

RESEARCH

Open Access

Probabilistic motion pixel detection for the reduction of ghost artifacts in high dynamic range images from multiple exposures

Jaehyun An^{1†}, Seong Jong Ha^{2†} and Nam Ik Cho^{1*†}

Abstract

This paper presents an algorithm for compositing a high dynamic range (HDR) image from multi-exposure images, considering inconsistent pixels for the reduction of ghost artifacts. In HDR images, ghost artifacts may appear when there are moving objects while taking multiple images with different exposures. To prevent such artifacts, it is important to detect inconsistent pixels caused by moving objects in consecutive frames and then to assign zero weights to the corresponding pixels in the fusion process. This problem is formulated as a binary labeling problem based on a Markov random field (MRF) framework, the solution of which is a binary map for each exposure image, which identifies the pixels to be excluded in the fusion process. To obtain the ghost map, the distribution of zero-mean normalized cross-correlation (ZNCC) of an image with respect to the reference frame is modeled as a mixture of Gaussian functions, and the parameters of this function are used to design the energy function. However, this method does not well detect faint objects that are in low-contrast regions due to over- or under-exposure, because the ZNCC does not show much difference in such areas. Hence, we obtain an additional ghost map for the low-contrast regions, based on the intensity relationship between the frames. Specifically, the intensity mapping function (IMF) between the frames is estimated using pixels from high-contrast regions without inconsistent pixels, and pixels out of the tolerance range of the IMF are considered moving pixels in the low-contrast regions. As a result, inconsistent pixels in both the low- and high-contrast areas are well found, and thus, HDR images without noticeable ghosts can be obtained.

Keywords: Exposure fusion; High dynamic range image; Image fusion; Ghost artifacts

Introduction

The dynamic ranges of most commercial image sensors and display devices are narrower than the radiance range of an actual scene, and hence, under- or over-exposure is often inevitable. In order to overcome such limitations of image sensors and displays, a number of multi-exposure capturing and processing techniques have been proposed, which can be roughly categorized into two approaches: high dynamic range imaging (HDRI) with tone mapping [1-4] and image fusion methods [5-10]. The former generates an image of higher dynamic range (i.e., higher bit depth for each pixel) from multiple images having

different exposures. To obtain this image, the camera response function (CRF) must be known or estimated, and a tone mapping process is needed when showing the synthesized HDR image on a low dynamic range (LDR) display. On the other hand, the latter generates a tone-mapped-like high-quality image by the weighted addition of multiple exposure images and thus does not need CRF estimation, HDR image generation, and tone-mapping process. Hence, fusion approaches tend to require fewer computations than conventional HDRI, while providing comparable image quality for the LDR displays. Of course, HDRI is the more appropriate solution when showing images on HDR devices.

The conventional exposure fusion and the HDRI work well for the static scene when multi-exposure images are well registered and there is no moving object. But the

*Correspondence: nicho@snu.ac.kr

†Equal contributors

¹Department of Electrical and Computer Engineering, INMC, Seoul National University, Seoul, Korea

Full list of author information is available at the end of the article

ghost artifact is often observed in the HDR image from the dynamic scene where the images are not aligned and/or some objects are moving. Hence, there have been much efforts to alleviate the ghosting problem in the case of HDRI approaches. Some of the existing algorithms consider misalignment of input frames and moving objects simultaneously, while others assume well-aligned input or pre-registration of misaligned frames and concentrate on the detection of moving objects that cause inconsistency. For example, the study [11] exploits the measure of local entropy differences to identify regions that might contain moving pixels, which are then excluded from the HDRI generation process. In addition, Khan et al. [12] proposed an iterative method that gives larger weights to static and well-exposed pixels, thereby diminishing the weights for pixels that can cause ghosts. Li et al. [13,14] proposed methods to detect and modify moving pixels based on the intensity mapping function (IMF) [15]. There are also patch-based methods, in which patches including moving objects are excluded [16,17]. To simultaneously deal with the misalignment and moving objects, Zimmer et al. [18] proposed an optical flow-based energy minimization method, and Hu et al. [19] used non-rigid dense correspondence and color transfer function for this task. Recently, low-rank matrix-based algorithms [20,21] have also been presented, based on the assumption that irradiance maps are linearly related to LDR exposures.

In the case of exposure fusion, there are also similar approaches for ghost removal. For example, the median threshold bitmap approach was proposed to detect clusters of inconsistent pixels [22], which are then excluded when fusing the images. In addition, a gradient domain approach was introduced that gives smaller weights to inconsistent pixels [23]. The IMF is used to exclude region of inconsistent pixels in the fusion process [24,25], where the images are over-segmented and the IMF is used to detect the inconsistent regions. In our previous work [26], we proposed a method to detect inconsistent pixels based on a test of the reciprocity law of exposure and the measure of zero-mean normalized cross-correlation (ZNCC). It is noted that the ZNCC between a region in an image and its corresponding region in the reference is close to 1 when there is no moving object. Hence, a pixel is considered to be inconsistent when the region around the pixel shows low ZNCC under a certain threshold, i.e., the hard thresholding of ZNCC was used.

In this paper, we propose a probabilistic approach to constructing a ghost map, which is a binary image depicting the pixels to be excluded in the exposure fusion process. We assume that the images are well registered, otherwise apply a registration algorithm. The basic measure is also based on the ZNCC, but probabilistic soft

thresholding is used instead of the hard thresholding used in our previous work. Specifically, ZNCC histogram is modeled as a Gaussian mixture function, where the parameters are found using an expectation maximization (EM) algorithm. Generating a ghost map is then posed as a binary labeling problem based on a Markov random field (MRF) framework, where the energy to be minimized is designed as a function of the ZNCC distribution parameters. It will be shown that the proposed method provides a less noisy and more accurate binary map than the simple hard thresholding method.

