Bottom-up Segmentation for Ghost-free Reconstruction of a Dynamic Scene from Multi-Exposure Images

Shanmuganathan Raman*
Vision and Image Processing Laboratory
Indian Institute of Technology Bombay
Mumbai, India
shanmuga@ee.iitb.ac.in

Subhasis Chaudhuri†
Vision and Image Processing Laboratory
Indian Institute of Technology Bombay
Mumbai, India
sc@ee.iitb.ac.in

ABSTRACT

High Dynamic Range (HDR) imaging requires one to composite multiple differently exposed images of a scene in the irradiance domain and perform tone mapping of the generated HDR image for displaying on Low Dynamic Range (LDR) devices. In the case of dynamic scenes, standard techniques may introduce artifacts called ghosts if the scene changes are not accounted for. In this paper, we consider the HDR problem for dynamic scenes. We develop a novel bottom-up segmentation algorithm through superpixel grouping which would enable us to detect scene changes. We then employ a piecewise patch-based compositing methodology to directly generate the ghost-free LDR image of the dynamic scene. The primary advantage of our approach is that we do not assume any knowledge of camera response function and exposure settings. Further, our approach performs well even in the case of significant scene changes.

Keywords

High Dynamic Range Imaging, Deghosting, Computational Photography

1. INTRODUCTION

Computational Photography aims at circumventing the restrictions of the common digital cameras using computational techniques. One of the major problems of the common digital cameras is the limited dynamic range due to the limited capacity of the digital sensors. Capturing the entire dynamic range of the scene is the primary goal of any digital imaging system. This would enable us to visualize the scene with the highest level of contrast. Consider a scene with both brightly and poorly illuminated regions. Such a scene has a very high dynamic range. All the brightness levels of the scene cannot be captured using a single snapshot of common digital cameras. The general approach is to capture multiple images of the scene with different exposure settings and combine them to generate the desired image of the scene.

As most of the real world scenes have a higher dynamic range than what can be captured using a digital camera, it is required to capture multiple images of the scene with different exposure settings. These images would together then span the entire dynamic range of the scene. The imaging methodology meant for combining these multiple, differently exposed images into a single image is popularly known as High Dynamic Range (HDR) imaging. We require the knowledge of camera response function (CRF) which relates irradiance and intensity values, to generate the HDR image. The generated HDR image needs to be tone mapped to a low dynamic range (LDR) image for compatibility with common displays and printers [29].

Most real world scenes are dynamic. While capturing multiple images of a scene, one does not have control over the movement of objects in the scene. If the changes in the scene are not detected before compositing multi-exposure images, the generated LDR image would have artifacts called ghosts. It is imperative that we need to detect any scene changes across these multi-exposure images to prevent ghosts from appearing in the generated LDR image. In this work, we address the problem of generating an LDR image of a dynamic scene directly from a set of multi-exposure images. Our contributions are to develop a robust algorithm for detecting scene changes and to compose different regions seamlessly.

We develop a novel bottom-up segmentation algorithm based on superpixel grouping [30] to segment out the scene changes. A characteristic function between a given pair of observations with different exposures enables us to identify decision regions for grouping the superpixels which belong to the foreground (scene changes). After detecting the regions of the image which show change with respect to a reference image, we compose the multi-exposure images to generate the LDR image without any ghosts. The primary advantage of our approach is that we do not assume any knowledge of the scene and camera settings. Further, we show that seamless LDR image can be generated even when there is an appreciable scene change across the multi-exposure images.

To start with, we shall provide an overview of the existing literature on HDR imaging and segmentation. We discuss the problem of reconstructing an ghost-free LDR image for a dynamic scene without the knowledge of CRF and exposure settings. We shall then look at the proposed approach for generating an LDR image from a set of multi-exposure images of a dynamic scene in detail. We shall then present the results of the proposed approach for a few dynamic scenes. We show that state-of-the-art LDR images can be generated using the proposed approach.

*Corresponding author, Thanks to Microsoft Research India, Bangalore

†Thanks to JC Bose Fellowship and Bharti Center for Communication

ICVGIP '10, December 12-15, 2010, Chennai, India

2. RELATED WORK

Capturing the entire dynamic range of a scene using a single image is a challenging problem in computational photography. Analog cameras can capture such scenes with very high dynamic range using a single snapshot. The common digital cameras are not capable of doing so due to the limited dynamic range of the sensor. However, one can capture multiple, differently exposed images of the scene and composite them in the irradiance domain and capture the entire dynamic range [29]. There are digital cameras which can capture the entire dynamic range of the scene using a single snapshot [24]. These cameras are very expensive in market. In this section, we would address some of the works done earlier in HDR imaging along with an overview of the segmentation techniques.

