MULTIFOCUS IMAGE FUSION BY ESTABLISHING FOCAL CONNECTIVITY

H. Hariharan, A. Koschan and M. Abidi

Imaging Robotics and Intelligent Systems Laboratory, ECE Department,
University of Tennessee, Knoxville -37996.

ABSTRACT

Multifocus fusion is the process of unifying focal information from a set of input images acquired with limited depth of field. In this effort, we present a general purpose multifocus fusion algorithm, which can be applied to varied applications ranging from microscopic to long range scenes. The main contribution in this paper is the segmentation of the input images into partitions based on focal connectivity. Focal connectivity is established by isolating regions in an input image that fall on the same focal plane. Our method uses focal connectivity and does not rely on physical properties like edges directly for segmentation. Our method establishes sharpness maps to the input images, which are used to isolate and attribute image partitions to input images. The partitions are mosaiced seamlessly to form the fused image. Illustrative examples of multifocus fusion using our method are shown. Comparisons against existing methods are made and the results are discussed.

Index Terms—Depth of focus, focal connectivity, image fusion, image partitioning, multifocus fusion.

1. INTRODUCTION

When obtaining images of a 3-dimensional scene it is desirable to have all objects in the scene to be in focus. Cameras suffer from the problem of limited depth of field (DOF) and this disallows a typical imaging system to obtain such an all-in-focus image. This is a prominent problem in inspection of microscopic scenes. In multifocus fusion, one aims to acquire information from different focal planes and fuse them into one image where all objects in the scene appear to be in focus, as shown in Figure 1. Previous work by others investigated solutions to this problem by using techniques based on tiling, multiscale decomposition (MSD), and learning methods. In this effort, we present a general purpose multifocus fusion technique that is able to fuse images acquired in varied applications such as microscopy, long range imaging, etc. Many datasets in the literature use input images wherein objects are placed well apart in the 3-D environment. Commonly, two input images are fused. Multifocus fusions of such datasets are relatively and easier task. In certain applications, such as microscopy,

the narrow DOF requires multiple shots of the 3D scene to gather all the information contained in the scene. Since the focal planes lie close to each other, there are overlapping sections of a focused object in consecutive frames making multifocus fusion a challenging task. We present a technique that fuses multiple focal planes with narrow overlapping sections of the scene. Typical segmentation based methods employ edges from objects for segmentation. The main contribution in this paper is that we segment regions from the input images based on focal connectivity, and not on object connectivity. We employ such partitions for unifying information from all the focal planes. Before going into details of our method in section 3, we present the related work in section 2. We have compared out results with seminal works in the literature and we present experimental results in section 4, before presenting our conclusions in section 5.

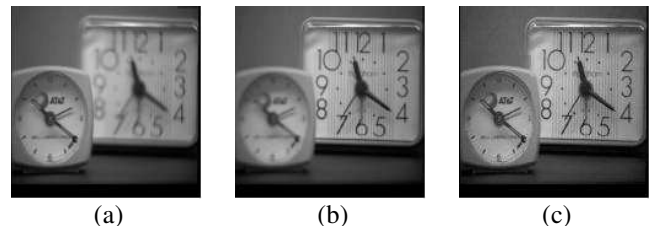(a)   (b)   (c)

Figure 1. An example of multifocus fusion (a) Clock face on left focused (b) Clock face on right focused and, (c) Both clock faces fused by our multifocus fusion.

