

Reconstruction of high contrast images for dynamic scenes

Shanmuganathan Raman · Subhasis Chaudhuri

Published online: 6 November 2011
© Springer-Verlag 2011

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 blind HDR problem for dynamic scenes. We develop a novel bottom-up segmentation algorithm through superpixel grouping which enables us to detect scene changes. We then employ a piecewise patch-based compositing methodology in the gradient domain to directly generate the ghost-free LDR image of the dynamic scene. Being a blind method, the primary advantage of our approach is that we do not assume any knowledge of camera response function and exposure settings while preserving the contrast even in the non-stationary regions of the scene. We compare the results of our approach for both static and dynamic scenes with that of the state-of-the-art techniques.

Keywords High dynamic range imaging · De-ghosting · Low dynamic range image generation · Computational photography

1 Introduction

Computational photography aims at circumventing the restrictions of the common digital cameras using either modification of camera internals or post-processing techniques. One of the major problems of 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 together 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 [45]. Estimating CRF accurately for most practical applications is a challenging task.

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

S. Raman (✉) · S. Chaudhuri
Department of Electrical Engineering,
Indian Institute of Technology, Bombay, Powai, Mumbai,
400076 India
e-mail: shanmuga@ee.iitb.ac.in

S. Chaudhuri
e-mail: sc@ee.iitb.ac.in

images, the generated LDR image is found to have artifacts called ghosts. It is imperative that we must 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 a high-contrast LDR image of a dynamic scene directly from a set of multi-exposure images. Our contribution lies in developing a robust algorithm for detecting scene changes and to compose different regions seamlessly. Additionally, we show that this can be achieved in the absence of any knowledge of both the CRF and the exposure times.

We develop a novel bottom-up segmentation algorithm based on superpixel grouping to segment out the changes in the scene. 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 a seamless LDR image can be generated even when there is an appreciable scene change across the multi-exposure images.

We shall now present some of the salient improvements over the work published earlier [44]. A novel general-purpose gradient domain solution for generating a low dynamic range (LDR) image from a set of multi-exposure images corresponding to both static and dynamic scenes is proposed. In [44], the approach was meant for dynamic scenes alone. The LDR image results of the proposed approach for three sets of multi-exposure images corresponding to static scenes are added and also a comparison with popular tone mapping operators and the other direct LDR image generation methods have been carried out in this paper.

The proposed gradient domain solution for static scenes is then extended for handling dynamic scenes and the results are discussed. The proposed approach replaces the existing “exposure fusion” approach [34] used in [44]. The added advantages of the gradient domain solution are explained in this paper. The LDR image results corresponding to dynamic scenes are additionally compared with those of the LDR image results of “Merge to HDR Pro” tool in Adobe® Photoshop CS5 [1]. This tool requires the knowledge of exposure settings while the proposed approach does not.

To start with, we shall provide an overview of the existing literature on HDR imaging and segmentation. We shall then look at the proposed approach for generating an LDR image from a set of multi-exposure images of a real world scene in detail. We shall then present the results of the proposed approach for a few static and dynamic scenes. We show that

state-of-the-art LDR images can be generated using the proposed approach. Such LDR images can be made compatible with HDR displays using inverse tone mapping [6].

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 [45]. There are digital cameras which can capture the entire dynamic range of the scene using a single snapshot [51]. These cameras are very expensive. In this section, we shall 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 [29]. 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 [11]. Mitsunaga and Nayar estimated the CRF by parameterizing it and then generated the HDR image [35]. CRF can also be recovered approximately from a single image [26]. However, this approach assumes that the image has significant edges present in the entire available dynamic range. 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 [45]. Assuming the CRF is known, Granados et al. recently developed a method to generate an HDR image in the presence of various types of noise [18].

The generated HDR image can then be encoded in Radiance RGBE (.hdr) [56] or OpenEXR (.exr) [7] 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. [14]. A generic tone mapping operator has been modeled to explain existing operators in [30]. An overview of the different types of tone mapping operators (spatial domain, gradient domain, frequency domain) can be found in [45].

For a static scene, there are several 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 [10, 41, 52]. An interactive method for compositing image regions was proposed by Agarwala et al. [2]. The method by Goshtasby uses entropy measures on blocks to combine multi-exposure images [17]. Exposure fusion combines multi-exposure images on a Laplacian pyramid using an appropriate weighting function [34]. A variational, iterative solution for combining multi-exposure images was introduced in [42]. Bilateral filter was used to define weighting function and composite the multi-exposure images in [43].

In the case of static scenes, capturing multi-exposure images with a hand-held camera can lead to registration mismatch. The images can be registered using bitmaps computed on image pyramids [55]. However, 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 move when there is wind in the scene. In other words, the scene is most probably dynamic. When the methods mentioned above are employed for compositing multi-exposure images of a dynamic scene, the objects in motion in the scene give rise to artifacts called ghosts. It is required that the changes in the scene are detected a priori before compositing is performed on the multi-exposure images. We shall first look at the methods previously used for removing ghosting artifacts.

The change detection across the multiple differently exposed images can be merged with the HDR image generation process. Then appropriately lesser weights can be given to the pixel locations of an image if it is found to have scene change [23]. This technique helps one to eliminate ghosts to certain extent. This approach suffers from the fact that even pixel locations which have scene change are given some weight which may lead to artifacts which are evident while looking closely. This approach also assumes the knowledge of CRF to generate the HDR image.

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 [22]. This method fills the motion regions using 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 [16]. This approach preserves contrast even in the motion regions as only the static regions from multiple images are combined. However, estimation of CRF of the imaging system required by this approach is

challenge and any error in estimation will lead to poor detection of dynamic regions of the scene. Another approach for eliminating ghosting artifacts while creating mosaic from images having possible exposure change was proposed in [53]. However, the emphasis of the work has been primarily on seam removal and not on HDR imaging.

A recent work assumed no knowledge of CRF and reconstructed the dynamic scene as an LDR image from multi-exposure images [57]. 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. [49]. 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 [24]. We focus more on the automatic segmentation approaches in this work as we intend to develop an automatic method to compensate for scene changes 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 [8, 9]. 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 [50]. The bottom-up segmentation methods have derived much interest within the computer vision community of late as they lead to better segmentation of the foreground objects.

The algorithm by Shi and Malik uses normalized cuts to split a given image into multiple regions which are homogeneous [50]. This approach was later extended to define different homogeneous regions of the image as superpixels and they are grouped for segmentation [48]. Each superpixel is a collection of a set of pixels inside a closed contour signifying uniformity in terms of color, intensity 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 ([36, 37]).

Additionally, even neighborhoods can be defined for superpixels to improve the segmentation [15].

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 has 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 [21].

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 images. The key challenge here is the identification of appropriate decision regions to group the superpixels. Given a set of K multi-exposure images corresponding to a dynamic scene, we need to define $K - 1$ different decision regions with respect to the reference image. We shall explain the steps involved in the modeling of decision regions, the grouping of superpixels and LDR image compositing in more detail in the next section.

3 Proposed approach

The 2-D image formation can be described by (1).

$$I(x, y) = f(\Delta t E(x, y)), \quad (1)$$

where $I(x, y)$ is the intensity value of the image, Δt is the exposure time, and $E(x, y)$ is the image irradiance. The image irradiance $E(x, y)$ and the image intensity $I(x, y)$ are related by a non-linear function called camera response function (CRF) f . The term $(\Delta t E(x, y))$ is called the exposure. Generally, CRF is estimated from a set of multi-exposure images and plotted as the logarithm of exposure vs. the image intensity.

While capturing multi-exposure images corresponding to a dynamic scene, we can rewrite (1) as shown in (2).

$$I_k(x, y) = f(\Delta t_k E_k(x, y)), \quad (2)$$

where $I_k(x, y)$ represent the intensity values of the k th image in the exposure stack with exposure time Δt_k . Here E_k is the irradiance of the scene corresponding to the k th exposure as the scene changes across the images. Given a set of observations $I_k(x, y)$ corresponding to a dynamic scene, we want to generate an artifact-free but high-contrast LDR image of the scene when the CRF f and exposure times Δt_k are not known. Further we have the issue of irradiance changing as well.