Mann and Picard introduced a method to recover the camera response function (CRF) and estimate the HDR image from multi-exposure images [19]. They used the derivative of the CRF, called certainty function, to weigh the multi-exposure irradiances. Debevec and Malik developed a practical algorithm for recovering the HDR image and used a simple hat function as the weighting function [7]. Mitsunaga and Nayar estimated the CRF by parameterizing it and then generated the HDR image [21]. CRF can also be recovered from a single image provided the image has edges of highly varying magnitudes [17]. An overview of all these different methods for the generation of HDR image from multi-exposure images of a static scene can be found in [29]. Assuming CRF is known, Granados *et al* recently developed a method to generate an HDR image in the presence of various types of noise [12].

The generated HDR image can then be encoded in Radiance RGBE (.hdr) or OpenEXR (.exr) formats which employ floating point numbers to store the intensity values. These image formats require a large amount of memory and need to be compressed for optimal storage. The methods discussed above for the generation of HDR image have to be complemented by an appropriate tone mapping operator for visualization in common displays and printers. An example of tone mapping is the gradient domain HDR compression method by Fattal *et al* [8]. An overview of the different types of tone mapping operators (global and local) can be found in [29].

For a static scene, there are alternate methods based on digital compositing which combine the multi-exposure images directly without the knowledge of CRF. These methods employ basic digital compositing principles with an appropriate weighting function. The basics of digital compositing methodology can be found in ([3], [6], [26], [33]). An interactive method for compositing image regions was proposed by Agarwala *et al* [1].The method by Goshtasby uses entropy measures on blocks to combine multi-exposure images [11]. Exposure fusion combines multi-exposure images on a Laplacian pyramid using an appropriate weighting function [20]. A variational, iterative solution for combining multi-exposure images was introduced in [27]. Bilateral filter was used to define weighting function and composite multi-exposure images in [28].

While capturing multi-exposure images of a scene, we cannot guarantee that the scene would not change. There are chances of new objects being introduced in the scene between the exposures due to motion. Also, objects such as leaves and branches of a tree would move when there is the presence of heavy wind in the scene. In other words, the scene would most probably be dynamic. When the methods mentioned above are employed for compositing multi-exposure images of a dynamic scene, the objects in motion in the scene would give rise to artifacts called ghosts. It is required that the changes in the scene are detected apriori before compositing is performed on the multi-exposure images. We shall first look at the methods previously used for removing ghosting artifacts.

Jacobs *et al.* proposed a method to identify the regions on the image grid which change across multi-exposure images using weighted variance and entropy measures [15]. This method fills the motion regions by details from one of the observations and thereby reducing contrast in such regions. Gallo *et al* proposed a method to detect motion regions in multi-exposure images when CRF is known and eliminate them while compositing [10]. This approach preserves contrast in the motion regions as regions from multiple images are combined. Another approach for eliminating ghosting artifacts while creating mosaic from images of different exposures was proposed in [34]. In this work, we address this problem from a bottom-up segmentation perspective.

A recent work assumed no knowledge of CRF and reconstructed the dynamic scene as an LDR image from multi-exposure images [36]. This method employs the gradient directions for detecting scene change and may perform poorly in the presence of even a small amount of noise of any form. Further, this approach cannot handle scenes which have tiny objects such as leaves of a tree in motion. Another problem is that this approach requires a number of parameters to be adjusted empirically.

Segmenting foreground from background in a single image is a classic vision problem. The segmentation algorithm can either be automatic or be interactive. One popular approach to achieve interactive segmentation is the Grabcut by Rother *et al* [31]. Interactive segmentation depends on the user input such as a bounding box or scribbles to perform segmentation. Another class of interactive algorithms which extract the alpha matte along with the foreground mask is known as matting. An example is the natural image matting [16]. We focus more on the automatic segmentation approaches in this work as we intend to develop an automatic method to compensate for scene change across different exposures.

Automatic segmentation approaches can be broadly classified into top-down and bottom-up methods. In the top-down approach, one tries to capture the entire object boundary directly by the learned features of the desired object class ([4], [5]). This approach tries to get the foreground separated from the background as a whole. While in the bottom-up approach, the entire image is split into homogeneous regions based on color, contours, and texture details. These homogeneous regions are then grouped to segment the foreground from the background [32]. The bottom-up segmentation methods have derived much interest among the computer vision community of late as they lead to better segmentation of the foreground objects.

In this paper, we specifically focus on the segmentation based on the bottom-up approach. The algorithm by Shi and Malik uses normalized cuts to split a given image into multiple regions which are homogeneous [32]. This approach was later extended to define different homogeneous regions of the image as superpixels [30]. Each superpixel is a collection of a set of pixels inside a closed contour signifying uniformity in terms of color and texture. Object recognition systems can then work on the level of superpixels instead of image pixels which can help in designing faster algorithms. For segmentation tasks, one can group the superpixels belonging to foreground object based on some criteria to segment out the foreground ([23], [22]). Even neighborhood can be defined for superpixels to improve the segmentation [9]. Another approach to extract homogeneous regions from an image is the quick shift algorithm by Vedaldi and Soatto [35].