2. RELATED WORK

Multifocus fusion has been performed using various approaches in the past. The three seminal approaches are based on region selection methods [1,2], multiscale decomposition methods [3,4] and learning based methods. In region selection methods, the input images are initially divided in regions, typically blocks [2] or into segments using a segmentation technique [5, 6]. From sets of such regions, one region per set is chosen based on a sharpness criterion and mosaiced or blended to form the final fused image. The value of the sharpness criterion increases and decreases as objects come into focus and go out of focus, or if the contrast changes in the scene [7]. In region based methods, regions are typically selected in the image based

on edges. Liao *et al* use Hough transform to select regions which are mosaiced for the fused image [5]. Lewis *et al* use an entropy priority map for segmentation of regions for fusion [6]. In segmentation based methods, primarily high frequency information in the area of DOF in each input image is used. When the DOF is narrow, edge correspondence between the input images are not the same, due to the optics of the system. Therefore, segmentation based on physical object boundaries becomes ambiguous. In methods using tiling, the most widely reported issues are from blocking effects [3].

In MSD based methods, many of the fusion regulations used, rely on pixel manipulation or replacement at a detail level and, these results in changes in the intensity values of the fused image. These effects are not very prominent to the end user when using simple datasets but stands out as an issue in accurate scene inspection. Learning based methods use training engines which learn to classify between sharp and blurred regions and are normally computationally expensive [5,8]. Training is normally done with prescribed focused and unfocussed training data sets. In the advent of region that is blurred in all the input images, i.e. unseen data, misclassification takes place and learning based methods employ averaging or force one arbitrary region as the fused image. Ringing effects have been widely reported. Furthermore, methods employing wavelets for a particular application may not be extendable to another application as it is difficult to realize a wavelet kernel that can handle multiscale datasets [3]. In our method, we base our segmentation on focal connectivity and not on object connectivity in the scene. There is no pixel manipulation and hence no artifacts such as ringing effects, etc. Our method is computationally straightforward and contains intelligence to choose between sharp and blurred regions.

3. FUSION BY ESTABLISHING FOCAL CONNECTIVITY

In segmentation based methods, segmentation is performed using physical edges of objects in the scene. Due to camera optics, an edge in one input image blurs in another. This makes segmentation based on edges ambiguous. To counter this problem, we propose segmentation of regions based on focal connectivity. Each image in a set of input images has certain regions of the scene in focus. A focally connected region is a region or a set of regions in an input image that fall on the same focal plane. These regions may be connected focally with or without physical continuities in object geometries. The central idea of our method is to isolate and attribute such partitions to one particular input image and synthesize the fused image. In our method, a sharpness map is calculated for every input image $I_{\{i\}}(x,y) \ \forall \ i = 1,2,\dots, N$. As a precursor to this step, the images are filtered with sobel masks to approximate horizontal and vertical gradients, $I_{x\{i\}}(x,y)$ and $I_{y\{i\}}(x,y)$

respectively, where the subscripts x and y denote directional gradient operations. These are used to calculate the sharpness maps $S_i(x,y)$'s for each of the N input images by, isolate and attribute such partitions to one particular input image. The chosen partition is in better focus than its relative counterparts from all the input images. We isolate partitions in the input images based on focal connectivity

$$S_{\{i\}}(x,y) = \sqrt{I_{x\{i\}}^2(x,y) + I_{y\{i\}}^2(x,y)} \ \ \forall \ i \in \{1,2,\dots,N\}. \quad (1)$$

To make the system less vulnerable to fluctuations dependent on sensor (e.g. noise), optics (e.g. magnification and side lobes), local contrast and illumination at the scene, we low pass filter the sharpness maps. This increases the accuracy of the decisions to follow by ensuring that areas with better focus influence the decision of its neighbors. These sharpness maps are examined for regions of higher focus with their respective counterparts. When the sharpness map of input image $I_{\{i\}}(x,y)$, of N input images, is compared with its N-1 counterparts, one focally linked region, $P_{\{i\}}(x,y)$ is isolated by,

$$\begin{aligned} P_{\{i\}}(x,y) &= S_{\{i\}}(x,y) > S_{\{k \neq i\}}(x,y) \\ &\forall \ i \in \{1,2,\dots,N\}, \forall k \in \{1,2,\dots,N\} \end{aligned} \quad (2)$$

