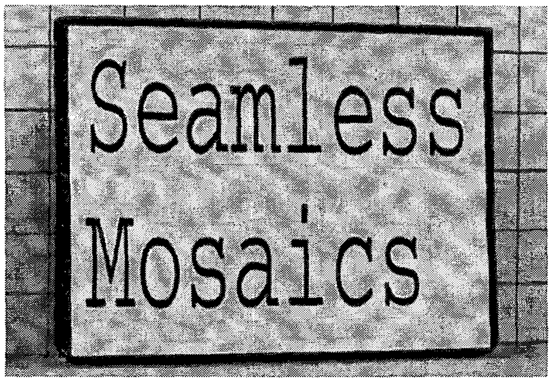
3. Experimental Results and Conclusions

In this section, we compare the performance of the proposed and the conventional algorithm for constructing large depth of field mosaics. In a conventional scheme, a focus measure is computed for every scene point across the image sequence and that point is picked from the image where the focus measure is the highest. We have used the Tenengrad focus measure for our experiments, with window size of 20 pixels. For the first experiment, the scene consists of single planar surface. A sequence of images at different focus settings are obtained and they are mosaiced using the conventional scheme (Fig. 2(a)) and the proposed scheme (Fig. 2(b)). The seams in the mosaic obtained using the conventional scheme are quite evident. It is interesting to note that the visual artifacts (*blotches*) surround the edges like a halo. This is quite consistent with what we had earlier concluded from the spatio-temporal plot of the focus measure. We had shown that regions within the 20-pixel wide averaging window around edges are selected from the correct frame, however, uniform regions beyond that window are selected from defocussed frames, where the intensity of those uniform regions might be quite different due to lighting variations. In a second experiment, we placed three objects at different distances from the camera. The background grid of lines, the newspaper-wrapped cylinder and the patterned grid are at distances 3.5', 2.5' and 1.5', respectively, from the camera. The corresponding mosaics are shown in Fig. 3. The mosaics obtained by using either the conventional or the proposed scheme are equally focussed, but the one generated using the proposed algorithm has no visual artifacts. We also note that the conventional scheme did not create the undesirable blotches in the newspaper and the patterned grid region. This was also expected, since all the uniform areas within these regions are within the averaging window of the textured areas. However, for the background grid of lines, this is not the case and consequently the blotches are present.

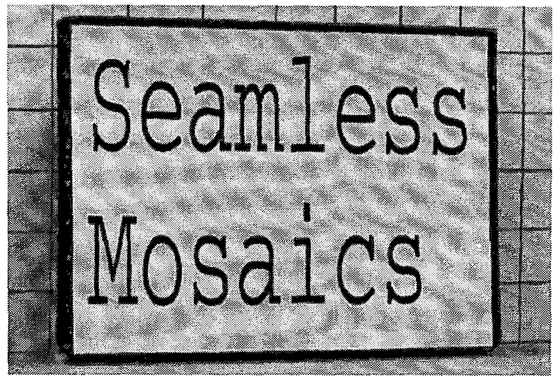
References

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

[2] N. Asada, H. Fujiwara, and T. Matsuyama. Seeing behind the scene: Analysis of photometric properties of occluding edges by the reversed projection blurring model. *IEEE*



(a)



(b)

Figure 2. Mosaic of a planar scene (a) Conventional method, (b) Proposed Method

Transactions on Pattern Analysis and Machine Intelligence, 20(2):155–167, February 1998.

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

[4] 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.

[5] A. Krishnan. *Non-frontal Imaging Camera*. PhD thesis, University of Illinois at Urbana-Champaign, 1997.

[6] E. P. Krotkov. *Active Computer Vision by Cooperative Focus and Stereo*. Springer-Verlag, 1989.

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

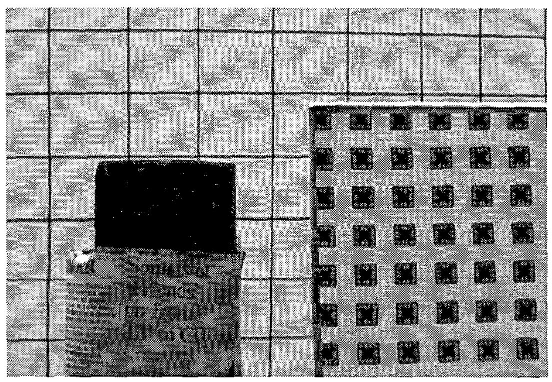
[8] S. Peleg. Panoramic mosaics by manifold projection. In *Conference on Computer Vision and Pattern Recognition*, pages 338–343, June 1997.

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

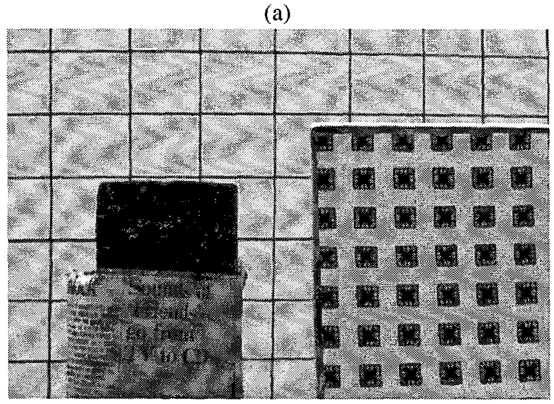
[10] H.-Y. Shum and R. Szeliski. Panoramic image mosaics. Technical Report MSR-TR-97-23, Microsoft Research, 1997.

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

[12] R. Szeliski. Image mosaicing for tele-reality applications. Technical Report CRL 94/2, Digital Equipment Corporation, 1994.



(a)



(b)

Figure 3. Mosaic of a scene with multiple objects at different depths (a) Conventional method, (b) Proposed method