A typical bottom-up approach for segmentation relies on the efficient grouping of the superpixels and recovering the foreground. Apart from basic segmentation, grouping of superpixels have a number of applications in computer vision. Consider the problem of estimating the depth map of a scene from a single image. Superpixels corresponding to objects at different depths of the scene can be grouped to recover the 3D information from a single image [14]. In the present work, we apply superpixel grouping for detecting scene change in the multi-exposure images of a dynamic scene. We assume one of the multi-exposure images as the reference and employ grouping of superpixels to recover the scene change in the other multi-exposure images.

3. PROPOSED APPROACH

In this section, we shall discuss the proposed approach for the generation of an artifact-free LDR image corresponding to a dynamic scene. We assume that we do not have the knowledge of CRF of the camera and the exposure settings. Given a set of multi-exposure images, our task is to identify the regions which have moving objects in each of the images and eliminate them while compositing. The proposed approach is shown in Figure 1. The salient feature of the proposed approach is that we composite multiple images even in regions which show motion, thereby preserving overall contrast of the scene in the generated LDR image. We shall first estimate the decision boundaries to classify dynamic and static regions and use that to reconstruct the ghost-free LDR image of the dynamic scene.

3.1 Weighted Variance Measure

Consider a set of multi-exposure images corresponding to a dynamic scene shown in Figure 2. These images are taken at different times of the day with different exposure settings. These images together are sufficient to recover the entire dynamic range of the scene. However the scene changes appreciably in the 2(b) and 2(d) due to the movement of people in the scene. When CRF is known, we can recover the HDR equivalent of the scene using the technique mentioned in [10]. In the absence of an accurate estimate of the CRF corresponding to the camera used, we need to figure out pixel locations which do not change in any of the multi-exposure images. The intensity values of the multi-exposure images corresponding to these pixel locations is used later to detect changes in the scene.

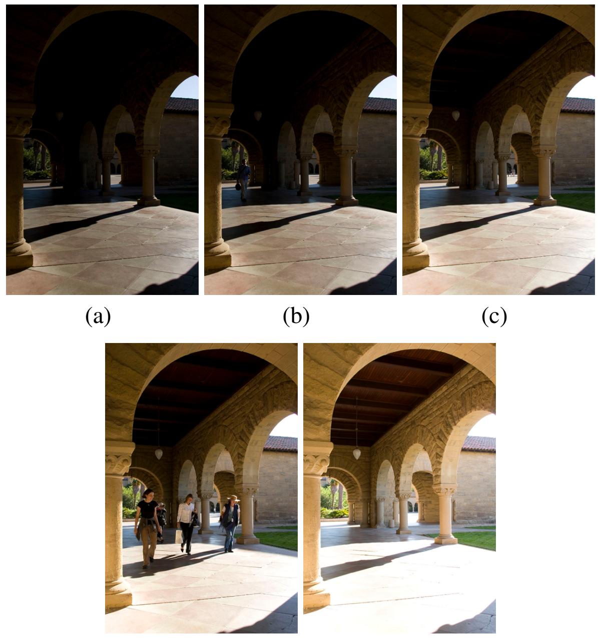
We employ the approach suggested in ([15], [29]) based on weighted variance to detect the pixel locations which do not have scene change in any of the multi-exposure images. The weighted variance measure $V(x,y)$ can be computed for K differently exposed images $I_i(x,y)$using the Equation 1.

$$V(x,y) = \frac{\dfrac{\sum_{i=1}^{K} w_i(x,y) I_i^2(x,y)}{\sum_{i=1}^{K} w_i(x,y)}}{\dfrac{\left(\sum_{i=1}^{K} w_i(x,y) I_i(x,y)\right)^2}{\left(\sum_{i=1}^{K} w_i(x,y)\right)^2}} - 1 \tag{1}$$

where the weight is given by the Gaussian function

$$w_i(x,y) = e^{-\frac{(I_i(x,y)-\mu)^2}{2\sigma^2}} \tag{2}$$

where $\mu = 0.5$ and $\sigma = 0.2$ respectively. This Gaussian function $w_i(x,y)$ is used as weight in order to provide lesser weight to the over-exposed and the under-exposed pixel locations. We use an appropriate threshold (0.25 times the maximum weighted variance) to detect pixel locations which show low weighted variance measure. In the case of noisy multi-exposure images, simple Gaussian spatial smoothing can be used prior to the computation of weighted variance.



(a) (b) (c) (d) (e)

Figure 2: (a-e) Multi-Exposure images of a dynamic scene. Images Courtesy: Orazio Gallo, UCSC.