The union of the such partitions, $P_{\{i\}}(x,y)$'s, form the fused image space, $F_N(x,y)$ and the intersection of the partitions is the null set, corresponding to blurred sections in all the input images. A corresponding mask is created and a pixel wise multiplication isolates the image partitions. The image partitions are then seamlessly mosaiced to form,

$$F_N(x,y) = \bigcup_{i=1}^{N} \left\{ P_{\{i\}}(x,y) \times I_{\{i\}}(x,y) \right\}, \quad (3)$$

the fused image, where $\times$ denotes pixel-wise multiplication. There is no duplication in the partitions and therefore no blending at the peripheries of the partitions is required. This allows us to capitalize on focal overlaps by selecting the areas under best focus, out of adjacent areas, and between multiple images. Our system is able to choose between blurred regions and offer the least blurred region for fusion. When a region falls under overlapping DOF of multiple images, we use its sharpness map in a dynamic system that serves to populate the corresponding region with the area of best focus. This is an useful property that surpasses learning based methods in that, it is able to handle unseen data.

4. EXPERIMENTAL RESULTS

While imaging a 3D scene, the size of the objects and their relative positions in the scene defines the complexity of multifocus fusion. If the objects are placed well apart such that there are no focal overlaps, the problem of multifocus
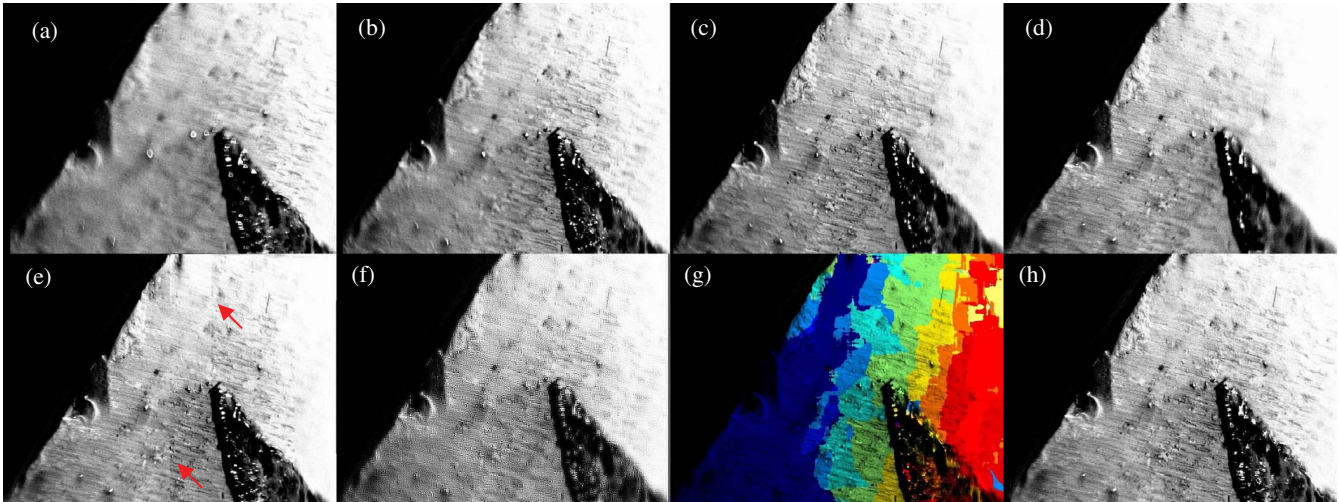
Figure 2. Comparison of fusion methods, (a-d) Raw input images, (e) Fusion by tiling (notice block effects shown by arrows), (f) Fusion using MSD based fusion [4] (g) color coded focally connected partitions from our partitioning scheme, and (h) multifocus fused image using our method. Due to space considerations, full resolution images of the test images are provided for inspection at: http://www.ece.utk.edu/~hharihar/fusion.htm. Active zoom has been enabled for close inspection in the electronic versions of this paper.