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 salient feature of the proposed approach is that we composite patches from multiple images even in regions which show scene change, 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 Estimation of the decision regions

Consider any two observations of a scene $I_1(x, y)$ and $I_2(x, y)$ of a static scene which differ only in the exposure times. These two images can be related by the linear equation in the log domain shown in (3).

$$\log f^{-1}(I_2(x, y)) = \log f^{-1}(I_1(x, y)) + \log\left(\frac{\Delta t_2}{\Delta t_1}\right). \quad (3)$$

This equation shows us that the knowledge of CRF f and hence its inverse will prove vital in the estimation of dynamic regions of the scene. This fact was employed by Gallo et al. to generate an HDR image corresponding to a dynamic scene without any ghosting artifacts [16]. In this work we do not assume the knowledge of CRF f . We shall now discuss how dynamic regions can be determined in each of the multi-exposure images with respect to a reference image without the knowledge of CRF f and the exposure times.

The intensity values of these two images can also be related by (4).

$$I_2(x, y) = f\left(f^{-1}(I_1(x, y))\left(\frac{\Delta t_2}{\Delta t_1}\right)\right) \quad (4)$$

which is of the form

$$I_2(x, y) = u_{2,1}(I_1(x, y)). \quad (5)$$

The intensity values of these two images can be characterized by a function called comparametric function $u_{2,1}$ [28]. The comparametric function is also referred to as intensity mapping function (IMF) and is estimated from the histograms when there is a minimal scene change between the two images [19]. 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 times. 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. This function can be accurately estimated when the scene is static as intensity values in any pixel locations can be used to estimate the IMF. However, in the case of a dynamic scene, we need to have a different technique to estimate IMF.

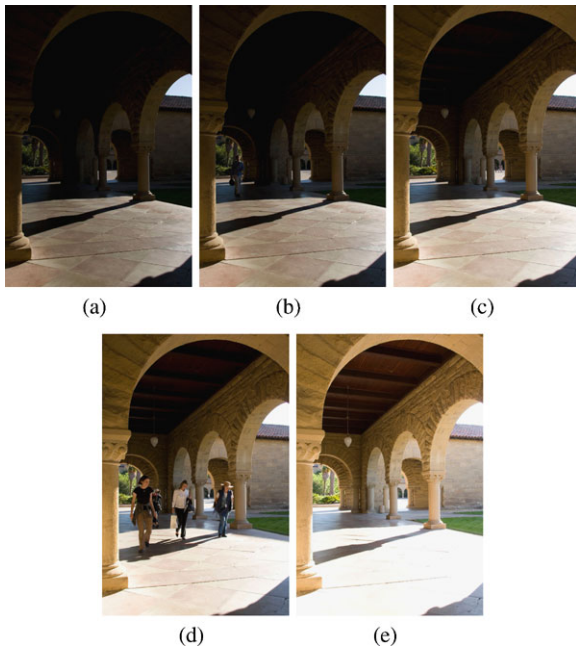


Fig. 1 (a–e) Multi-exposure images of a dynamic scene. Images courtesy: Orazio Gallo, UCSC

Consider a set of multi-exposure images corresponding to a dynamic scene shown in Fig. 1. 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, we observe that the scene changes appreciably in Fig. 1(b and 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 [16]. 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 are used to estimate the IMF between a pair of images from the exposure stack in Fig. 1.

It is worth noting that the normalized intensity values of the multi-exposure images are in the range $[0, 1]$. The weighted variance measure $V(x, y)$ can be computed for K differently exposed images $I_i(x, y)$ using (6).

$$V(x, y) = \frac{\sum_{k=1}^K \zeta_k(x, y) I_k^2(x, y) / \bar{\zeta}(x, y)}{(\sum_{k=1}^K \zeta_k(x, y) I_k(x, y))^2 / (\bar{\zeta}(x, y))^2} - 1, \quad (6)$$

where $\bar{\zeta}(x, y) = \sum_{k=1}^K \zeta_k(x, y)$ and the weight is given by the Gaussian function

$$\zeta_k(x, y) = e^{\frac{-(I_k(x, y) - 0.5)^2}{2 \times 0.2^2}}. \quad (7)$$

This Gaussian function $\zeta_i(x, y)$ is used as weight in order to provide lesser weight to the overexposed and the underexposed pixel locations. We use an appropriate threshold

(0.25 times the maximum weighted variance) to detect pixel locations which show low weighted variance measures. In the case of noisy multi-exposure images, simple Gaussian spatial smoothing or anisotropic diffusion can be used prior to the computation of weighted variance [40]. The weighted variance measure has already been used to fill intensity values from one of the images in the exposure stack in the dynamic regions in [22, 45]. However, such an approach reduces contrast in regions where there is scene change in any of the multi-exposure images.

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. We assume that we have enough pixel locations on the image grid where there is no scene change in any of the images. This assumption is, in general, true for a natural scene and may not apply to complex scenes such as those involving crowd motion in most of the pixel locations. However, such occurrences are a rarity while processing natural scenes and we employ the weighted variance measure as a tool to determine the pixel locations from where we get data points to estimate the IMFs.

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 $S \subset \mathbb{R}^2$ 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 $\psi \subset S$. As a generality, we select one of the multi-exposure images as representing the static scene. We now estimate IMFs between the intensity values of this image and the rest of the images in ψ . Given a set of K multi-exposure images, we have a total of $(K - 1)$ IMFs estimated with respect to the reference image.

We fit a polynomial of order four (chosen empirically) in order to estimate IMF. Such a choice is motivated from the previous work of modeling IMF and is reported to fit the function accurately [19]. The pixel locations of the multi-exposure images in ψ should follow this IMF with respect to the reference image in order to get classified as static. The pixel locations which have some appreciable scene change from the reference would not follow this IMF. The estimated IMF between a pair of images in Fig. 1(b and e) (reference) is shown in Fig. 2(a).

These IMFs, thus estimated, are now employed to find the decision regions to perform bottom-up segmentation. Having estimated the IMFs for each of the multi-exposure images with respect to the reference image, we need to define a constant-width region around the IMF. This constant-width region represents the pixel locations of the test image which do not have any appreciable change with respect to the reference as shown in Fig. 2(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 exploit this constant-width region

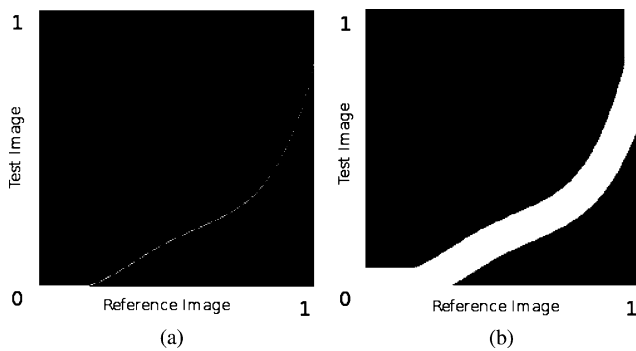


Fig. 2 (a) The IMF between a pair of images in Fig. 1 (b and e) (reference), and (b) The constant-width region which defines the decision boundary for static and dynamic regions

for decision regions as this provides us a means to deal with noisy images. The width of the region around the IMFs can be increased in the case of multi-exposure images with more noise.

3.2 Superpixel grouping

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 excluding the reference image [48]. Alternately, one can use a fast algorithm to speed up the over-segmentation process [25, 54]. We use superpixels as it saves us from the huge load of classifying each pixel and grouping them. Further, it allows us to take care of the object boundaries during segmentation which is not possible when other types of patches are used. The superpixels corresponding to the images shown in Fig. 1 are shown in Fig. 3. As can be observed in Fig. 3(e), the superpixels do not cross the object boundaries and grouping them enables us to recover the exact silhouette of the scene change. Instead of classifying every pixel, we classify the superpixels for the possible scene change with respect to the reference (Fig. 1(e)). 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) shown in Fig. 2. 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 effectively lets us group all the superpixels of the given image which show changes with respect to the reference image unlike other methods based on square patches [16]. This enables us to recover an exact boundary

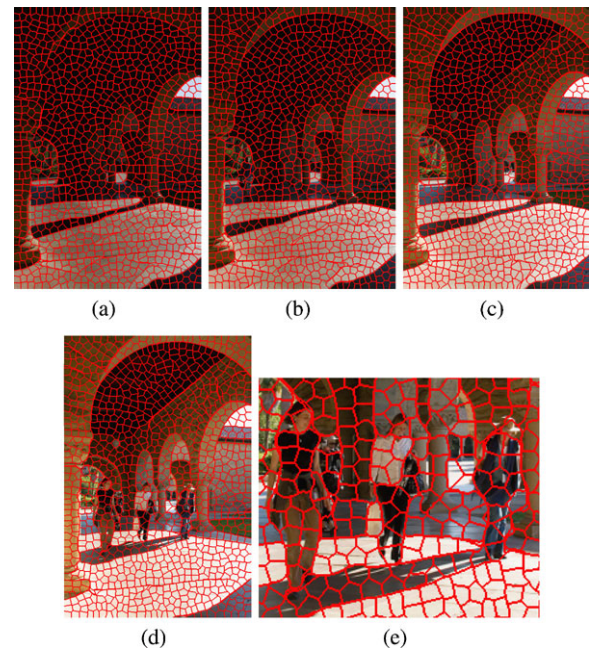


Fig. 3 (a–d) Superpixels estimated corresponding to the first four multi-exposure images in Fig. 1, and (e) Zoomed out superpixels corresponding to a region of (d)

of the object present in the dynamic region of the scene. This is the novel bottom-up segmentation algorithm we have 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 Fig. 4. 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.