3.2 Estimation of the Intensity Mapping Functions

Consider any two observations of a scene which differ only in the exposure setting. The intensity values of these two images can be characterized by a function called comparametric function [18]. The comparametric function is also referred to as intensity mapping function (IMF) [13]. We use the term IMF to refer to this function henceforth. This function defines how the intensity values of two images of a static scene should relate when there is only a difference of exposure settings. IMF is a non-linear function whose slope can be computed and this would be greater than 1 if the exposure time of one image is greater than that of the reference image.

The weighted variance measure provides us the pixel locations of the scene where there are no appreciable changes in any of the multi-exposure images. Now, we would be able to estimate a unique IMF for a given pair of images using the intensity values at these pixel locations. Let $S \in \Re^2$ be the set of all pixel locations in the image grid. Now, the pixel locations where there are no motion in any of the images is given by the set $\Omega \subset S$. As a generality, we select one of the multi-exposure images as representing the static scene. We would now estimate IMFs between the intensity values of this image and the rest of the images in Ω. Given a set of K multi-exposure images, we would have a total of $(K-1)$ IMFs with respect to the reference image.

It is worth noting that the normalized intensity values of the multi-exposure images are in the range $[0,1]$. We need to fit a polynomial of order four (chosen empirically) in order to estimate IMF. The pixel locations of the multi-exposure images in Ω should fol-

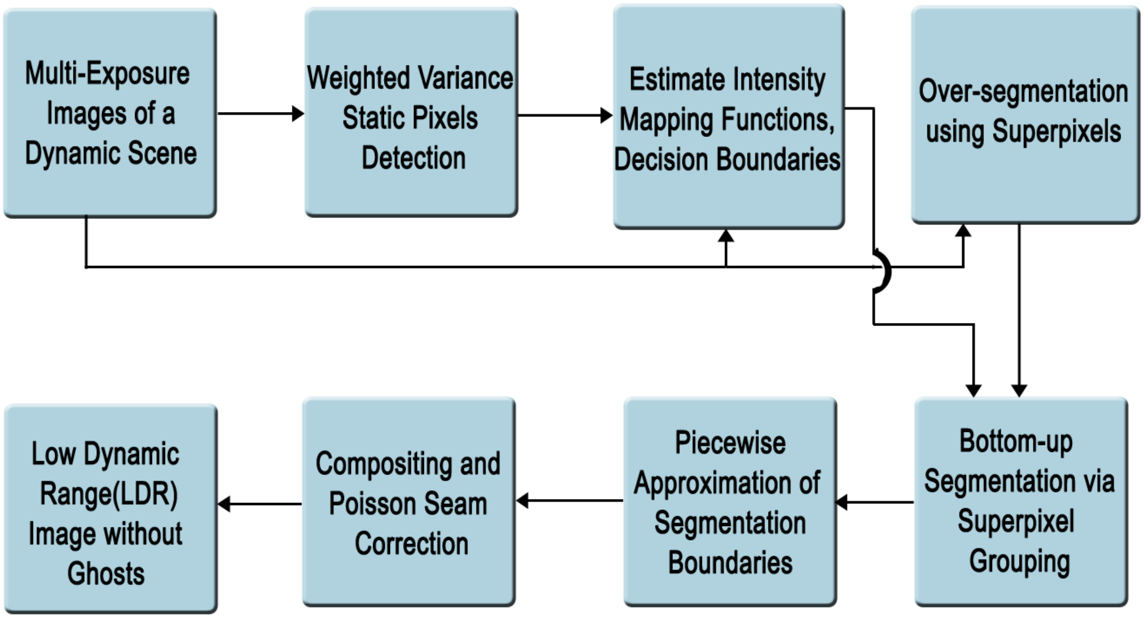
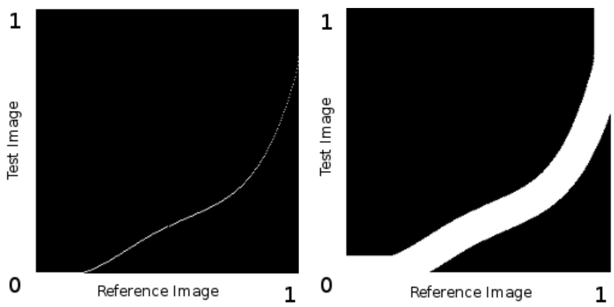


Figure 1: Schematic Representation of the Proposed Approach

low this IMF with respect to the reference image for them to be classified as static. The pixel locations which have some appreciable scene change from the reference would not follow this IMF. The estimated IMF between the images Figure 2(b) and Figure 2(e) (reference) is shown in Figure 3 (a).



(a) (b)

Figure 3: (a) The IMF between a pair of images 2(b) and 2(e) and (b) The constant width region which defines the decision boundary for static and dynamic regions

3.3 Superpixel Grouping

Having estimated the IMFs for each of the multi-exposure images, we need to define a constant width region around the IMF. This constant width region would represent the pixel locations of the test image which do not have any appreciable change with respect to the reference as shown in Figure 3 (b). This constant width region provides us with a decision boundary between static and dynamic regions of the given multi-exposure image with respect to the reference. We now propagate the decisions from pixel level to region level. We exploit over-segmentation of images using superpixels to recover regions of images which have homogeneous color and texture.

