

De-ghosting in High Dynamic Range Imaging Based on Intensity Scaling Cue

Seong-O SHIM¹, Soltan ALHARBI¹, Ishtiaq Rasool KHAN², Wajid AZIZ^{2,3}

¹Department of Computer & Network Engineering, College of Computer Science and Engineering, University of Jeddah, Jeddah 23890, Saudi Arabia

²Department of Computer Science and AI, College of Computer Science and Engineering, University of Jeddah, Jeddah 23890, Saudi Arabia

³Department of Computer Sciences and Information Technology, University of Azad Jammu and Kashmir, Muzaffarabad, 13100, Pakistan
sshim@uj.edu.sa

Abstract—A High Dynamic Range (HDR) image produced from a sequence of low dynamic range (LDR) images can contain motion artefacts (ghosting) if the scene contains moving objects. Conventional de-ghosting methods first detect moving objects in the scene, and then either remove those moving objects totally or reconstruct them. However, these methods are computationally expensive. This paper proposes a de-ghosting method that does not require explicit detection of moving regions. First, the ratio between camera exposure times of a target image and a reference image, which is called the intensity scaling factor in this paper, is computed. Since the information about camera exposure time is not available always, we propose a novel method to estimate the intensity scaling factor from non-saturated and non-moving pixels. Then, the estimated scaling factor is used as a cue to label every pixel in the target image as either static or moving pixel. Finally, the values of moving pixels are corrected with their expected values which can be estimated from the intensity scaling factor. Experimental results show that the proposed method generates more accurate ghost-free HDR images than the existing state of the art methods.

Index Terms—image sequence analysis, image fusion, image reconstruction, image motion analysis, image quality.

I. INTRODUCTION

The human visual system (HVS) has high dynamic range (HDR) of 20 stops (1,000,000:1) in well-lit condition which means that the brightest part human eye can perceive is 2^{20} ($\approx 1,000,000$) times brighter than the darkest part in the same scene. However, most of the compact cameras have dynamic range of just 5-7 stops while DSLR cameras have 8-11 stops. Although the gap is reduced by high-precision modern cameras like Nikon D810 which has 15 stops of dynamic range, the rich contrast information of the scene perceived by the human eye can still not be captured by a single image. One way to mimic the dynamic range of human eye is to generate a high dynamic range (HDR) image from multiple low dynamic range (LDR) images of the same scene taken at various exposures. Several algorithms have been presented for this and are generally referred as multiple exposure methods [1–3]. In these methods, a sequence of LDR images is acquired by changing the exposure time setting of the camera. These LDR images are merged to produce a HDR image. To

render a HDR image on an LDR display, tone mapping techniques are applied which generally use a model of HVS to compress the dynamic range [4–9]. Different techniques for evaluation of the tone-mapping results have also been proposed in literature [10–13].

A common problem involved in making a HDR image is that if there are any moving objects in the scene during the time of capturing LDR image sequence at different exposures, they will appear blended in the reconstructed HDR image at more than one location. This artifact is called the ghosting effect [14]. A lot of research has been conducted on construction of ghost-free HDR images. However, most of the proposed methods are accurate only on certain types of scenes and are slow in general. A recent work in [15] groups the de-ghosting algorithms into three categories: moving object removal methods, moving object reconstruction methods and moving object registration methods.

Moving object removal methods [16–19] check consistency for each pixel across exposures and completely remove the objects that are not stationary in majority of the exposures. These methods assume that the majority pixels of the captured images are static. However, failing this assumption such as in case of deformable-body motions with overlapping regions between exposures, dynamic background in the scene, or insufficient number of input exposures can lead to poor de-ghosting quality. Moving object reconstruction methods [20–33] select one or more reference images for the regions affected by motion. The moving objects are then reconstructed or replaced by the same moving objects extracted from the reference images. While the moving object reconstruction methods are computationally efficient compared to the moving object removal methods, their resultant HDR image may lack in dynamic range on the reconstructed moving objects. On the other hand, moving object registration methods try to align the pixels affected by motion by searching the best corresponding pixels in other exposures. Moving object registration methods can be further classified into optical flow methods and patch-based methods. Optical flow methods [34–39] conduct a pixel-wise matching between two input images with different exposures. In these methods, the matching accuracy determines the final quality of the output HDR image. In addition, optical flow methods have difficulty in dealing with the occlusion, large

This work was funded by Deanship of Scientific Research (DSR), University of Jeddah, under grant No. UJ-02-034-DR.

displacements in the scene, and noise. Patch-based methods [40–43] conduct patch-based optimization that solves for HDR merging and alignment of images simultaneously. Even though these methods produce high quality HDR images, they are computationally expensive due to their intensive searching and patching operations.

The de-ghosting techniques mentioned above require two main steps in general—detection and removal of motion. Accuracy of these two steps affects the quality of the output HDR image. The algorithms used for these steps have high computational complexity in general. We propose a new moving object registration method, which skips the motion detection step and carries it out implicitly with the motion removal step for enhanced speed. In our method, we first select a reference image from the LDR image sequence, manually or by an automatic algorithm. It is desirable that the reference image represents the scene better than any other single image in the sequence. Next, an intensity scaling factor between the reference and each target image is determined. For this step, we use only non-saturated and non-moving pixels to ensure accuracy of the scaling factor. These scaling factors are used as cues to determine the moving pixels in the target image. Finally, the values of the moving pixels in each target image are replaced with their expected values. The corrected LDR images are merged to generate ghosting-free HDR images. In comparison to the best known HDR imaging methods in the literature and a commercial HDR software, the proposed method achieved better de-ghosting quality while capturing more dynamic range of the background scene.

This work is the extended version of our conference paper [32]. In this manuscript, we extensively improved the work: 1) the theory of estimating the intensity scaling factor was elaborated more in detail and we proposed a scheme to verify the validity of the idea, 2) de-ghosting method in fully white and black regions was developed and 3) two more benchmark methods were compared with the proposed method in experiment.

The rest of the paper is organized as follows. Section II presents the proposed de-ghosting algorithm. In this section, we present algorithms to categorize pixels in each exposure as moving or static, and saturated or un-saturated. Moreover, the methods to estimate the intensity scaling factors using non-saturated and non-moving pixels, and to reconstruct the moving objects in the target images to match the reference image are explained. Section III presents some experimental results on different datasets and compares them with the results of some existing methods. Conclusions are given in section IV.