fusion becomes relatively easier. If the DOF is very narrow and the scene is imaged with many individual frames, with focal overlap the fusion algorithm requires more intelligence and finesse. It is possible to image an object at adjacent planes that appears to have common regions in focus. This is a prominent problem in microscopic imaging. In such a scenario, block based methods start exhibiting boundary issues more prominently and an a priori selection of the block sizes becomes more complicated. Data collected for this effort vary from microscopic to long range (>30 feet apart) data sets. Multifocus fusion is a significant problem in microscopic applications and we present an example (a microscopic scene of a damaged drill bit) in Figure 2. Acquisition conditions were chosen such that the DOF is very narrow and there is focal overlap between adjacent frames. Multiple shots of the scene are required to gather information from all the focal planes, especially given the extremely narrow depth of focus. In Figure 2 (a-d), we show a few images from the set of input images acquired. In Figure 2 (e), the results of the seminal tiling approach are presented. The input images were divided into sets of blocks and one block per set was chosen based on the Tenengrad sharpness criterion [9]. Block sizes were chosen empirically and the fused image with the highest overall Tenengrad measure was chosen to establish a fair comparison against our method. While a reasonable understanding on the scene can be obtained, there are visible blocking effects. In Figure 2 (f), results from a widely used MSD based fusion method are presented, following [4]. This method was chosen as it was designed for fusion with multiple frames, and uses the coiflet wavelet (level 2) family, reported in the literature repeatedly as one suitable for multifocus fusion [4]. In Figure 2 (g), image partitioning of the fused image space

is shown using a color coding scheme. Each color coded section in this image is one focally connected area in an image and represents areas from 1 input image. In Figure 2 (h), we show results from our work, wherein regions selected from different focally connected partitions of the input images are mosaiced seamlessly to synthesize a multifocus fused image. Our method is suitable for applications where the illumination is poor and DOF is very narrow.

In Figure 3, we show input images acquired of 2 individuals standing about 30 feet apart in a long range application. This is a case where focusing both the subjects s difficult even with an extremely narrow aperture and extended illumination. In Figure 3 (a-b), input images are shown where one individual is blurred in each of the images. In Figure 3 (c), an image fused used the seminal tiling approach is shown. Window selection was extremely difficult in this case due to the scale of the subjects in the scene. An optimum window size was selected based on trials with different window sizes and computed Tenengrad measures. In Figure 3(d), fusion by MSD based fusion [4] is presented. The fusion shows sufficient detail from both focal planes. Upon close inspection ringing effects are visible (marked by arrows). In Figure 3 (e), an image fused using our method is shown.

Our method is able to select partitions regardless of the scale differences of the input images, consistently over other methods implemented for comparison. In our experiments, apart from subjective evaluations we have also performed objective evaluations of the fused images. The fused outputs of the various methods were evaluated for sharpness using the Tenengrad measure [9]. The Tenengrad measure is given by,

$$T = \sum_{i=1}^{m} \sum_{j=1}^{n} \sqrt{F_x{}^2(x,y) + F_y^2(x,y)} , \qquad (4)$$

where, F is the fused image whose quality of fusion is to be evaluated, m×n is the total number of pixels in F, and x and y denote directional operations. This metric has been recorded as being optimal for sharpness evaluation by Krotkov [9].