We compute superpixels on all multi-exposure images including the reference image [30]. We use superpixels as it saves us from the huge load of classifying each pixel and grouping them. Further, it lets take care of the object boundaries while segmentation which is not possible when patches are used. The superpixels corresponding to the images shown in Figure 2 are shown in Figure 4. As can be observed in Figure 4(e), the superpixels do not cross the object boundaries and grouping them would enable us to recover the exact silhoutte of the scene change.

Instead of classifying every pixel, we classify the superpixels for the possible scene change with respect to the reference. Given the reference image and any other multi-exposure image, we now find out the fraction of pixels in a given superpixel which lie inside the constant width region (0.12 in this work). We define a parameter γ which defines the minimum fraction of pixels which should be present inside the constant width region for the superpixel to be classified as having no change. We used γ to be equal to 0.9 in our experiments. We classify all the superpixels corresponding to the given image with respect to the reference as either dynamic or static.

This operation would effectively let us group all the superpixels of the given image which show change with respect to the reference image. This is the novel bottom-up segmentation algorithm developed for multi-exposure images using the estimated IMF for the computation of decision boundaries. The bottom-up segmentation which has been performed on the multi-exposure images is shown in Figure 5. One can clearly see that the proposed algorithm for segmentation is able to group all the superpixels which convey an appreciable scene change. We need to ignore these regions while compositing multi-exposure images in order to avoid ghosting artifacts.

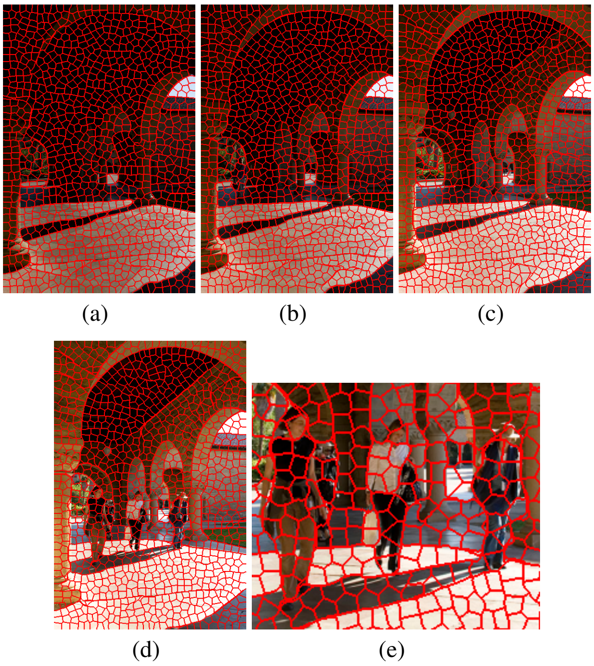


Figure 4: (a-d) Superpixels estimated corresponding to the first four multi-exposure images in Figure 2, and (e) Magnified superpixels corresponding to a region of (d).

3.4 Scene Reconstruction

Having detected the scene changes with respect to the reference image, we need to develop a method to reconstruct the final LDR image of the scene. The LDR image needs to be generated from the reference image and the superpixels from other images marked as having no scene change (static). One problem with segmentation by grouping superpixels is that it leads to irregular segmentation boundaries which are difficult to account for while reconstruction. One cannot guarantee that these grouped regions would be closed as evident in Figure 5(d). This would lead us into trouble while defining boundary conditions.

We split the images into patches of size (say 6 X 6). We detect patches of the images which have more than 90 percent of the pixels lying inside the static superpixels. These patches are classified as not having appreciable motion. Such an operation would result in piece-wise rectangular approximation of the bottom-up segmentation boundaries as visible in Figure 6. We can now use any of the static multi-exposure compositing algorithms on the patches marked as static ([20], [27], [28]). This compositing operation would result in seams across patch boundaries as different number of patches are combined at each patch location.

To avoid these seams, we use overlapping patches (of size 6 X 6) with an overlap of one pixel in each direction. This would ensure that there is information conveyed across the patches. In the present work, we perform multi-exposure compositing using exposure fusion [20] on these overlapping patches while ignoring the patches marked as dynamic. The composited patches would still produce seams on their boundaries if arranged spatially on the image grid S.