3.3 Piecewise rectangular approximation

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). The segmented images shown in Fig. 4 present a new challenge while compositing the final LDR image. As one can observe from Fig. 4(d), the segmented regions can be very irregular which leads to problems when we want the LDR image to be generated without any visible seams. Also, one cannot guarantee that these grouped regions would be closed as evident in Fig. 4(d). Possible solution can be the use of gradient domain processing with Dirichlet boundary conditions and employ a Poisson solver to reconstruct the LDR image. However, we do not have the actual values of the LDR image on the segmentation boundaries which makes this approach

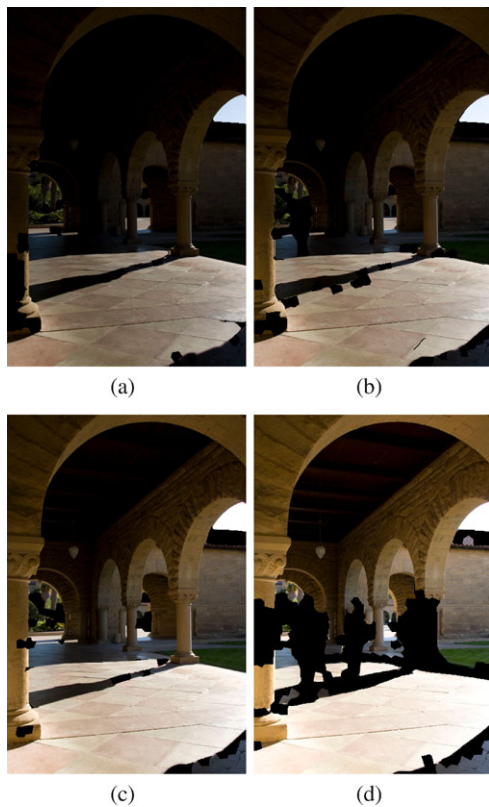


Fig. 4 (a–d) Bottom-up segmentation through superpixel grouping performed on the images in Fig. 3

infeasible. We need to adopt a different strategy to handle these irregular boundaries.

We shall now discuss the piecewise approximation approach to handle these irregular boundaries. We split the images into overlapping patches of certain size (say 6×6) with an overlap of one pixel in each direction. 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 results in piecewise rectangular approximation of the bottom-up segmentation boundaries as visible in Fig. 5. As one can visualize from Fig. 5(d), we have the freedom to choose the size of the patches which approximate the segmentation boundary. If faster compositing is desired, one can use larger patches to approximate the segmentation boundaries. However, a choice of larger patch size will make the piecewise rectangular approximation very erroneous.

3.4 Compositing static patches

We can now use any of the static multi-exposure compositing algorithms on the patches marked as static regions [34, 42, 43]. But they tend to generate seams in the composited image when unequal number of images are stitched at different patch locations. The gradient domain compositing

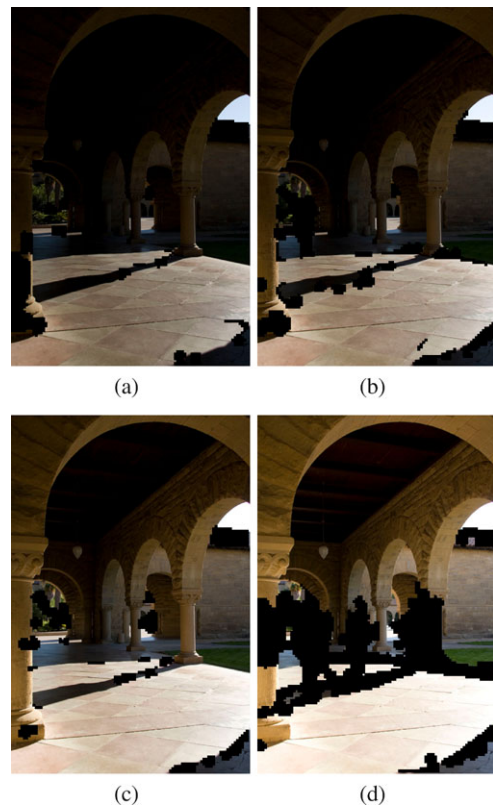


Fig. 5 (a–d) Piecewise rectangular approximation of the segmentation boundaries shown in Fig. 4

helps us to prevent seams from appearing in the final composite image apart from enabling us to perform faster compositing compared to the existing methods. To avoid these seams across patch boundaries while combining different number of images at each patch location, we use overlapping patches (of size 6×6) with an overlap of one pixel in each direction.

3.4.1 Preprocessing of gradients

This step is meant to process the image gradients in a such a way that saturation is reduced if the patch is over- or under-exposed. This is done by carefully modifying their corresponding gradient fields. A method to correct the nearly saturated regions in the image domain was recently proposed by Masood et al. [33] and Guo et al. [20]. We perform a similar operation in this work in the gradient domain.

Given a set of multi-exposure images, we subject their gradients to an illumination change function specified in [39]. This operation is performed on each patch independently in order to brighten the less exposed image patches and darken the more exposed image patches as a preprocessing operation. We modify the gradient fields corresponding to each patch location as shown in (8).

$$\tilde{g}_k(x, y) = \left(\frac{\text{Mean}^P[|\nabla I_k(x, y)|]}{|\nabla I_k(x, y)|} \right)^{\beta_k} \nabla I_k(x, y), \quad (8)$$

where $\text{Mean}^P[|\nabla I_k(x, y)|]$ is the mean gradient magnitude of the p th patch in the k th image, $\nabla I_k(x, y)$ is the gradient vector of the k th image, $\tilde{g}_k(x, y)$ is the corresponding modified gradient vector, and β_k is a real number signifying the relative exposure time corresponding to each image. This parameter has to be modeled in such a way as to reduce the over- and under-exposure in the patches. We define β_k as given in (9).

$$\beta_k = 0.5(0.5 - \text{Mean}^P[I_k(x, y)]), \quad (9)$$

where $\text{Mean}^P[I_k(x, y)]$ is the mean intensity corresponding to the p th patch in the k th image. This formulation enables us to penalize the gradients whenever there is over- or under-saturation in the intensity domain.

3.4.2 Gradient domain compositing

The gradient domain compositing involves weighing the gradients of input images at each patch location appropriately to obtain the resultant gradient field $\hat{g}(x, y)$. The compositing process can be better explained by (10).

$$\hat{g}(x, y) = \sum_{k=1}^K \alpha_k(x, y) \tilde{g}_k(x, y), \quad (10)$$

where $\alpha_k(x, y)$ is the weighting function which weighs the gradients $\tilde{g}_k(x, y)$ and $\sum_{k=1}^K \alpha_k(x, y) = 1$.

Consider the images $I_k(x, y)$ to have intensity values in the normalized range $[0, 1]$. We composite the modified gradients $\tilde{g}_k(x, y)$ using the weighting function in (11).

$$\alpha_k(x, y) = \delta(b_k(x, y) + c_k(x, y)) + (1 - \delta)(b_k(x, y)c_k(x, y)), \quad (11)$$

where $0 \leq \delta \leq 1$, $b_k(x, y) = e^{\frac{-(I_k(x, y) - 0.5)^2}{2 \times 0.2^2}}$ (same as the function in (7)) is the brightness term used to provide less weight to overexposed and underexposed regions and $c_k(x, y)$ is the normalized local contrast term which equals the mean of the gradient magnitudes in the neighborhood of (x, y) as defined in (12).

$$c_k(x, y) = \frac{1}{N_\Omega} \sum_{k=1}^K \tilde{g}_k(x, y), \quad (12)$$

where N_Ω is the neighborhood of the pixel location (x, y) . We use the 8-neighbors in this work when $N_\Omega = 9$. Equation (11) has significance in always assigning more weight to that image which has less saturation and more local contrast. Note that the term $b_k(x, y)$ is high whenever the observation is in the middle of the available dynamic range, providing a larger weight for the weighting function. Similarly, if the local contrast $c_k(x, y)$ is high, the gradient is given

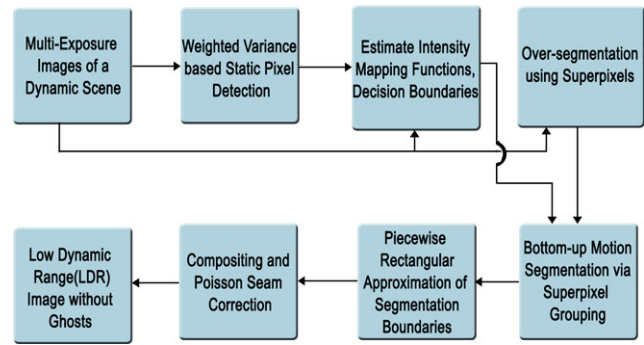


Fig. 6 Schematic representation of the proposed approach

a higher weight. Since presence of noise in the observation $I_k(x, y)$ may adversely affect the local contrast, the measure is smoothed out locally through an averaging process. We found that the sum and the product of the brightness and the contrast terms are both equally effective in achieving the task, which justifies the choice of wright given in (11).