<center>(a)   (b)   (c)</center>

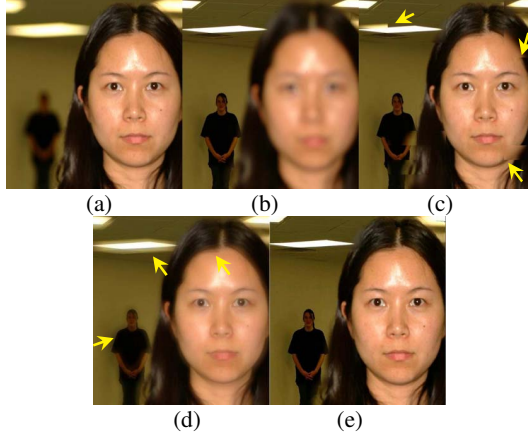<center>(d)   (e)</center>

Figure 3. Demonstrations of multi focus fusion in a long range application wherein each focal plane is about 30 feet apart from each other. (a) Individual in foreground is in focus, (b) Individual in background is in focus, (c) Fusion using tiling (block effects shown with arrows), (d) Fusion using MSD based fusion (ringing effects shown with arrows) [4] and (e) Fusion using our method. Notice both individuals are in focus with no blocking or ringing effects in the fused image. Active zoom has been enabled for close inspection in the electronic versions of this paper.

The objective results are consistent with visual inspection and show that our method produces images with better overall sharpness. The results of the objective testing are summarized in Table 1 wherein it is shown that the images fused using our method have the most measured sharpness when compared against the other methods used for comparison.

Table 1: Objective evaluation of the output images of fusion methods (using the Tenengrad Metric [9]) studied in this paper

	Tiling	MSD	Proposed Method
Microscopic	8.67 E+07	5.31 E+07	9.70 E+07
Long Range	1.14 E+09	0.77 E+09	2.05 E+09

5. CONCLUSIONS AND FUTURE WORK

A multifocus fusion scheme where regions are segmented from input image frames using the virtue of focal connectivity is presented. Images are partitioned based on focal connectivity and each partition is attributed to one input image. This method capitalizes on redundant focal information and narrow DOF to synthesize a fused image that appears to have an extended DOF while retaining the visual verisimilitude of the scene. This method has been found resilient to noise and the fused images exhibit focal information from all possible focal planes imaged. Rather than test them on examples with isolated objects in the scene, tests were done on challenging scenes where the focal planes are adjacent to each and simple visual inspection is insufficient during acquisition to determine planes of focus. Additionally, long range datasets have been used to demonstrate the versatility of our technique. Experimental results and a comparison with other methods have been shown.

7. REFERENCES

[1] D. Fedorov, B. Sumengen, and B. S. Manjunath, "Multi-focus imaging using local focus estimation and mosaicking," in *IEEE International Conference on Image Processing*, Atlanta, Georgia, 2006, pp. 227-231.

[2] A. Goshtasby, "Fusion of multi-focus images to maximize image information," *Defense and Security Symposium*, Orlando, Florida, 2006, pp. 17-21.

[3] I. De and B. Chanda, "A simple and efficient algorithm for multi-focus image fusion using morphological wavelets," *Signal Processing*, vol. 86, pp. 924-936, 2006.

[4] S. Frechette and V. K. Ingle, "Gradient based multi-focus video image fusion," in *Proceedings of IEEE Conference on Advanced Video and Signal Based Surveillance*, 2005, pp. 486-492.

[5] Z. W. Liao, S. X. Hu, and Y. Y. Tang, "Region-based multi-focus image fusion based on hough transform and wavelet domain hidden markov models," in *Proceedings of International Conference on Machine Learning and Cybernetics*, vol. 9, 2005, pp. 5490-5495.

[6] J. J. Lewis, R. J. O'Callaghan, S. G. Nikolov, D. R. Bull, and C. N. Canagarajah, "Region-based image fusion using complex wavelets," in *Proceedings of the Seventh International Conference on Information Fusion*, vol. 1, 2004, pp. 555–562.

[7] M. B. Ahmad and T-S. Choi, "A heuristic approach for finding best focused shape," in *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 15, no.4, pp. 566–574, 2005.

[8] L. Shutao, T. K. James, and W. Yaonan, "Multifocus image fusion using artificial neural networks," in *Proceedings of International Conference on Machine Learning and Cybernetics*, 2005, pp. 985 – 997.

[9] E. P. Krotkov, *Active Computer Vision by Cooperative Focusing*. USA: Springer Verlag, 1989.