The seams across the patch boundaries can be avoided using an operation called Poisson seam correction. The gradients cor-

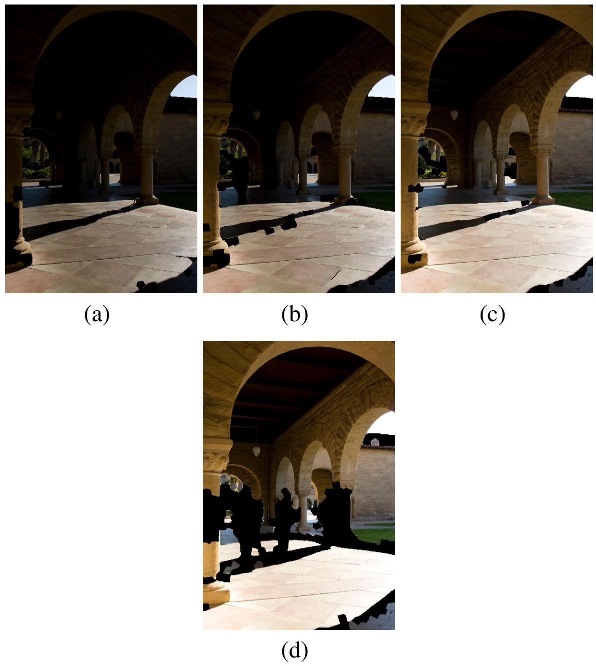


Figure 5: (a-d) Bottom up segmentation through superpixel grouping performed on the images in Figure 4.

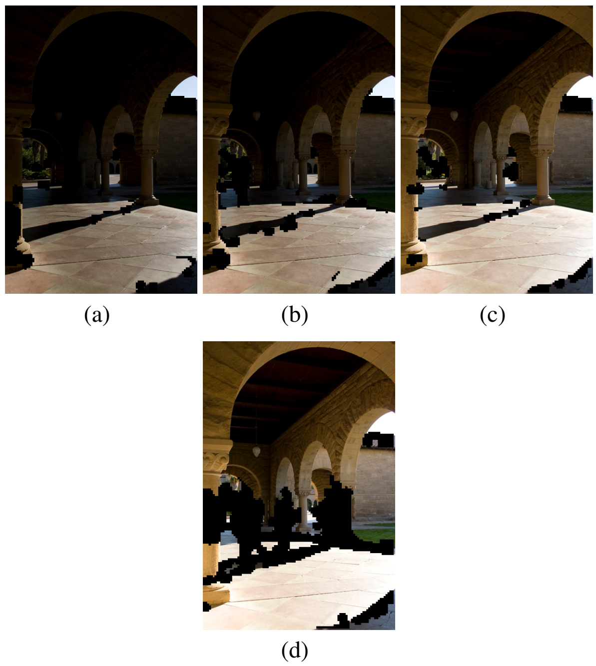


Figure 6: (a-d) Piece-wise rectangular approximation of the segmentation boundaries shown in Figure 5.

responding to the composited patches (of size 5 X 5) are computed and arranged on the image grid S. The resultant vector field will not be conservative. We employ a direct Poisson solver with Neumann boundary conditions to generate the scalar field closest to the vector field ([2], [25]). The scalar field obtained through Poisson seam correction operation is the desired LDR image without any artifacts.

4. RESULTS

In this section, we consider multi-exposure images of a dynamic scene. We shall present the LDR images generated using the proposed approach and compare the results with that of the tone-mapped LDR image obtained using the method of [10]. It is worth noting again that the proposed approach does not require the knowledge of CRF and the exposure settings. Further, we do not explicitly generate the HDR image of the scene and therefore do not require tone reproduction. These are the key advantages of the proposed approach over that of [10].

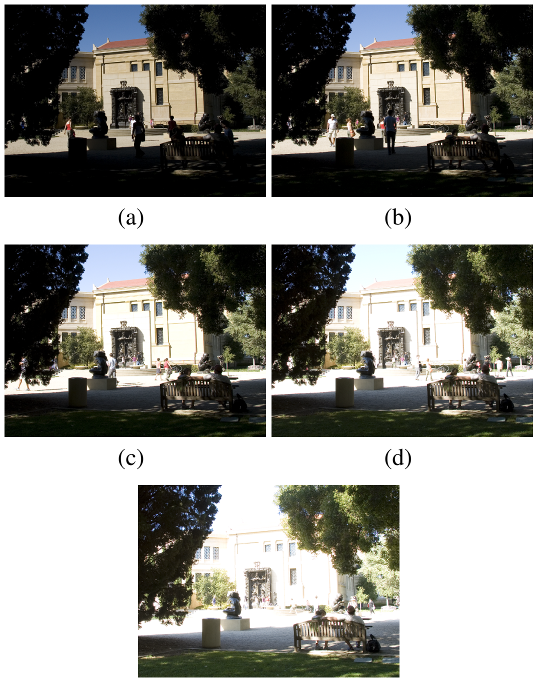
Consider the multi-exposure images of a dynamic scene in Figure 2. Figure 2(e) is picked as the reference image. Figure 9(a) shows the result of an existing approach [28]. As scene change is not accounted for, ghosting artifacts are clearly visible in the generated LDR image. The tone mapped image corresponding to the HDR image generated by Gallo *et al* is as shown in Figure 9(b). The LDR image generated using the proposed approach is shown in Figure 9(c). One can observe that the proposed approach is able to generate an artifact-free LDR image from a set of multi-exposure images. The details in both the brightly and poorly illuminated regions are clearly visible.