II. DE-GHOSTING ALGORITHM BASED ON INTENSITY SCALING FACTOR

In this section, we present our proposed de-ghosting algorithm, which is computationally fast and outperforms the existing methods.

A. Motivation

The amount of light reaching a sensor location p for Δt units of time, while camera shutter is open, defines a sensor exposure $E(p) \cdot \Delta t$, where $E(p)$ is a sensor irradiation at p .

Then, the pixel intensity value $L(p)$ is defined as a function of sensor exposure as

$$L(p) = f(E(p) \cdot \Delta t) \quad (1)$$

where f is a camera response function. While many different methods have been proposed to estimate the camera response function [44]–[46], many camera companies are using gamma curve approximation [47]. If a camera response function is assumed to be a gamma curve, the intensity value at pixel location p in the reference image L_{ref} is represented as

$$L_{ref}(p) = f(E(p) \cdot \Delta t_{ref}) \approx (E(p) \cdot \Delta t_{ref})^\gamma \quad (2)$$

The reference image is picked from the LDR sequence captured at various exposures such that it represents the scene best. Our method for this is to pick the image which has the least number of under-exposed and over-exposed pixels. Other criteria can also be used, or alternatively an image can be picked manually depending on the location and shape of the moving objects in the image. In the final constructed HDR image, the moving objects will be at the positions where they are in the reference image. Rest of the LDR images in the sequence are referred as the target images. As in (2), the pixel intensity value $L_n(p)$ in the target image L_n can be represented as

$$L_n(p) = f(E(p) \cdot \Delta t_n) \approx (E(p) \cdot \Delta t_n)^\gamma \quad (3)$$

Combining (2) and (3), $L_n(p)$ can be expressed as

$$L_n(p) \approx \left(\frac{\Delta t_n}{\Delta t_{ref}} \right)^\gamma \cdot L_{ref}(p) = \alpha_{ref}^n \cdot L_{ref}(p) \quad (4)$$

which leads to

$$\frac{L_n(p)}{L_{ref}(p)} \approx \left(\frac{\Delta t_n}{\Delta t_{ref}} \right)^\gamma = \alpha_{ref}^n \quad (5)$$

where Δt_n , Δt_{ref} , γ are all constant values. Equations (4) and (5) show that the pixel intensities of the target image L_n are scaled values of the corresponding pixel intensities in the reference image L_{ref} by a factor α_{ref}^n , referred as the scaling factor hereinafter. This scaling factor α_{ref}^n is not dependent on the pixel location p and has a constant value for all pixels. Therefore, we can assume that if $L_n(p)$ is different from $\alpha_{ref}^n \cdot L_{ref}(p)$, the pixel at location p is corresponding to a point on a moving object, which is called a moving pixel in this paper. In section B, we propose a novel algorithm to estimate the scaling factor α_{ref}^n from non-saturated and non-moving pixels of L_{ref} and L_n . In section C, we propose an algorithm to select moving pixels by utilizing the estimated scaling factor as a supporting index. In section D, the selected moving pixels are corrected with their expected values.

B. Estimation of Intensity Scaling Factor

In (5), the scaling factor α_{ref}^n is determined from Δt_n , Δt_{ref} and γ values. However, the exact information about camera exposure time for the captured image is not always available, and generally γ is unknown.