3.5 Poisson seam correction

Had we used non-overlapping patches, there would not have been any information passage between adjacent patches. This would have resulted in visible seams on the patch boundaries. The use of overlapping patches for approximating the segmentation boundaries enables us to avoid these visible seams on the patch boundaries and reduce artifacts in the LDR image. The seams across patch boundaries are avoided by the use of overlapping patches and this process is called Poisson seam correction. We shall now use the gradient domain solution to operate on the patches which do not have scene change. We will get a composited gradient patch at each of the patch location. These gradients corresponding to patch locations are now used for the reconstruction of the desired high-contrast LDR image.

The gradients corresponding to the composited patches (of size 5×5 , one less than the size of overlapping patches) are arranged on the image grid S . The resultant vector field may not be conservative. We employ a direct Poisson solver with Neumann boundary conditions to generate the scalar field closest to the vector field [3, 39]. The scalar field obtained through Poisson seam correction operation is the desired LDR image. The LDR image will not have any ghosting artifacts as we have eliminated the patches of the multi-exposure images which correspond to the dynamic regions of the scene. The complete schematic representation of the steps involved for dynamic scenes is shown in Fig. 6.

4 Results

In this section, we shall first discuss the results of the proposed approach in the case of static scenes. We consider

three sets of multi-exposure images corresponding to static scenes. These sets of images are labeled as lizard, star, and house. For the first two sets, we know the exposure times used to capture the images and hence HDR image of the scene can be generated. We use “Merge to HDR” tool in Adobe® Photoshop CS4 to generate the HDR images corresponding to the first two sets. We use the open source HDR imaging software called pfs-tools discussed for tone mapping the HDR images using different tone mapping operators [32].

There are many recent approaches which can composite multi-exposure images and produce an LDR image of a static scene: variational compositing [42], exposure fusion [34], and bilateral compositing [43]. These methods assume no knowledge of camera response function and exposure times. However, these methods tend to produce a final image which does not preserve the contrast of the real world scene and are not applicable for dynamic scenes. However, since our approach performs compositing in the gradient domain, there is a lesser loss of contrast and we obtain seamless blending.

Aydin et al. have developed a metric which can generate a distortion map by comparing two images having different dynamic ranges [5]. The metric requires one to provide both the HDR image and the LDR image obtained from the HDR image using a tone mapping operator. The distortion maps show the probability of the occurrence of different distortions at every pixel location—over amplification of visible contrast, loss of visible contrast, and reversal of visible contrast. This metric measures the distortions suffered at each pixel location when the dynamic range of an HDR image is compressed. This quality metric is applicable for the tone mapping operators which obtain an LDR image from the corresponding HDR image.

We cannot use the dynamic range independent metric by Aydin et al. for producing distortion maps for the generated LDR images [5]. This metric requires a reference HDR image to be used as the exact replica of the dynamic scene captured. The quality of the HDR image corresponding to a dynamic scene depends on the scene change detection algorithm used. We do not yet have a method to produce completely artifact-free HDR image in the case of dynamic scenes. We, therefore, employ visual examination to compare LDR images obtained using different approaches.

We first present the computational time required for the direct LDR image generation methods. For compositing 9 images of size 2464×1632 using Matlab® in an Intel Xeon® machine with 4 GB RAM, and without involving the motion segmentation modules, variational compositing took more than 15 minutes, exposure fusion took 107 seconds, bilateral filter based compositing took 160 seconds, while the proposed gradient domain compositing took 113 seconds.

Let us now consider the scene ‘lizard’ which is captured using the 8 multi-exposure images shown as thumbnails in

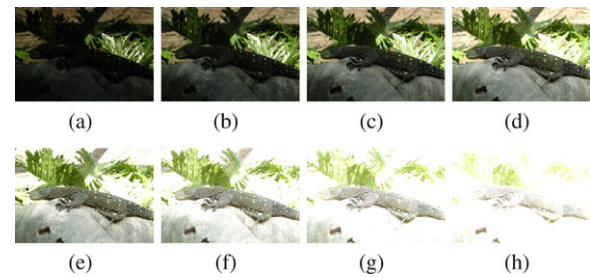


Fig. 7 (a–h) Multi-exposure images of a static scene labeled as ‘lizard.’ Images courtesy: Erik Reinhard, University of Bristol

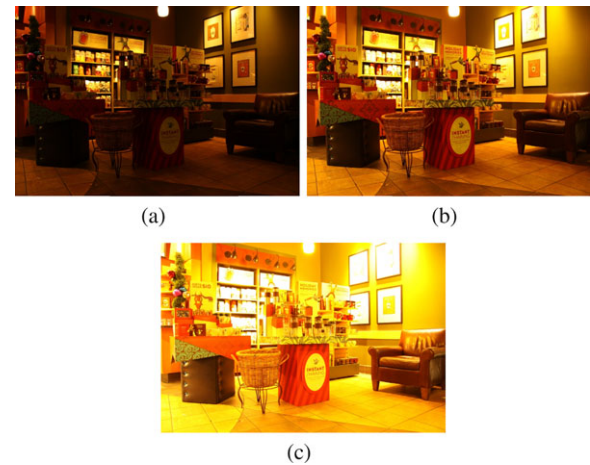


Fig. 8 (a–c) Multi-exposure images of a static scene labeled as ‘star.’ Images courtesy: Gopinath Sivanesan, San Antonio

Fig. 7. There is a lizard lying on a rock which is located in a tree shade while the bushes behind the rock are sunlit. This is a typical scene having a very high dynamic range where both brightly (bushes) and poorly (lizard and rock) illuminated regions are present in the same scene. The images in parts (a)–(d) of Fig. 7 have low exposure times and hence capture the details present in the brightly illuminated regions (bushes) of the scene. The images in parts (e)–(g) of the figure have high exposure times and hence capture the details present in the poorly illuminated regions of the scene (lizard and rock). The images in Fig. 7(h) are so highly exposed that all details are completely saturated. Later, we shall consider the multi-exposure images corresponding to static scenes shown in Fig. 8(a–c) and Fig. 9(a–d) for comparison of LDR image results of various approaches. To start with, let us first examine the LDR image results shown in Fig. 10(a–l) corresponding to the multi-exposure images shown in Fig. 7(a–h).

The LDR image results obtained using methods that do not require camera parameters are shown in Fig. 10(a–d). We also provide the LDR results of some of the popular tone mapping methods for comparison in Fig. 10(e–l). On visual examination, we can observe that the tone-mapped LDR image shown in Fig. 10(f) yields the best result in terms of both brightness and contrast. This method is the gradient do-

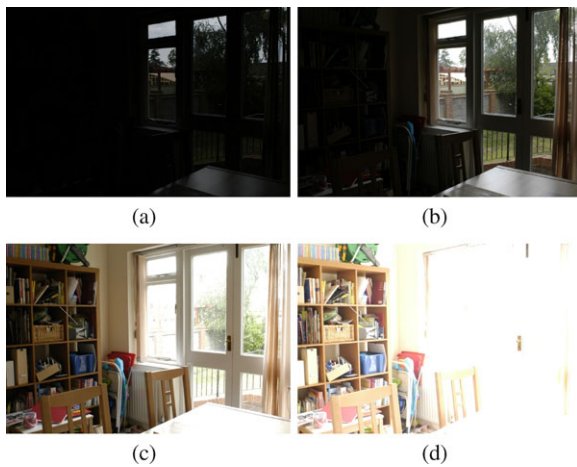


Fig. 9 (a–d) Multi-exposure images of a static scene labeled as ‘house’. Images courtesy: Min H. Kim

main tone-mapping operator proposed by Fattal et al. [14]. However, this method requires the knowledge of camera parameters. As a matter of fact, all methods corresponding to Fig. 10(e–l) require this knowledge. The proposed blind, gradient domain solution, yields the best LDR image after this approach. This can be observed in Fig. 10(d).