Consider another set of multi-exposure images of a dynamic scene as shown in Figure 7. This scene is complex in the sense that there are many people moving in and out of the scene across these images. Further there is a sun-lit region (brightly illuminated) and tree shade (poorly illuminated). As expected, common digital cameras cannot capture the entire dynamic range of this scene with varied levels of brightness. We assume the image in Figure 7(c) to be the reference image. We detect scene changes on the other images with respect to this image. Compositing and Poisson seam correction enables us to generate the LDR image shown in Figure 10(b). Figure 10(a) shows the tone mapped LDR image of [10]. One can visualize that the proposed approach is able to capture the details in both brightly and poorly lit regions without any ghosting artifacts.

Let us consider another dynamic scene captured using differently exposed images shown in Figure 8. This scene has the branches of the tree moving and there is a person introduced in the scene while capturing the last image (Figure 8(d)). This is an example of a scene where there is a significant motion in the majority of the pixel locations. We pick the image Figure 8(b) as the reference image. The reconstructed LDR image using the proposed approach is shown in Figure 11(b). The tone mapped result of [10] is shown in Figure 11(a). We can see that the proposed approach is able to reconstruct the scene without any artifacts even in the case of significant motion of objects across the multi-exposure images.

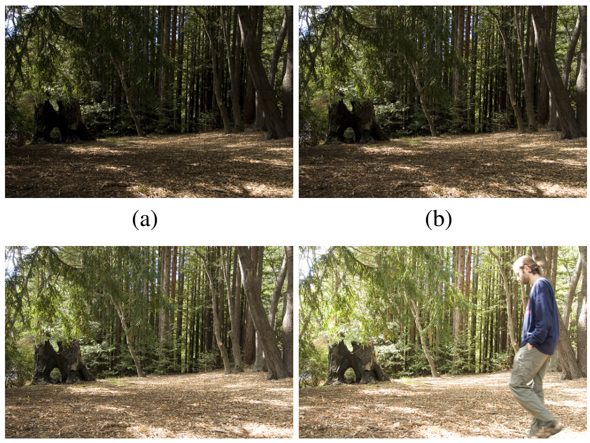
5. CONCLUSION

We have proposed a novel bottom-up motion segmentation approach for detecting motion in multi-exposure images corresponding to a dynamic scene. We then approximated the segmentation boundaries and reconstructed the dynamic scene with an LDR image without any artifacts. The proposed approach is a quite useful tool in digital photography where photographers like to use multiple differently exposed images to capture a dynamic natural scene.



(a) (b) (c) (d) (e)

Figure 7: (a-e) Multi-Exposure images of a dynamic scene. Images Courtesy: Orazio Gallo, UCSC.



(a) (b) (c) (d)

Figure 8: (a-d) Multi-Exposure images of a dynamic scene. Images Courtesy: Orazio Gallo, UCSC.

The proposed approach has added advantages of not requiring either CRF or tone mapping. The generated LDR image is compatible with common displays and occupies lesser memory compared to the corresponding HDR image.

The proposed approach can either be included in the digital camera firmware or common image manipulation tools like Adobe Photoshop. The proposed approach would be a worthy alternative to the 'Merge to HDR'tool available in latest Photoshop releases (CS2 onwards). Also, the proposed approach works well even when there is a significant change in the scene while capturing multi-exposure images.

6. REFERENCES

[1] A. Agarwala, M. Dontcheva, M. Agrawala, S. Drucker, A. Colburn, B. Curless, D. Salesin, and M. Cohen. Interactive digital photomontage. In *SIGGRAPH*, 2004.

[2] A. Agrawal, R. Raskar, and R. Chellappa. What is the range of surface reconstructions from a gradient field? In *ECCV*, Graz, Austria, 2006.

[3] J. F. Blinn. Compositing, part 1: Theory. *IEEE Computer Graphics & Applications*, 14(5):83–87, 1994.

[4] E. Borenstein and S. Ullman. Class-specific, top-down segmentation. In *ECCV*, 2002.

[5] E. Borenstein and S. Ullman. Learning to segment. In *ECCV*, 2004.

[6] R. Brinkmann. *The Art and Science of Digital Compositing*. Morgan Kaufmann Publishers, 1999.

[7] P. Debevec and J. Malik. Recovering high dynamic range radiance maps from photographs. In *SIGGRAPH*, 1997.

[8] R. Fattal, D. Lischinski, and M. Werman. Gradient domain high dynamic range compression. In *SIGGRAPH*, pages 249–256, San Antonio, Texas, 2002.