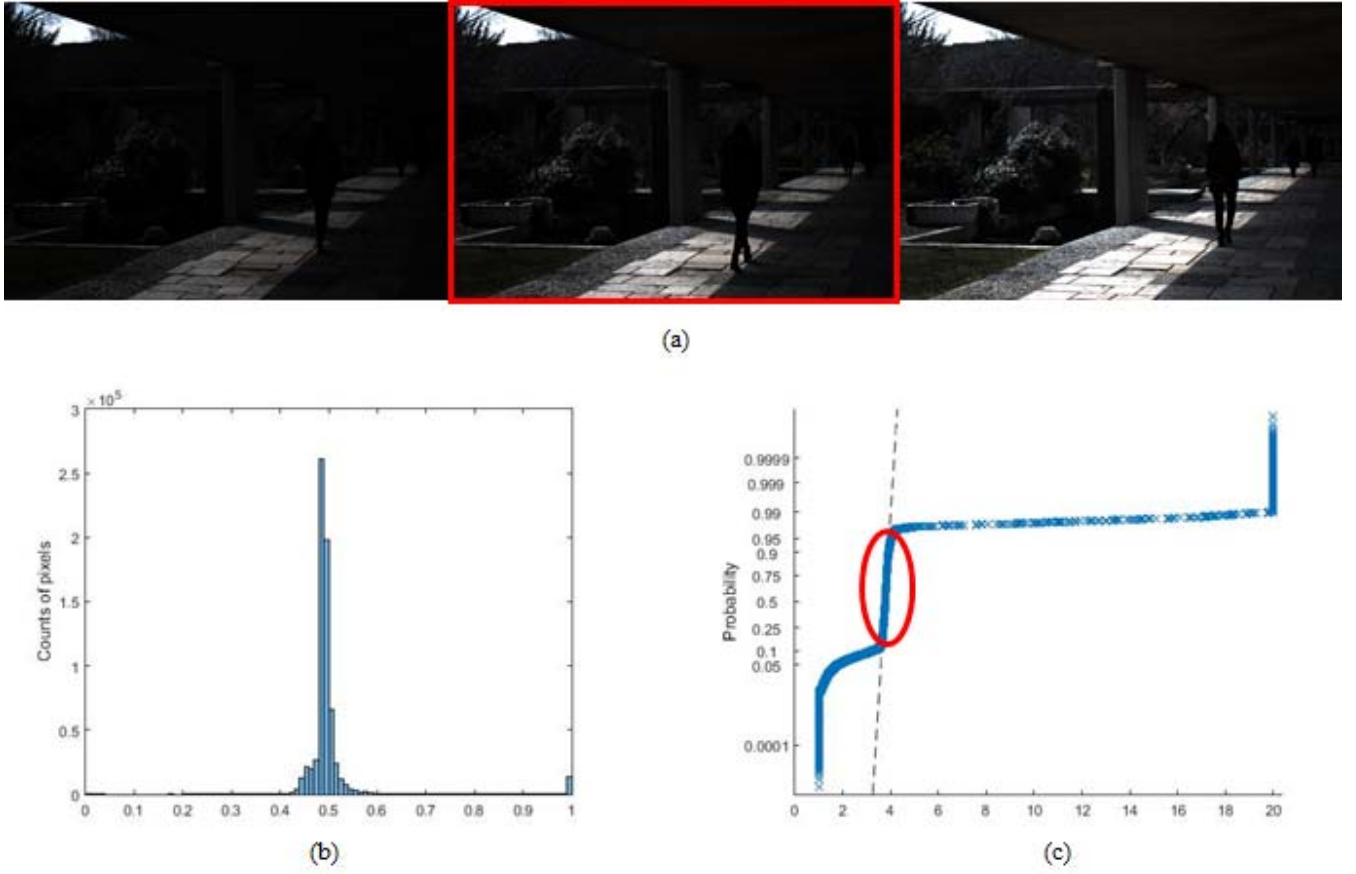


Figure 1. Analysis for the distribution of pixels' local intensity scaling value $\hat{\alpha}_{ref}^n$ between target image \hat{L}_n and reference image L_{ref} for 'Yard' scene. (a) Sequence of LDR images L_1 , L_{ref} , L_2 where $\Delta t_1 / \Delta t_{ref} = 0.5$ and $\Delta t_2 / \Delta t_{ref} = 4$; (b) Histogram of $\hat{L}_1 / L_{ref} (= \hat{\alpha}_{ref}^1)$ values; (c) Cumulative probability of $\hat{L}_2 / L_{ref} (= \hat{\alpha}_{ref}^2)$ values

In such a case, the scaling factor α_{ref}^n can be estimated from the ratio of two corresponding pixel intensity values in the observed images $L_n(p)$ and $L_{ref}(p)$. However, if a pixel p is a moving pixel or its intensity value is saturated, the computed ratio will not reflect the correct scaling factor between the images.

Here, we propose a maximum voting scheme to determine the scaling factor accurately, nullifying the effect of the moving pixels. We define the pixels set comprising the pixels of the same intensity in the reference image as $P_{ref}^i = \{p \mid L_{ref}(p) = i\}$. The intensity values of non-moving pixels in this set P_{ref}^i are expected to be $\alpha_{ref}^n \cdot i$ in the target image L_n as in (4) and moving pixels will have various different intensity values.

For the pixel locations $p \in P_{ref}^i$ in the target image L_n , we compute the representative value $\hat{L}_n(p)$ by taking median of their intensity values as

$$\hat{L}_n(p) = \text{median}(I_n^i), \quad I_n^i = \{L_n(p) \mid p \in P_{ref}^i\} \quad (6)$$

Assuming uniform distribution of noisy intensity values and that not all the pixels are affected by motion, the intensity values of all non-moving pixels will be the same and ranked in the middle when the intensity values in I_n^i are sorted in ascending order. Therefore, the median value of I_n^i can be considered to be the intensity value for non-moving pixels in I_n^i . We define the local scaling factor $\hat{\alpha}_{ref}^n$ as

$$\hat{\alpha}_{ref}^n = \frac{\hat{L}_n(p)}{L_{ref}(p)} \quad (7)$$

The computed $\hat{\alpha}_{ref}^n$ should be same for all intensity values i as mentioned in section 2.1. However, if majority in P_{ref}^i are moving pixels or saturated pixels (black or white), the computed $\hat{\alpha}_{ref}^n$ will have faulty values. Fig. 1(b) shows a typical histogram of \hat{L}_1 / L_{ref} values quantized to 100 bins where $\Delta t_1 / \Delta t_{ref} = 0.5$. The value of \hat{L}_1 / L_{ref} corresponding to maximum vote (frequency) is 0.49 which is very close to their expected value 0.5. Another way to visualize this is to plot the cumulative probability histogram. We have shown the cumulative probability distribution of a typical case of \hat{L}_2 / L_{ref} values where $\Delta t_2 / \Delta t_{ref} = 4$. More than 90 percent of pixels (red circle) have values very close to their expected value 4. Both of Fig. 1(b) and Fig. 1(c) show that the majority of pixels have scaling values the same as their expected scaling value, and they can be considered as non-moving pixels and the rest as moving pixels. Therefore, the final global scaling factor α_{ref}^n can be determined from non-moving pixels by selecting the value that has maximum vote as

$$\alpha_{ref}^n = \underset{\hat{\alpha}_{ref}^n}{\operatorname{argmax}} \operatorname{hist}(\hat{\alpha}_{ref}^n) \quad (8)$$

where $\operatorname{hist}(\hat{\alpha}_{ref}^n)$ is a histogram of $\hat{\alpha}_{ref}^n$ as in Fig. 1(b). For further validation of this classification scheme, we analyze

the accuracy of the estimated scaling factor α_{ref}^n in Table I on 9 different scenes from the dataset in [48]. The actual exposure time information for all the images in the dataset is available and can be used to calculate the true scaling factors, which are shown in the top row in the table.

TABLE I. ESTIMATED SCALING FACTOR FROM NON-MOVING AND NON-SATURATED PIXELS

Scene	$\Delta t_n / \Delta t_{ref}$					
	1/8	1/4	1/2	2	4	8
Yard	0.13	0.25	0.49	2	4	8
Cars	0.13	0.25	0.50	2	4	8
Building	0.13	0.26	0.51	2	4	8
Pedestrians	0.13	0.25	0.51	2	4	8
Museum1	0.13	0.26	0.50	2	4	8
StreetDay	0.13	0.27	0.49	2	4	8
Toytrain	0.13	0.25	0.49	2	4	7
Plants	0.13	0.27	0.48	2	4	8
Turtles	0.14	0.26	0.52	2	4	8

We compare them with the estimated scaling factors obtained by our method. The results show that the estimated values are almost the same as the real values for all 9 scenes.