The other methods either do not preserve brightness in both brightly and poorly illuminated regions or suffer from reduction in contrast. A typical example of the under-saturation of the poorly illuminated regions (lizard and rock) is shown in Fig. 10(l) where time-dependent visual adaptation tone-mapping operator is used [38]. An approach which leads to a loss of contrast for this scene is the exposure fusion approach as can be seen in Fig. 10(b) [34]. These are the inferences one can arrive at through visual examination.

Consider the set of multi-exposure images labeled as ‘star’ shown in Fig. 8. These images are captured in a coffee shop. The dynamic range of the scene is not that large as the sunlit scene and hence 3 images appear to be sufficient to capture the entire dynamic range of the scene as shown in Fig. 8. In this scene, the lamp and the adjacent regions such as paintings mounted on the wall and the boxes on the cupboard are brightly illuminated. The chairs located on the floor which are far away from the lamp are poorly illuminated. The images are arranged in the increasing order of exposure times. The lamp and the adjoining regions are better captured in Fig. 8(a) while the regions far away from the lamp are better captured in Fig. 8(b, c).

The LDR images generated using the proposed approach along with that of other tone mapping methods and exposure fusion method are presented in Fig. 11. It can be seen that the best LDR image result is the one generated using the bilateral filter (Fig. 11(c)). The proposed gradient domain approach also yields a high-contrast LDR image of the scene without much over- or under-saturation (Fig. 11(d)). For this scene, all the direct LDR image generation methods yield

better results compared to those of the tone mapping operators as can be observed in Fig. 11(a–d). The bilateral filter based tone mapping operator produces an over-saturated LDR image irrespective of the values for different parameters (Fig. 11(e)) [13]. The best LDR result among the tone mapping operators is obtained while using the gradient domain tone-mapping operator (Fig. 11(f)) [14].

Consider a scene labeled as ‘house’ with a very high dynamic range shown in Fig. 9. In this scene, the interior of the room is very poorly illuminated while the trees visible through the window and the door are very brightly illuminated by the sun. Such scenes are quite challenging as the dynamic range of the entire scene is very large. For this set of multi-exposure images, we do not have the knowledge of exposure times. The HDR image cannot hence be reliably estimated for such a scene because of its enormous dynamic range. We will have to apply only the direct methods of LDR image generation and compare the performance of these methods for this scene.

The LDR images generated using different approaches are shown in Fig. 12. For this set of images, exposure fusion method (Fig. 12(b)) and the proposed gradient domain solution (Fig. 12(d)) give the best results. The variational solution yields an LDR image (Fig. 12(a)) which is over-saturated in the brightly illuminated regions due to the inability of this approach to handle highly overexposed regions in Fig. 12(d). The bilateral filter based solution provides an LDR image with better contrast compared to the other methods and is slightly under-saturated in the poorly illuminated regions (Fig. 12(c)).

We shall now 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 Gallo et al. [16] and “Merge to HDR pro” tool in Adobe® Photoshop CS5 [1]. Gallo et al. employ an interactive tone mapping method proposed by Lischinski et al. [27]. 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 those of Gallo et al. [16] and Adobe® Photoshop CS5 [1]. Both these methods require the knowledge of exposure settings in order to generate the HDR image and then the LDR image. In case exposure settings of the multi-exposure images are unavailable, the results might be very erroneous. We are not aware of the technique used in building “Merge to HDR pro” tool in Adobe® Photoshop CS5 [1] as it has not been published by Adobe® in any of their technical reports to the best of our knowledge.

We again employ visual examination to compare LDR images obtained using different approaches. The difference

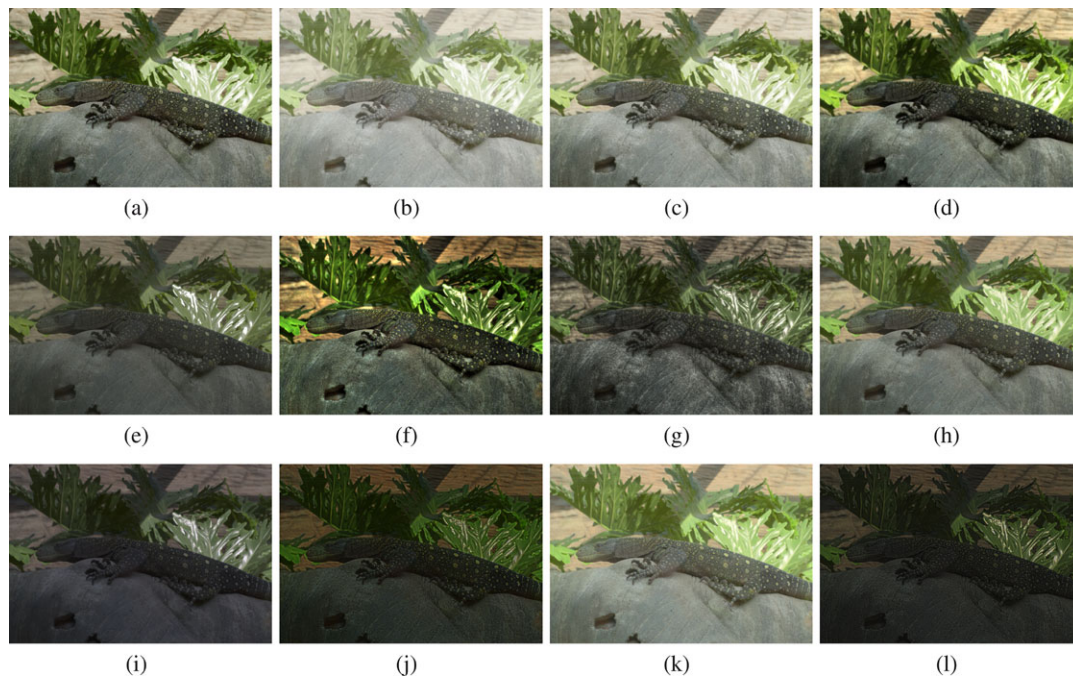


Fig. 10 LDR results of the static scene ‘lizard’ by (a) Variational compositing [42], (b) Exposure fusion [34], (c) Bilateral filter based compositing [43], (d) Proposed gradient domain compositing, (e) Fast bilateral filtering tone-mapping operator [13], (f) Gradient domain tone-mapping operator [14], (g) Contrast domain tone-mapping op-

erator [31], (h) Photographic tone mapping operator [47], (i) Photoreceptor tone mapping operator [46], (j) Ashikhmin’s tone mapping operator [4], (k) Adaptive logarithmic tone-mapping operator [12], and (l) Time-dependent visual adaptation tone-mapping operator [38]

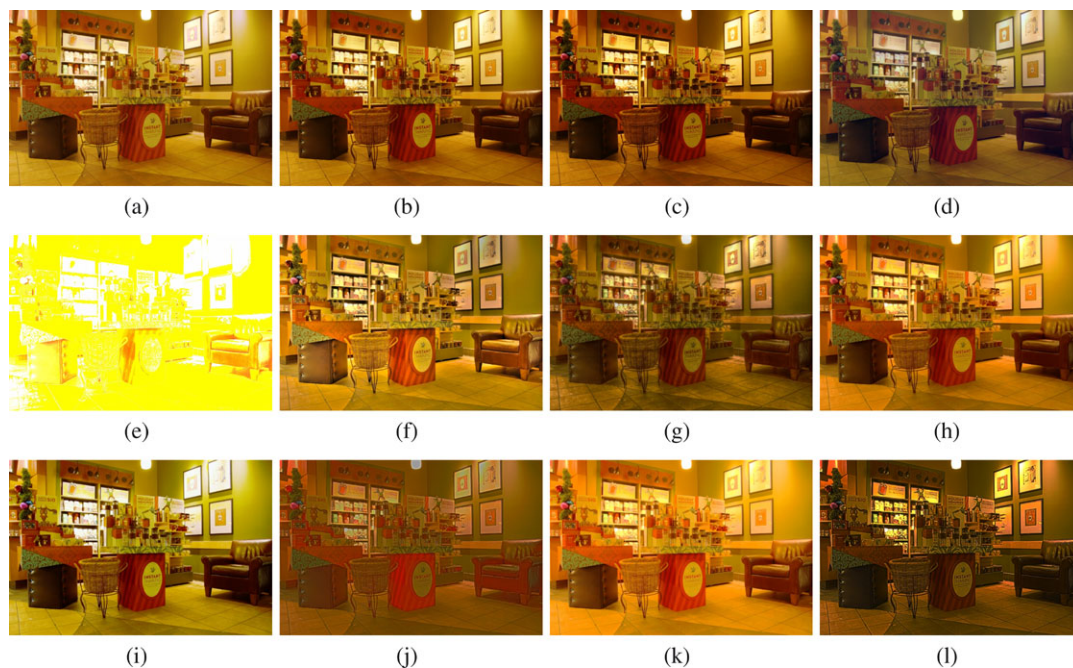


Fig. 11 LDR results of the static scene ‘star’ by (a) Variational compositing, (b) Exposure fusion, (c) Bilateral filter based compositing, (d) Proposed gradient domain compositing, (e) Fast bilateral filtering tone-mapping operator, (f) Gradient domain tone-mapping operator, (g) Contrast domain tone-mapping operator, (h) Photographic

tone mapping operator, (i) Photoreceptor tone mapping operator, (j) Ashikhmin’s tone mapping operator, (k) Adaptive logarithmic tone-mapping operator, and (l) Time-dependent visual adaptation tone-mapping operator