[9] B. Fulkerson, A. Vedaldi, and S. Soatto. Class segmentation and object localization with superpixel neighborhoods. In *IEEE ICCV*, 2009.

[10] O. Gallo, N. Gelfand, W. Chen, M. Tico, and K. Pulli. Artifact-free high dynamic range imaging. In *ICCP*, 2009.

[11] A. Goshtasby. Fusion of multi-exposure images. *Image and Vision Computing*, 23:611–618, 2005.

[12] M. Granados, B. Ajdin, M. Wand, C. Theobalt, H. P. Seidel, and H. P. A. Lensch. Optimal hdr reconstruction with linear digital cameras. In *IEEE CVPR*, 2010.

[13] M. D. Grossberg and S. K. Nayar. Determining the camera response from images: What is knowable? *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25(11):1455–1467, 2003.

[14] D. Hoiem, A. A. Efros, and M. Hebert. Automatic photo pop-up. *ACM Trans. Graph.*, 24(3):577–584, 2005.

[15] K. Jacobs, C. Loscos, and G. Ward. Automatic high-dynamic range image generation for dynamic scenes. *IEEE Computer Graphics and Applications*, 28(2):84–93, 2008.

[16] A. Levin, D. Lischinski, and Y. Weiss. A closed form solution to natural image matting. In *IEEE CVPR*, 2006.

[17] S. Lin, J. Gu, S. Yamazaki, and H.-Y. Shum. Radiometric calibration from a single image. In *IEEE CVPR*, 2004.

[18] S. Mann. Comparametric equations with practical applications in quantigraphic image processing. *IEEE Transactions on Image Processing*, 9(8):1389–1406, 2000.

[19] S. Mann and R. W. Picard. On being undigital with digital cameras: extending dynamic range by combining exposed pictures. In *In Proc. of IS & T 48th annual conference*, pages 422–428, 1995.

[20] T. Mertens, J. Kautz, and F. V. Reeth. Exposure fusion: A simple and practical alternative to high dynamic range photography. *Computer Graphics Forum*, 28(1):161–171, 2009.

[21] T. Mitsunaga and S. K. Nayar. Radiometric self calibration. In *CVPR*, 1999.

[22] G. Mori. Guiding model search using segmentation. In *IEEE ICCV*, 2005.

[23] G. Mori, X. Ren, A. Efros, and J. Malik. Recovering human body configurations: Combining segmentation and recognition. In *IEEE CVPR*, 2004.

[24] S. Nayar and T. Mitsunaga. High Dynamic Range Imaging: Spatially Varying Pixel Exposures. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, volume 1, pages 472–479, Jun 2000.

[25] P. Perez, M. Gangnet, and A. Blake. Poisson image editing. In *SIGGRAPH*, 2003.

[26] T. Porter and T. Duff. Compositing digital images. In *SIGGRAPH*, pages 253–259, 1984.

[27] S. Raman and S. Chaudhuri. A matte-less, variational approach to automatic scene compositing. In *ICCV*, 2007.

[28] S. Raman and S. Chaudhuri. Bilateral filter based compositing for variable exposure photography. In *EUROGRAPHICS Short Papers*, 2009.

[29] E. Reinhard, G. Ward, S. Pattanaik, and P. Debevec. *High Dynamic Range Imaging: Acquisition, Display and Image-Based Lighting*. Morgan Kaufmann Publishers, 2005.

[30] X. Ren and J. Malik. Learning a classification model for segmentation. In *IEEE ICCV*, 2003.

[31] C. Rother, V. Kolmogorov, and A. Blake. Grabcut: interactive foreground extraction using iterated graph cuts. *ACM Trans. Graph.*, 23(3):309–314, 2004.

[32] J. Shi and J. Malik. Normalized cuts and image segmentation. *IEEE Trans on PAMI*, 22(8):888–905, 2000.

[33] R. Szeliski. Image alignment and stitching: A tutorial. *Foundations and Trends in Computer Graphics and Vision*, 2(1), 2008.

[34] M. Uyttendaele, A. Eden, and R. Szeliski. Eliminating ghosting and exposure artifacts in image mosaics. In *IEEE CVPR*, 2001.

[35] A. Vedaldi and S. Soatto. Quick shift and kernel methods for mode seeking. In *ECCV*, 2008.

[36] W. Zhang and W.-K. Cham. Gradient-directed composition of multi-exposure images. In *CVPR*, 2010.

(a) (b) (c)

Figure 9: (a) LDR image generated by multi-exposure compositing without motion detection showing ghosts [28], (b) Tone mapped LDR image using [10], and (c) LDR image generated using the proposed approach.

(a) (b)

Figure 10: (a) Tone mapped LDR image using [10], and (b) LDR image generated using the proposed approach.

(a) (b)

Figure 11: (a) Tone mapped LDR image using [10], and (b) LDR image generated using the proposed approach.