C. Selection of moving pixels

Once α_{ref}^n is determined in (8), we use this scaling factor as a supporting index to determine the moving pixels. As shown in Fig. 1(b) and Fig. 1(c), majority of pixels have their local scaling factor $\hat{\alpha}_{ref}^n$ nearly the same as the global scaling factor α_{ref}^n with small variance. We consider those pixels which have their local scaling factors close to the global scaling factor as non-moving pixels and the others as moving pixels. In this section, the scaling factor α_{ref}^n computed in (8) is used as a guiding index to select the moving pixels. First, we define a pixel set P_{ref}^i collecting all the pixels in image L_{ref} and divide it into static pixels set P_{ref}^{i-s} and moving pixels set P_{ref}^{i-m} such that

$$P_{ref}^i = P_{ref}^{i-s} \cup P_{ref}^{i-m} \quad (9)$$

The pixels in the reference image can be classified into three categories – completely black who have intensity value $i = 0$, saturated who have intensity value $i = \max_i$, and the rest with intensity value $0 < i < \max_i$. Below we present our technique to label the pixels in each of these groups as static or moving.

1) For a pixel $p \in P_{ref}^i$ whose intensity value i is in between zero (minimum value, full black) and \max_i (maximum value, full white), its intensity value in the target image can be estimated as $\alpha_{ref}^n L_{ref}(p)$ from (4). Thus, if the estimated intensity value $\alpha_{ref}^n L_{ref}(p)$ is close to the observed intensity value $L_n(p)$, the pixel p in the reference image can be thought of as static pixel such that

$$P_{ref}^{i-s} = \left\{ p \mid L_{ref}(p) = i, \left| \frac{L_n(p)}{L_{ref}(p)} - \alpha_{ref}^n \right| \leq \varepsilon \cdot \alpha_{ref}^n \right\}$$

$$P_{ref}^{i-m} = \left\{ p \mid L_{ref}(p) = i, \left| \frac{L_n(p)}{L_{ref}(p)} - \alpha_{ref}^n \right| > \varepsilon \cdot \alpha_{ref}^n \right\} \quad (10)$$

where ε is predefined threshold. In most images of our

dataset, the pixels labelled as static pixels by the above method have their scaling factor within 10 percent of the global scaling factor. Fig. 1(c) shows that the static pixels in red circle are having very close values to its expected values. Therefore, ε was heuristically determined to be 0.1 in this work.

2) If the intensity value of a pixel $p \in P_{ref}^i$ is \max_i , the estimated intensity value in the target image also needs to be \max_i if $\alpha_{ref}^n > 1$. Thus, if the observed intensity value $L_n(p)$ is \max_i , the pixel p in the reference image can be thought of as static pixel. However, if $\alpha_{ref}^n < 1$, all the intensity values greater than $\alpha_{ref}^n \cdot \max_i$ can be the candidates for the estimated intensity value in the target image. We assume the pixel p in the reference image is static pixel if its observed intensity value $L_n(p)$ is any of these candidate values.

If $\alpha_{ref}^n > 1$,

$$P_{ref}^{i-s} = \{ p \mid L_{ref}(p) = \max_i, L_n(p) = \max_i \}$$

$$P_{ref}^{i-m} = \{ p \mid L_{ref}(p) = \max_i, L_n(p) \neq \max_i \} \quad (11)$$

If $\alpha_{ref}^n < 1$,

$$P_{ref}^{i-s} = \{ p \mid L_{ref}(p) = \max_i, L_n(p) \geq \alpha_{ref}^n \cdot \max_i - \varepsilon \}$$

$$P_{ref}^{i-m} = \{ p \mid L_{ref}(p) = \max_i, L_n(p) < \alpha_{ref}^n \cdot \max_i - \varepsilon \} \quad (12)$$

where ε is a tolerance factor. Again ε was set to 10 percent of the expected value ($= 0.1 \cdot \alpha_{ref}^n \cdot \max_i$).

3) If the intensity value of a pixel $p \in P_{ref}^i$ is zero, the estimated intensity value in the target image also needs to be zero if $\alpha_{ref}^n < 1$. However, if $\alpha_{ref}^n > 1$, all the intensity values where $L_n(p) / \alpha_{ref}^n$ is less than 0.5 can cause the intensity value of the reference image to be zero after rounding. Thus, the static pixels and the moving pixels can be determined as

If $\alpha_{ref}^n < 1$,

$$P_{ref}^{i-s} = \{ p \mid L_{ref}(p) = 0, L_n(p) = 0 \}$$

$$P_{ref}^{i-m} = \{ p \mid L_{ref}(p) = 0, L_n(p) \neq 0 \} \quad (13)$$

If $\alpha_{ref}^n > 1$,

$$P_{ref}^{i-s} = \{ p \mid L_{ref}(p) = 0, L_n(p) / \alpha_{ref}^n < 0.5 + \varepsilon \}$$

$$P_{ref}^{i-m} = \{ p \mid L_{ref}(p) = 0, L_n(p) / \alpha_{ref}^n \geq 0.5 + \varepsilon \} \quad (14)$$

where the tolerance factor ε was again set to 10 percent of 0.5 which is 0.05.

D. Final de-ghosting

Once the moving pixels are selected, the intensity values of the moving pixels lying on the target image are replaced by their expected background intensity values as

1) For pixels $p \in P_{ref}^i$ whose intensity is $0 < i < \max_i$:

$$L_n^{Deghosted}(p) = \begin{cases} L_n(p) & \text{if } p \in P_{ref}^{i-s} \\ \alpha_{ref}^n \cdot L_{ref}(p) & \text{if } p \in P_{ref}^{i-m} \end{cases} \quad (15)$$

2) For pixels $p \in P_{ref}^i$ whose intensity is $i = \max_i$:

$$L_n^{Deghosted}(p) = \begin{cases} L_n(p) & \text{if } p \in P_{ref}^{i-s} \\ \frac{\sum_{p \in P_{ref}^{i-s}} L_n(p)}{n(p \in P_{ref}^{i-s})} & \text{if } p \in P_{ref}^{i-m}, \alpha_{ref}^n < 1 \\ \max_i & \text{if } p \in P_{ref}^{i-m}, \alpha_{ref}^n > 1 \end{cases} \quad (16)$$