Fig. 12 LDR results of the static scene ‘house’ (a) Variational compositing, (b) Exposure fusion, (c) Bilateral filter based compositing, and (d) Proposed gradient domain compositing

in color tone between the results is mainly due to the type of tone mapping employed by Gallo et al. [16] and Adobe® Photoshop CS5 [1].

Consider the multi-exposure images of a dynamic scene in Fig. 1. Figure 1(e) is picked as the reference image. Figure 13(a) shows the result of an existing approach [43]. As scene change is not accounted for, ghosting artifacts are clearly visible in the generated LDR image. A similar LDR image is obtained when the proposed gradient domain solution without scene change detection is employed. The tone-mapped image corresponding to the HDR image generated by Gallo et al. is as shown in Fig. 13(b). It can be seen that there are some artifacts near the bottom of the pillar and the floor. “Merge to HDR Pro” tool in Adobe® Photoshop CS5 yielded the LDR image shown in Fig. 13(c). One can see a loss of contrast details on the floor which is over-saturated. The LDR image generated using the proposed approach is shown in Fig. 13(d). One can observe that the proposed approach is able to generate an artifact-free LDR image blindly from a set of multi-exposure images. The details in both the brightly and poorly illuminated regions are clearly visible and there are no artifacts.

Consider another two sets of multi-exposure images corresponding to two dynamic scenes as shown in Fig. 14 and Fig. 15. Let us first examine the set of multi-exposure images shown in Fig. 14. 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 sunlit 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 Fig. 14(c) to be the reference image. We detect scene changes on the other images with respect to this image. Figure 16(a) shows the tone-mapped LDR image of Gallo et al. [16]. Though it has higher contrast, this

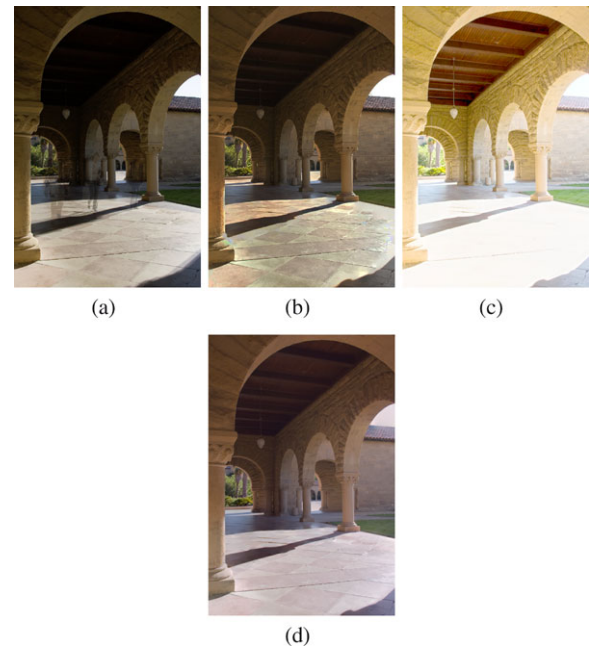


Fig. 13 Results for input images in Fig. 1: (a) LDR image generated by bilateral filter based solution without motion detection showing ghosts, (b) Tone-mapped LDR image using [16], (c) LDR image obtained using Adobe® Photoshop CS5, and (d) LDR image generated using the proposed approach

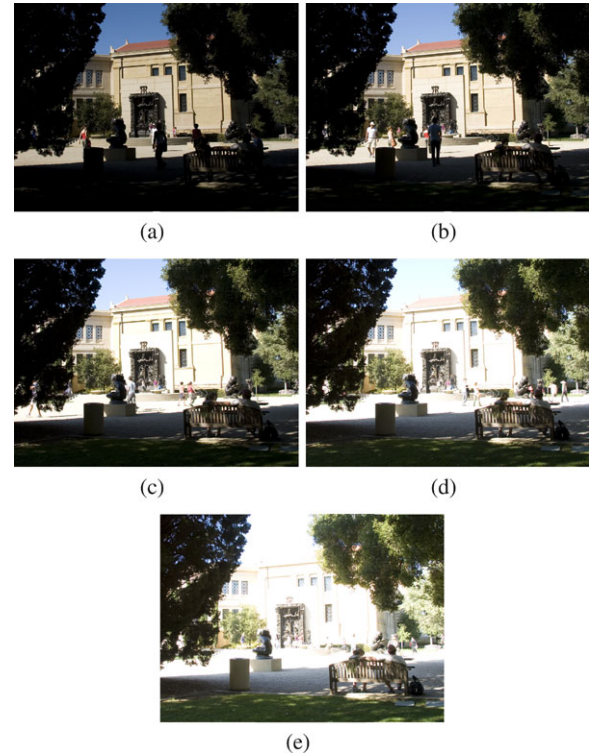


Fig. 14 (a–e) Multi-exposure images of another dynamic scene. Images courtesy: Orazio Gallo, UCSC

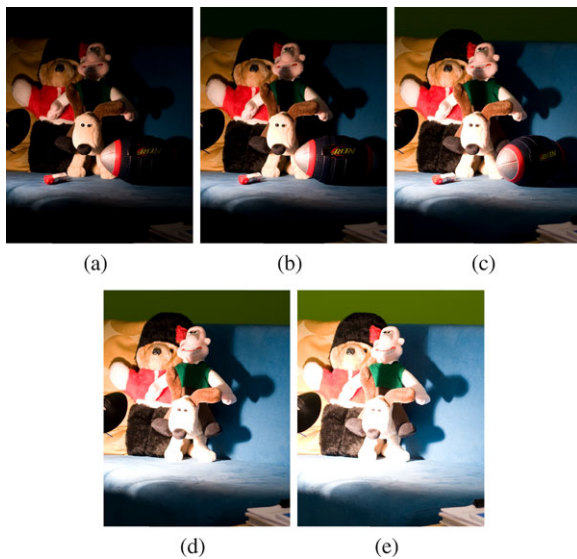


Fig. 15 (a–e) Multi-exposure images of another dynamic scene. Images courtesy: Orazio Gallo, UCSC

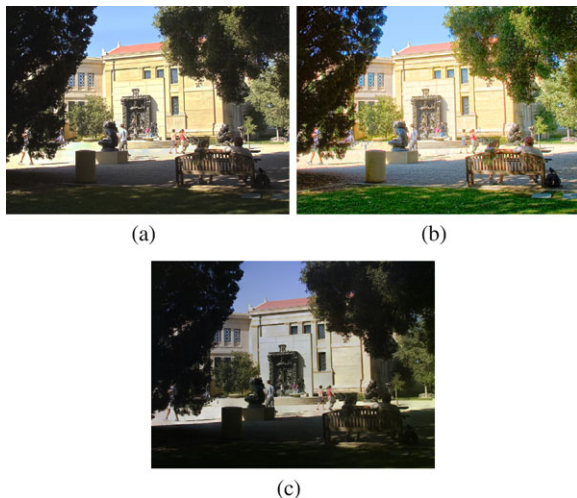


Fig. 16 Results for images given in Fig. 14: (a) Tone-mapped LDR image using [16], (b) LDR image obtained using Adobe® Photoshop CS5, and (c) LDR image generated using the proposed approach

image shows some blue colored artifacts around the walking people. “Merge to HDR Pro” tool in Adobe® Photoshop CS5 provided the LDR image shown in Fig. 16(b). One can observe that the overall contrast is far better than the method by Gallo et al. [16]. However, the ghosting artifacts are quite visible and the LDR image is not visually pleasing. Compositing and Poisson seam correction using the proposed approach enabled us to generate the LDR image shown in Fig. 16(c). 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 Fig. 15. This scene has three dolls, pens, and a ball on the floor (Fig. 15(c)). The