3) For pixels $p \in P_{ref}^i$ whose intensity is $i = 0$:

$$L_n^{Deghosted}(p) = \begin{cases} L_n(p) & \text{if } p \in P_{ref}^{i-s} \\ \frac{\sum_{p \in P_{ref}^{i-s}} L_n(p)}{n(p \in P_{ref}^{i-s})} & \text{if } p \in P_{ref}^{i-m}, \alpha_{ref}^n > 1 \\ 0 & \text{if } p \in P_{ref}^{i-m}, \alpha_{ref}^n < 1 \end{cases} \quad (17)$$

where $n(p \in P_{ref}^{i-s})$ is the total number of static pixels in the reference image having the intensity value i . For a pixel $p \in P_{ref}^i$ whose intensity value i is in between zero and \max_i , its intensity value in the target image can be replaced by its estimated value $\alpha_{ref}^n \cdot L_{ref}(p)$ computed from (4). For

pixels that are white (saturated), their expected value in a target image is also white if the target image has higher exposure time than the reference image ($\alpha_{ref}^n > 1$). But, if the reference image has higher exposure time than the target image ($\alpha_{ref}^n < 1$), all the intensity values in the target image that can be saturated in the reference image are candidates for the replacement value of moving pixels. Thus, the mean value of all the candidate values is chosen to replace the moving pixel. On the other hand, for the pixels which are black ($i = 0$), the replacement value of the moving pixels in the target image is 0 when $\alpha_{ref}^n < 1$ and the mean value of the candidates when $\alpha_{ref}^n > 1$. This procedure is repeated on each color channel separately and the results are merged together. If any of three color channel values of a pixel is saturated, that pixel is considered as a saturated pixel.



Figure 2. (a) Original LDR images with different exposures; (b) Reference image on top and target image at bottom; (c) De-ghosted target image with respect to reference image from the patch-based method [42]; (d) De-ghosted target image with respect to reference image from the proposed method

III. EXPERIMENTAL EVALUATION

We compared the proposed method with the patch-based method proposed by Sen *et al.* [42] which has shown one of the best de-ghosting results in recent comparative researches [15, 49], ghost bitmap-based method by Tursun *et al.* [15], and the commercial HDR software Photomatix Pro 6.1. The dataset in [48] is used for experiments.

Fig. 2 compares results of Sen *et al.* [42] and our algorithm on the dataset ‘Yard’. The images in this dataset taken at different exposures are shown at top in Fig. 2(a). It can be seen that some persons are moving and have different locations in each exposure. We take two images, one as reference and the other as target image shown in Fig. 2(b), and apply both algorithms to reconstruct the target image such that the moving objects in both reference and the target images are aligned. Results are shown in Fig. 2(c) and Fig. 2(d) respectively. It can be seen that the reconstructed image by Sen *et al.* [42] has ghost artefacts on head of the large person shown inside red rectangle. Moreover, the image produced by Sen *et al.* [42] is darker than the original target image and hence some details of background seem to be

lost. On the other hand, the proposed method does not produce the ghost artefacts at all and maintains the fine details and global intensity of the target image very well. All non-moving pixels in our reconstructed image remain unchanged, and only the moving pixels are shifted to match their positions in the reference image. Background is reconstructed perfectly at the previous positions of the moving pixels. The results clearly show superiority of the proposed method over Sen *et al.* [42]. It is clear that the moving objects are relocated and backgrounds are reconstructed perfectly. Therefore, when the output images produced by our method are merged to a HDR image, there will be no duplicate objects or ghost artefacts.

We demonstrate this in Fig. 3-6, which compare our method with three other methods mentioned earlier in this section, by showing the final constructed HDR images for four different scenes. For a fair comparison, we used the same HDR merging method as Photomatix for all, except for Tursun *et al.* [15] which uses its own merging technique in the provided implementation code.

Again, for the sake of fair comparison, we used the tone mapping technique of Photomatix for all four methods to

display the images. For ‘Yard’ scene in Fig. 3, the head part was blurred off in Sen *et al.* [42] and the background of top-left corner has white stains in Tursun *et al.* [15], and the fine details of a tree inside red rectangle is blurred in Photomatix. However, the proposed method produced ghost-free HDR image preserving all the fine details of the background.



Figure 3. Comparison of constructed HDR images of ‘Yard’ scene using Sen *et al.* [42] (top-left), Tursun *et al.* [15] (top-right), Photomatix (bottom-left), and the proposed method (bottom-right)



Figure 4. Comparison of constructed HDR images of ‘Museum1’ scene using Sen *et al.* [42] (top-left), Tursun *et al.* [15] (top-right), Photomatix (bottom-left), and the proposed method (bottom-right)



Figure 5. Comparison of constructed HDR images of ‘Pedestrians’ scene using Sen *et al.* [42] (top-left), Tursun *et al.* [15] (top-right), Photomatix (bottom-left), and the proposed method (bottom-right)

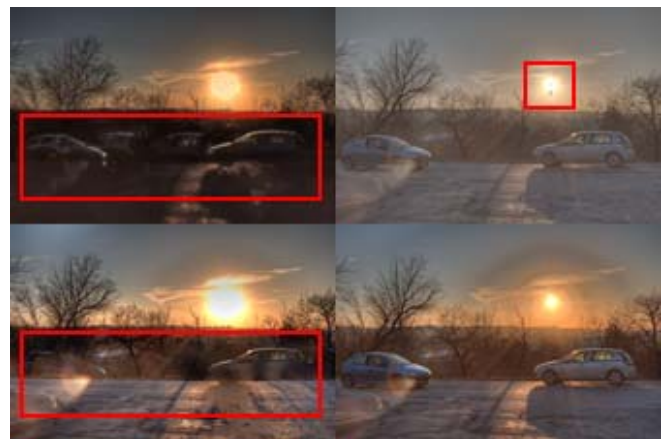


Figure 6. Comparison of constructed HDR images of ‘Cars’ scene using Sen *et al.* [42] (top-left), Tursun *et al.* [15] (top-right), Photomatix (bottom-left), and the proposed method (bottom-right)

For ‘Museum1’ scene in Fig. 4, most of the background is blurred in Sen *et al.* [42] and head of the person has ghost effect in Photomatix.

For ‘Pedestrians’ and ‘Cars’ scenes in Fig. 5 and Fig. 6, all three methods except the proposed method have ghost effects. The artefacts are easily visible on bodies of two persons and on cars, and the saturated region of sun is corrupted with stains in Tursun *et al.* [15].



Figure 7. Comparison of constructed HDR images of ‘Building’ scene using Sen *et al.* [42] (top-left), Tursun *et al.* [15] (top-right), Photomatix (bottom-left), and the proposed method (bottom-right)

For ‘Building’ scene in Fig. 7, the body of the person has ghost effect in Sen *et al.* [42] and Photomatix, while shadow of the person at the end of the aisle is lost in Tursun *et al.* [15]. Overall, the proposed method produced clean ghost-free HDR images that retain rich dynamic contents of the backgrounds for all the scenes. We have marked the regions with major artefacts with rectangles in the figures, and they can be seen more clearly by zooming in the view.

For objective analysis of the methods, the structural similarity index measure (SSIM) between the reference image and its final HDR image was compared in Table II.

Since the final HDR image is constructed with respect to the reference image, the higher SSIM value implies the better de-ghosting quality in HDR image. Table II shows the results from the proposed method have higher SSIM values than the other methods.

TABLE II. SSIM VALUES FOR DIFFERENT DE-GHOSTING METHODS

Scene	De-ghosting Methods		
	Sen et al.	Photomatrix	Proposed
Yard	0.2003	0.1623	0.2209
Cars	0.1826	0.1938	0.2174
Building	0.2303	0.2427	0.2802
Pedestrians	0.1854	0.2182	0.2408
Museum1	0.0748	0.0766	0.0848

The effect of the various predefined thresholds in (10) is analyzed in Table III. Since the pixel value of each target image is restored from the reference image if the difference between the estimated value and the observed value in target image is less than the threshold, higher threshold produce an HDR image which is more similar to the reference image, and thus produce higher SSIM values.

TABLE III. SSIM VALUES FOR DIFFERENT THRESHOLD IN (10)

Scene	Threshold (%)				
	5	10	15	20	25
Yard	0.2188	0.2209	0.2211	0.2211	0.2211
Cars	0.2153	0.2174	0.2178	0.2178	0.2178
Building	0.2746	0.2802	0.2802	0.2802	0.2802
Plants	0.2355	0.2408	0.2413	0.2413	0.2413
Turtles	0.0846	0.0848	0.0849	0.0849	0.0849

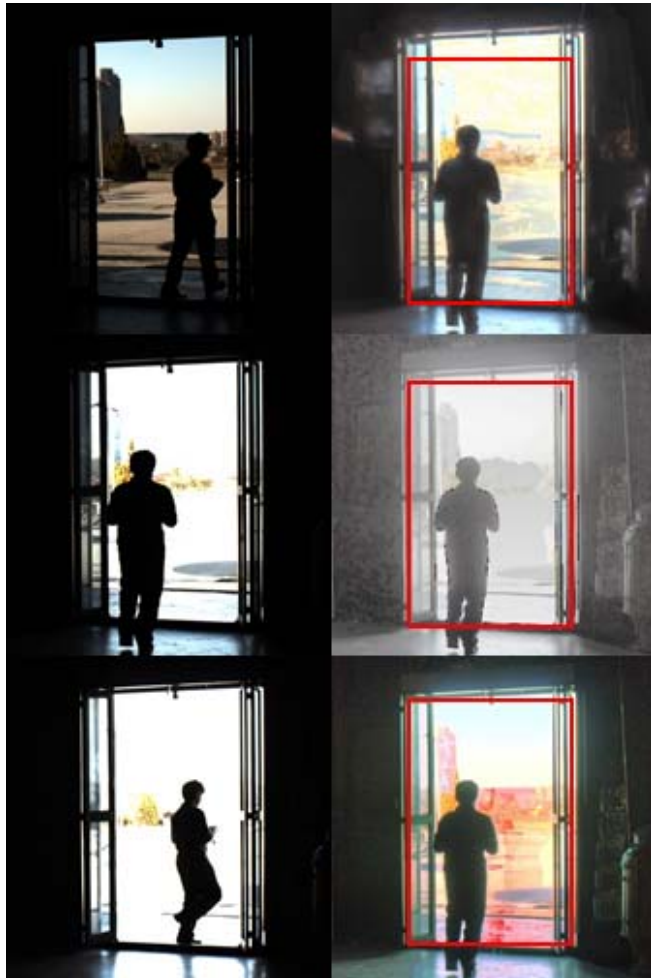


Figure 8. Left column: original LDR images of 'Exit' scene with different exposures (middle one is the reference image). Right column: constructed HDR images of 'Exit' scene using Sen et al. [42] (top), Tursun et al. [15] (middle), and the proposed method (bottom).

However, if higher threshold is used, less original

information of the target image is kept. Table III shows that the increase in SSIM values above the threshold 10 percent is negligible in most of the scenes, thus 10 percent was chosen for threshold to preserve maximum original information while keeping SSIM value near maximum.

Fig. 8 shows an example in which the proposed method failed to recover high quality HDR image. In this example, the reference image has the part where the background is fully saturated (outside the door). Result from the proposed method has ghost artefacts outside the door, but still better than the results from Sen *et al.* [42] and Tursun *et al.* [15] and the background information was recovered better in the proposed method. The proposed method tries to recover HDR content from the reference image by using the estimated scaling factor with respect to each target image. However, the computed scaling factor might not be accurate if large portion of reference image is full white or black, and this causes deterioration in the quality of final reconstructed HDR image.

IV. CONCLUSION

A de-ghosting algorithm based on the pixel intensity scaling factors was proposed to remove the artefacts caused by moving objects in HDR imaging. The intensity scaling factor which is the ratio of exposure times between two LDR images is hard to estimate if there are moving objects in the scene. A novel algorithm was proposed to estimate this value accurately in such situations, and use it as a cue to mark the pixels in the target images as moving or non-moving with respect to the reference image. The intensity values of the pixels classified as moving pixels in the target images were replaced with their expected values which were estimated again using the intensity scaling factors. Experimental evaluation of the proposed method in comparison to the existing state of the art methods was presented using several different scenes. The results showed that the proposed method can remove the ghost artefacts completely and retain the rich details of the scene in the constructed HDR image if the reference image does not have large portion of saturated area. In case the reference image contains saturated area where no background information is available, the quality of reconstructed HDR image is degraded, but still better in comparison to the existing methods.