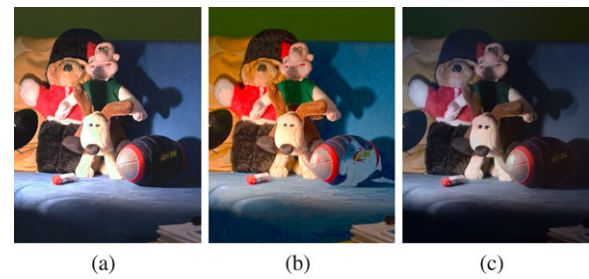


Fig. 17 LDR image results for input images given in Fig. 15: (a) Tone mapped output using [16], (b) Image obtained using Adobe® Photoshop CS5, and (c) Output generated using the proposed approach

ball and the pens are missing in the last two images of the multi-exposure stack (Fig. 15(d, e)). Further we can notice the changes in the positions of the objects across the images. We pick the image (c) in Fig. 15 as the reference image. The interactively tone-mapped result of Gallo et al. [16] is shown in Fig. 17(a). Figure 17(b) shows the LDR image generated using “Merge to HDR Pro” tool in Adobe® Photoshop CS5. For this scene, the LDR image generated using this tool has very good contrast. However, one may observe that there are appreciable ghosting artifacts on the ball. The reconstructed LDR image using the proposed approach is shown in Fig. 17(c). We can see that the proposed approach is able to reconstruct the scene well, albeit with few ghosting artifacts on the ball and the eyes of a doll. These artifacts are due to the fact that there is slight change in the positions of these objects in different images which could not be detected accurately by superpixel grouping. The proposed approach can be enhanced by employing much smaller superpixels or smaller patches which can predict such small changes in positions of the objects in the scene. However, this comes at the cost of reduced computational efficiency. Another instance when such ghosting artifacts can occur is when there is very less common static region present in the multi-exposure images as explained in Sect. 3.1. The LDR image in Fig. 17(a) yields the best contrast among all the images and without any artifacts due to the knowledge of CRF. The proposed approach needs to be improved further in order to generate the LDR image as visually pleasing as that of the approach by Gallo et al.

5 Conclusions

We have proposed a novel bottom-up segmentation approach for detecting motion in multi-exposure images corresponding to a dynamic scene. We then approximated the segmentation boundaries as piecewise rectangles 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. The proposed approach has added advantages

of not requiring both the knowledge of CRF and the tone mapping operation. Further, exposure settings of the multi-exposure images are also not needed. The generated high-contrast 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) [1]. The generated LDR image can be made compatible with HDR displays by using any of the inverse tone mapping algorithms. In this paper, we explained how the algorithms for static scenes can be extended to handle dynamic scenes through bottom-up segmentation approach which does not require any additional information along with the multi-exposure images.

The algorithm discussed in this paper can be used to work on more complex data sets which would enable one to model the solutions better. The computational complexity of the algorithms can be reduced once faster Poisson solvers are integrated into the solutions as and when they are available. The results obtained using the proposed algorithm can be validated more thoroughly based on the human visual system model for accurate compositing of LDR images. This also requires us to arrive at a suitable algorithm for a given digital display device. The proposed approach can be suitably modified to perform region based compositing of multi-exposure images by defining suitable boundary conditions on the regions.

The LDR image generation algorithms thus developed for static and dynamic scenes can be implemented in real time using a GPU. Parallel implementation might help in a fast computation of the desired LDR image. More complex scenes such as crowd motion can be addressed by improving the scene change detection algorithm. The proposed algorithm can also be used to build freeware similar to Enfuse and Tufuse which perform exposure fusion for static scenes.

Acknowledgements The first author would like to acknowledge the financial support provided by Microsoft Research India through a Ph.D. fellowship. The second author would like to thank JC Bose National fellowship and Bharti Center for Communication for the financial support.

References

1. Adobe Photoshop CS5 User Guide. Adobe Systems Incorporated, San Jose (2010)
2. Agarwala, A., Dontcheva, M., Agrawala, M., Drucker, S., Colburn, A., Curless, B., Salesin, D., Cohen, M.: Interactive digital photomontage. *ACM Trans. Graph.* **23**, 294–302 (2004)
3. Agrawal, A., Raskar, R.: What is the range of surface reconstructions from a gradient field. In: *ECCV*, pp. 578–591. Springer, Graz (2006)
4. Ashikhmin, M.: A tone mapping algorithm for high contrast images. In: *Proceedings of the 13th Eurographics Workshop on Rendering*, pp. 145–156. Eurographics Association, Pisa (2002)
5. Aydin, T.O., Mantiuk, R., Myszkowski, K., Seidel, H.P.: Dynamic range independent image quality assessment. *ACM Trans. Graph.* **27**, 69:1–69:10 (2008)
6. Banterle, F., Ledda, P., Debattista, K., Chalmers, A.: Inverse tone mapping. In: *GRAPHITE '06: Proceedings of the 4th International Conference on Computer Graphics and Interactive Techniques in Australasia and Southeast Asia*, pp. 349–356. ACM, Kuala Lumpur (2006)
7. Bogart, R., Kainz, F., Hess, D.: The openexr file format. In: *SIGGRAPH Technical Sketch*, San Diego, USA (2003)
8. Borenstein, E., Ullman, S.: Class-specific, top-down segmentation. In: Heyden, A., Sparr, G., Nielsen, M., Johansen, P. (eds.) *Computer Vision ECCV 2002. Lecture Notes in Computer Science*, vol. 2351, pp. 639–641. Springer, Berlin (2002)
9. Borenstein, E., Ullman, S.: Learning to segment. In: *Computer Vision-ECCV*, pp. 315–328. Springer, Prague (2004)
10. Brinkmann, R.: *The Art and Science of Digital Compositing*. Morgan Kaufmann, San Mateo (1999)
11. Debevec, P.E., Malik, J.: Recovering high dynamic range radiance maps from photographs. In: *Proceedings of the 24th Annual Conference on Computer Graphics and Interactive Techniques, SIGGRAPH '97*, pp. 369–378. ACM, Los Angeles (1997)
12. Drago, F., Myszkowski, K., Annen, T., Chiba, N.: Adaptive logarithmic mapping for displaying high contrast scenes. *Comput. Graph. Forum* **22**(3), 419–426 (2003)
13. Durand, F., Dorsey, J.: Fast bilateral filtering for the display of high-dynamic-range images. *ACM Trans. Graph.* **21**, 257–266 (2002)
14. Fattal, R., Lischinski, D., Werman, M.: Gradient domain high dynamic range compression. *ACM Trans. Graph.* **21**, 249–256 (2002)
15. Fulkerson, B., Vedaldi, A., Soatto, S.: Class segmentation and object localization with superpixel neighborhoods. In: *Computer Vision, 2009 IEEE 12th International Conference on*, pp. 670–677. IEEE, Kyoto (2010)
16. Gallo, O., Gelfand, N., Chen, W., Tico, M., Pulli, K.: Artifact-free high dynamic range imaging. In: *IEEE International Conference on Computational Photography (ICCP)*, San Francisco, CA, USA (2009)
17. Goshtasby, A.: Fusion of multi-exposure images. *Image Vis. Comput.* **23**(6), 611–618 (2005)
18. Granados, M., Ajdin, B., Wand, M., Theobalt, C., Seidel, H., Lensch, H.: Optimal HDR reconstruction with linear digital cameras. In: *Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on*, pp. 215–222. IEEE, San Francisco (2010)
19. Grossberg, M., Nayar, S.: Determining the camera response from images: what is knowable? *IEEE Trans. Pattern Anal. Mach. Intell.* **25**(11), 1455–1467 (2003)
20. Guo, D., Cheng, Y., Zhuo, S., Sim, T.: Correcting over-exposure in photographs. In: *Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on*, San Francisco, CA, USA, pp. 515–521 (2010). doi:[10.1109/CVPR.2010.5540170](https://doi.org/10.1109/CVPR.2010.5540170)
21. Hoiem, D., Efros, A.A., Hebert, M.: Automatic photo pop-up. *ACM Trans. Graph.* **24**, 577–584 (2005)
22. Jacobs, K., Loscos, C., Ward, G.: Automatic high-dynamic range image generation for dynamic scenes. *IEEE Comput. Graph. Appl.* **28**(2), 84–93 (2008)
23. Khan, E., Akyuz, A., Reinhard, E.: Ghost removal in high dynamic range images. In: *Image Processing, 2006 IEEE International Conference on*, Atlanta, GA, USA, pp. 2005–2008 (2006)
24. Levin, A., Lischinski, D., Weiss, Y.: A closed-form solution to natural image matting. *IEEE Trans. Pattern Anal. Mach. Intell.* **30**, 228–242 (2008)