REFERENCES

- [1] F. Banterle, A. Artusi, K. Debattista, A. Chalmers, *Advanced High Dynamic Range Imaging: Theory and Practice*. Boca Raton, FL, USA: CRC Press, 2017. doi: 10.1201/b11373
- [2] S. K. Nayar, T. Mitsunaga, "High dynamic range imaging: Spatially varying pixel exposures," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, pp. 472–479, June 2000. doi: 10.1109/CVPR.2000.855857
- [3] M. A. Robertson, S. Borman, R. L. Stevenson, "Estimation-theoretic approach to dynamic range enhancement using multiple exposures," *J. Electron. Imaging*, vol. 12, no. 2, pp. 219–229, April 2003. doi: 10.1117/1.1557695
- [4] Z. Li and J. Zheng, "Visual-saliency-based tone mapping for high dynamic range images," *IEEE Trans. Ind. Electron.*, vol. 61, no. 12, pp. 7076–7082, March 2014. doi: 10.1109/TIE.2014.2314066
- [5] B. K. Kim, R. H. Park, S. Chang, "Tone mapping with contrast preservation and lightness correction in high dynamic range imaging," *Signal, Image and Video Processing*, vol. 10, no. 8, pp. 1425–1432, July 2016. doi: 10.1007/s11760-016-0942-1
- [6] S. Ferradans, M. Bertalmio, E. Provenzi, V. Caselles, "An analysis of visual adaptation and contrast perception for tone mapping," *IEEE*

- Trans. Pattern Anal. Mach. Intell., vol. 33, no. 10, pp. 2002–2012, March 2011. doi: 10.1109/TPAMI.2011.46
- [7] K. Kim, J. Bae, J. Kim, "Natural HDR image tone mapping based on retinex", IEEE Trans. Consum. Electron., vol. 57, no. 4, pp. 1807–1814, November 2011. doi: 10.1109/tce.2011.6131157
- [8] I. R. Khan, S. Rahardja, M. M. Khan, M. M. Movania, F. Abed, "A tone-mapping technique based on histogram using a sensitivity model of the human visual system," IEEE Trans. Ind. Electron., vol. 65, no. 4, pp. 3469–3479, October 2017. doi: 10.1109/TIE.2017.2760247
- [9] C. Jung, T. Sun, "Optimized perceptual tone mapping for contrast enhancement of images," IEEE Trans. Circuits Syst. Video Technol., vol. 27, no. 6, pp. 1161–1170, February 2016. doi: 10.1109/TCSVT.2016.2527339
- [10] G. Eilertsen, R. K. Mantiuk, and J. Unger, "A comparative review of tone - mapping algorithms for high dynamic range video," Comput. Graph. Forum, vol. 36, no. 2, pp. 565–592, May 2017. doi : 10.1111/cgf.13148
- [11] G. Yue, C. Hou, T. Gu, S. Mao, W. Xiang, "Biologically inspired blind quality assessment of tone-mapped images," IEEE Trans. Ind. Electron., vol. 65, no. 3, pp. 2525–2536, March 2018. doi: 10.1109/TIE.2017.2739708
- [12] G. Yue, C. Hou, T. Zhou, "Blind quality assessment of tone-mapped images considering colorfulness, naturalness and structure," IEEE Trans. Ind. Electron., vol. 66 no. 5, pp. 3784–3793, July 2018. doi: 10.1109/TIE.2018.2851984
- [13] H. Yeganeh, Z. Wang, "Objective Quality Assessment of Tone-Mapped Images," IEEE Trans. Image Process., vol. 22, no. 2, pp. 657–667, February 2013. doi: 10.1109/TIP.2012.2221725
- [14] E. A. Khan, A. O. Akyüz, E. Reinhard, "Ghost Removal in High Dynamic Range Images," in Proc. IEEE Int. Conf. Image Process., pp. 2005–2008, October 2006. doi: 10.1109/ICIP.2006.312892
- [15] O. T. Tursun, A. O. Akyüz, A. Erdem, E. Erdem, "The state of the art in HDR deghosting: a survey and evaluation," Comput. Graph. Forum, vol. 34, no. 2, pp. 683–707, June 2015. doi: 10.1111/cgf.12593
- [16] M. Granados, H. P. Seidel, H. Lensch, "Background estimation from non-time sequence images," in Proc. Graphics & Interface, Canadian Information Processing Society, pp. 33–40, May 2008.
- [17] A. Srikantha, D. Sidibé, "Ghost detection and removal for high dynamic range images: Recent advances," Signal Process. Image., vol. 27, no. 6, pp. 650–662, July 2012. doi: 10.1016/j.image.2012.02.001
- [18] S. Silk, J. Lang, "Fast high dynamic range image deghosting for arbitrary scene motion," in Proc. Graphics & Interface, Canadian Information Processing Society, pp. 85–92, May 2012.
- [19] W. Zhang, W. K. Cham, "Gradient-directed composition of multi-exposure images," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., pp. 530–536, June 2010. doi: 10.1109/CVPR.2010.5540168
- [20] F. Pece, J. Kautz, "Bitmap movement detection: HDR for dynamic scenes," in IEEE Conf. Visual Media Production, pp. 1–8, November 2010. doi: 10.1109/CVMP.2010.8
- [21] O. Galo, N. Gelfandz, W. C. Chen, M. Tico, K. Pulli, "Artifact-free high dynamic range imaging," in IEEE Int. Conf. Computational Photography, pp. 1–7, April 2009. doi: 10.1109/ICCPHOT.2009.5559003
- [22] K. Jacobs, C. Loscos, G. Ward, "Automatic high-dynamic range image generation for dynamic scenes," IEEE Comput. Graph., vol. 28, no. 2, pp. 84–93, March 2008. DOI: 10.1109/MCG.2008.23
- [23] H. Y. Lin, W. Z. Chang, "High dynamic range imaging for stereoscopic scene representation," in Proc. IEEE Int. Conf. Image Process., pp. 4305–4308, November 2009. doi: 10.1109/ICIP.2009.5413665
- [24] D. K. Lee, R. H. Park, S. Chang, "Improved histogram based ghost removal in exposure fusion for high dynamic range images," in Proc. IEEE Int. Symposium Consum. Electron., pp. 586–591, June 2011. doi: 10.1109/ISCE.2011.5973898
- [25] T. H. Min, R. H. Park, S. Chang, "Histogram based ghost removal in high dynamic range images," in Proc. IEEE Int. Conf. Multimedia and Expo, pp. 530–533, July 2009. doi: 10.1109/ICME.2009.5202550
- [26] S. Wu, S. Xie, S. Rahardja, Z. Li, "A robust and fast anti-ghosting algorithm for high dynamic range imaging," in Proc. IEEE Int. Conf. Image Process., 2010, pp. 397–400, September 2010. doi: 10.1109/ICIP.2010.5654196
- [27] W. Zhang, W. K. Cham, "Reference-guided exposure fusion in dynamic scenes," J. Vis. Commun. Image. R., vol. 23, no. 3, pp. 467–475, April 2012. doi: 10.1016/j.jvcir.2012.01.006
- [28] T. H. Oh, J. Y. Lee, I. S. Kweon, "High dynamic range imaging by a rank-1 constraint," in Proc. IEEE Int. Conf. Image Process., pp. 790–794, September 2013. doi: 10.1109/ICIP.2013.6738163
- [29] M. Granados, K. I. Kim, J. Tompkin, C. Theobalt, "Automatic noise modeling for ghost-free HDR reconstruction," ACM Trans. Graphics, vol. 32, no. 6, p. 201, November 2013. doi: 10.1145/2508363.2508410
- [30] C. Wang, C. Tu, "An exposure fusion approach without ghost for dynamic scenes," in Proc. IEEE Int. Congress Image and Signal Processing, vol. 2, pp. 904–909, December 2013. doi: 10.1109/CISP.2013.6745293
- [31] K. R. Prabhakar, R. Arora, A. Swaminathan, K. P. Singh, R. V. Babu, "A fast, scalable, and reliable deghosting method for extreme exposure fusion," in IEEE Int. Conf. Computational Photography, pp. 1–8, May 2019. doi: 10.1109/ICCPHOT.2019.8747329
- [32] S. O. Shim, I. R. Khan, "Removal of ghosting artefacts in HDRI using intensity scaling cue," in SIGGRAPH Asia 2017 Technical Briefs, p. 16, November 2017. doi: 10.1145/3145749.3149433
- [33] Q. Yan, D. Gong, P. Zhang, Q. Shi, J. Sun, I. Reid, Y. Zhang, "Multi-Scale Dense Networks for Deep High Dynamic Range Imaging," in Proc. IEEE Winter Conf. Appl. Comput. Vis. (WACV), pp. 41–50, January 2019. doi: 10.1109/WACV.2019.00012
- [34] L. Bogoni, "Extending dynamic range of monochrome and color images through fusion," in Proc. IEEE Int. Conf. Pattern Recognit., vol. 3, pp. 7–12, September 2000. doi: 10.1109/ICPR.2000.903475
- [35] H. Zimmer, A. Bruhn, J. Weickert, "Freehand HDR imaging of moving scenes with simultaneous resolution enhancement," Comput. Graph. Forum, vol. 30, no. 2, pp. 405–414, April 2011. doi: 10.1111/j.1467-8659.2011.01870.x
- [36] T. Jinno, M. Okuda, "Multiple exposure fusion for high dynamic range image acquisition," IEEE Trans. Image Process., vol. 21, no. 1, pp. 358–365, June 2011. doi: 10.1109/TIP.2011.2160953
- [37] S. Ferradans, M. Bertalmio, E. Provenzi, V. Caselles, "Generation of HDR images in non-static conditions based on gradient fusion," in VISAPP, pp. 31–37, February 2012. doi: 10.5220/0003840700310037
- [38] D. Hafner, O. Demetz, J. Weickert, "Simultaneous HDR and optic flow computation," in Proc. IEEE Int. Conf. Pattern Recognit., pp. 2065–2070, August 2014. doi: 10.1109/ICPR.2014.360
- [39] S. Wu, J. Xu, Y. W. Tai, C. K. Tang, "Deep high dynamic range imaging with large foreground motions," in Proc. European Conf. Comput. Vis., pp. 117–132, September 2018. doi: 10.1007/978-3-030-01216-8_8
- [40] S. C. Park, H. H. Oh, J. H. Kwon, S. D. Lee, "Motion artifact-free HDR imaging under dynamic environments," in Proc. IEEE Int. Conf. Image Process., 2011, pp. 353–356, September 2011. doi: 10.1109/ICIP.2011.6116479
- [41] J. Hu, O. Gallo, K. Pulli, "Exposure stacks of live scenes with hand-held cameras," in Proc. European Conf. Comput. Vis., pp. 499–512, October 2012. doi: 10.1007/978-3-642-33718-5_36
- [42] P. Sen, N. K. Kalantari, M. Yaesoubi, S. Darabi, D. B. Goldman, E. Shechtman, "Robust patch-based HDR reconstruction of dynamic scenes," ACM Trans. Graph., vol. 31, no. 6, pp. 203–1, November 2012. doi: 10.1145/2366145.2366222
- [43] J. Hu, O. Gallo, K. Pulli, X. Sun, "HDR deghosting: How to deal with saturation," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., pp. 1163–1170, June 2013. doi: 10.1109/CVPR.2013.154
- [44] H. Farid, "Blind Inverse Gamma Correction," IEEE Trans. Image Process., vol. 10, no. 10, pp. 1428–1433, October 2001. doi: 10.1109/83.951529
- [45] T. Mitsunaga, S. K. Nayar, "Radiometric Self Calibration," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., vol. 1, pp. 374–380, June 1999. doi: 10.1109/CVPR.1999.786966
- [46] M. D. Grossberg, S. K. Nayar, "Determining the Camera Response from Images: What is Knowable?" IEEE Trans. Pattern Anal. Mach. Intell., vol. 25, no. 11, pp. 1455–1467, October 2003. doi: 10.1109/TPAMI.2003.1240119
- [47] M. D. Grossberg, S. K. Nayar, "What is the Space of Camera Response Functions," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., vol. 2, pp. II–602, June 2003. doi: 10.1109/CVPR.2003.1211522
- [48] O. T. Tursun, A. O. Akyüz, A. Erdem, E. Erdem, "An objective deghosting quality metric for HDR images," Comput. Graph. Forum, vol. 35, no. 2, pp. 139–152, May 2016. doi: 10.1111/cgf.12818
- [49] Y. Fang, H. Zhu, K. Ma, Z. Wang, "Perceptual quality assessment of HDR deghosting algorithms," in Proc. IEEE Int. Conf. Image Process., pp. 3165–3169, September 2017. doi: 10.1109/ICIP.2017.8296866