25. Levinstein, A., Stere, A., Kutulakos, K., Fleet, D., Dickinson, S., Siddiqi, K.: Turbopixels: Fast superpixels using geometric flows. *IEEE Trans. Pattern Anal. Mach. Intell.* **31**(12), 2290–2297 (2009)
26. Lin, S., Gu, J., Yamazaki, S., Shum, H.Y.: Radiometric calibration from a single image. In: *Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, CVPR'04*, pp. 938–945. IEEE Computer Society, Washington (2004)
27. Lischinski, D., Farbman, Z., Uyttendaele, M., Szeliski, R.: Interactive local adjustment of tonal values. *ACM Trans. Graph.* **25**, 646–653 (2006)
28. Mann, S.: Comparametric equations with practical applications in quantigraphic image processing. *IEEE Trans. Image Process.* **9**(8), 1389–1406 (2000)
29. Mann, S., Picard, R.W.: On being undigital with digital cameras: Extending dynamic range by combining exposed pictures. In: *Proc. of IS & T 48th Annual Conference*, pp. 422–428 (1995)
30. Mantiuk, R., Seidel, H.: Modeling a generic tone-mapping operator. *Comput. Graph. Forum* **27**(2), 699–708 (2008)
31. Mantiuk, R., Myszkowski, K., Seidel, H.P.: A perceptual framework for contrast processing of high dynamic range images. *ACM Trans. Appl. Percept.* **3**, 286–308 (2006)
32. Mantiuk, R., Krawczyk, G., Mantiuk, R., Seidel, H.P.: High dynamic range imaging pipeline: Perception-motivated representation of visual content. In: Rogowitz, B.E., Pappas, T.N., Daly, S.J. (eds.) *Human Vision and Electronic Imaging XII. Proceedings of SPIE*. SPIE, San Jose (2007)
33. Masood, S., Zhu, J., Tappen, M.: Automatic correction of saturated regions in photographs using cross-channel correlation. *Comput. Graph. Forum* **28**(7), 1861–1869 (2009)
34. Mertens, T., Kautz, J., Reeth, F.V.: Exposure fusion: A simple and practical alternative to high dynamic range photography. *Comput. Graph. Forum* **28**(1), 161–171 (2009)
35. Mitsunaga, T., Nayar, S.: Radiometric self calibration. In: *Computer Vision and Pattern Recognition, 1999. CVPR 1999. Proceedings of the 1999 IEEE Computer Society Conference on*, Ft. Collins, CO, USA, p. 1374 (1999)
36. Mori, G.: Guiding model search using segmentation. In: *Computer Vision, 2005. ICCV 2005. Tenth IEEE International Conference on*, Beijing, China, vol. 2, pp. 1417–1423 (2005)
37. Mori, G., Ren, X., Efros, A., Malik, J.: Recovering human body configurations: Combining segmentation and recognition. In: *Computer Vision and Pattern Recognition, 2004. CVPR 2004. Proceedings of the 2004 IEEE Computer Society Conference on*, Washington, DC, USA, vol. 2, pp. II-326–II-333 (2004)
38. Pattanaik, S.N., Tumblin, J., Yee, H., Greenberg, D.P.: Time-dependent visual adaptation for fast realistic image display. In: *Proceedings of the 27th Annual Conference on Computer Graphics and Interactive Techniques, SIGGRAPH '00*, pp. 47–54. ACM, New York (2000)
39. Pérez, P., Gangnet, M., Blake, A.: Poisson image editing. *ACM Trans. Graph.* **22**, 313–318 (2003)
40. Perona, P., Malik, J.: Scale-space and edge detection using anisotropic diffusion. *IEEE Trans. Pattern Anal. Mach. Intell.* **12**(7), 629–639 (1990)
41. Porter, T., Duff, T.: Compositing digital images. *SIGGRAPH Comput. Graph.* **18**, 253–259 (1984)
42. Raman, S., Chaudhuri, S.: A matte-less, variational approach to automatic scene compositing. In: *Computer Vision, 2007. ICCV 2007. IEEE 11th International Conference on*, Rio de Janeiro, Brazil, pp. 1–6 (2007)
43. Raman, S., Chaudhuri, S.: Bilateral filter based compositing for variable exposure photography. In: *EUROGRAPHICS Short Papers*, Munich, Germany (2009)
44. Raman, S., Chaudhuri, S.: Bottom-up segmentation for ghost-free reconstruction of a dynamic scene from multi-exposure images. In: *Proceedings of the Seventh Indian Conference on Computer Vision, Graphics and Image Processing, ICVGIP '10*, pp. 56–63. ACM, Chennai (2010)
45. Reinhard, E., Heidrich, W., Pattanaik, S., Debevec, P., Ward, G., Myszkowski, K.: *High Dynamic Range Imaging: Acquisition, Display, and Image-Based Lighting*. Morgan Kaufmann Series in Computer Graphics. Elsevier, Amsterdam (2010)
46. Reinhard, E., Devlin, K.: Dynamic range reduction inspired by photoreceptor physiology. *IEEE Trans. Vis. Comput. Graph.* **11**(1), 13–24 (2005)
47. Reinhard, E., Stark, M., Shirley, P., Ferwerda, J.: Photographic tone reproduction for digital images. *ACM Trans. Graph.* **21**, 267–276 (2002)
48. Ren, X., Malik, J.: Learning a classification model for segmentation. In: *Computer Vision, 2003. Proceedings. Ninth IEEE International Conference on*, Nice, France, vol. 1, pp. 10–17 (2003)
49. Rother, C., Kolmogorov, V., Blake, A.: “Grabcut”: Interactive foreground extraction using iterated graph cuts. *ACM Trans. Graph.* **23**, 309–314 (2004)
50. Shi, J., Malik, J.: Normalized cuts and image segmentation. *IEEE Trans. Pattern Anal. Mach. Intell.* **22**(8), 888–905 (2000)
51. Spheron: Spherocam HDR. <http://www.spheron.com/>
52. Szeliski, R.: Image alignment and stitching: A tutorial. *Found. Trends Comput. Graph. Vis.* **2**(1) (2008)
53. Uyttendaele, M., Eden, A., Skeliski, R.: Eliminating ghosting and exposure artifacts in image mosaics. In: *Computer Vision and Pattern Recognition, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on*, Kauai, HI, USA, vol. 2, pp. II-509–II-516 (2001)
54. Vedaldi, A., Soatto, S.: Quick shift and kernel methods for mode seeking. In: *Proceedings of the European Conference on Computer Vision (ECCV)*, Marseille, France (2008)
55. Ward, G.: Fast robust image registration for compositing high dynamic range photographs from hand-held exposures. *J. Graph. Tools* **8**(2), 17–30 (2003)
56. Ward, G.J.: The radiance lighting simulation and rendering system. In: *Proceedings of the 21st Annual Conference on Computer Graphics and Interactive Techniques, SIGGRAPH '94*, pp. 459–472. ACM, New York (1994)
57. Zhang, W., Cham, W.K.: Gradient-directed composition of multi-exposure images. In: *Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on*, San Francisco, CA, USA, pp. 530–536 (2010)



Shanmuganathan Raman is currently a Postdoctoral Research Associate at the Indian Institute of Science, Bangalore, India. He received his Bachelors in Engineering and Masters in Technology from PSG College of Technology, Coimbatore and Indian Institute of Technology Bombay, India, respectively. He obtained his Doctor of Philosophy degree at the Indian Institute of Technology Bombay, India under the supervision of Prof. Subhasis Chaudhuri. His research interests include high dynamic range imaging, computational photography, computer vision, and computational neuroscience. He was awarded Microsoft Research India Ph.D. Fellowship for the year 2007.



Subhasis Chaudhuri received his B.Tech. degree in Electronics and Electrical Communication Engineering from the Indian Institute of Technology (IIT), Kharagpur in 1985. He received the M.Sc. and the Ph.D. degrees, respectively, from the University of Calgary, Canada and the University of California, San Diego. He is currently serving as a Professor and the Dean of International Relations at IIT Bombay. He is the co-author of the books 'Depth from defocus: A real aperture imaging approach,' and

'Motion-free super-resolution,' both published by Springer, NY. His research interests include image processing, computer vision and multimedia.