However, as in other feature-based methods, the ZNCC shows meaningful differences only for well-contrasted and highly textured regions. Hence, feature-based methods often give incorrect results in low-contrast regions where the pixel values are about to be saturated due to over- or under-exposure and also in low-textured regions. For these regions, we exploit the IMF between the images, which was successfully used in [13,14,24,25]. In this paper, the IMF is estimated from regions having high ZNCC only, because other regions are saturated or moving object regions that have low credibility in estimating the IMF. Then, the pixels lying outside the IMF tolerance are considered pixels on the faint moving objects. To determine the ghost map in this region, we also develop an optimization technique, which yields less noisy results than conventional IMF-based thresholding methods. Experimental results show that the proposed method constructs plausible ghost maps and hence yields pleasing HDR images without noticeable ghost artifacts.

The rest of this paper is organized as follows: In the second section, we review the conventional weight map generation method [6]. In the third section, we describe the proposed algorithm that excludes the ghost pixels from the weight map. Then, we show some experimental results, and finally, conclusions are given in the last section.

Review of exposure fusion

Conventional exposure fusion methods create an output as a weighted sum of multiple exposure images, in which the weights reflect the quality of pixels in terms of contrast, saturation, and well-exposedness. The contrast is computed by Laplacian filtering [27], and the saturation is defined as the standard deviation of the pixels in each component. The measure of well-exposedness is designed to have the largest value when a pixel value is around the center of the dynamic range. The weight map for each exposure image is calculated by using these measures as

$$W_k(p) = (C_k(p))^{\omega_C} \times (S_k(p))^{\omega_S} \times (E_k(p))^{\omega_E} \quad (1)$$

where p is the pixel index; k means the k -th exposure image; C , S , and E represent contrast, saturation, and well-exposedness, respectively; and ω_C , ω_S , and ω_E are corresponding weighting factors. After the weighted images are added, multi-resolution blending is performed by using pyramidal image decomposition [28]. Figure 1 presents weight maps for the corresponding multi-exposure images, as well as the output as a result of the weighted sum of images. The ghost artifacts can be observed in the red box of Figure 1c, caused by moving people.

Proposed algorithm

The core of the ghost reduction algorithm is to find inconsistent pixels that can cause artifacts, for excluding them from the fusion process. For this task, we first determine a reference frame among the multi-exposure images, one that has the largest well-contrasted region. Then, in all other input frames except for the reference frame, we find regions that have moving objects with respect to the

reference. More specifically, we construct a *ghost map* (binary image) for each input frame except for the reference, which indicates which pixel to exclude or include in the image fusion process. When a pixel in the ghost map is 1, the corresponding pixel in the input frame will be included in the fusion process and vice versa.

The proposed method begins by finding the reference frame that has the largest well-contrasted region (i.e., smallest saturated region) as in conventional methods [16,24-26,29,30]. It needs to be noted that our method identifies the inconsistent pixels in high-contrast regions and low-contrast regions separately. For this, we define a *saturation map* \mathbf{b} , which is also a binary matrix with the size of input image. Note that this matrix can be constructed when finding the reference frame, because we check the contrast of regions at this time. Precisely, if we denote the element of \mathbf{b} at the pixel position p as $b(p)$, then it is given 1 when the p belongs to a well-contrast region in the reference frame and 0 when it belongs to the low-contrast region. In summary, for each of the input frames



Figure 1 An example of the exposure fusion. (a) A sequence of multi-exposure images. (b) Weight maps for the corresponding input images. (c) Fused image where ghost effect can be found in the red box due to moving people.

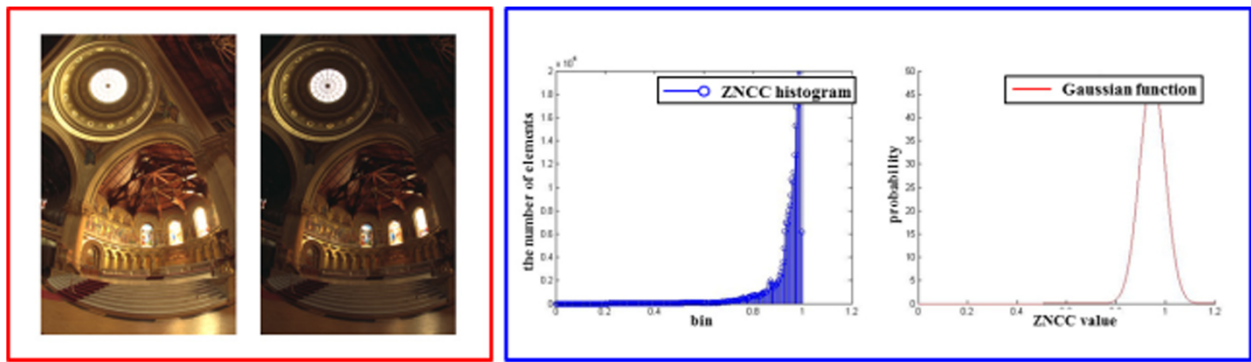


Figure 2 A pair of static images without moving objects (left) and the histogram of ZNCC for this set of images (right).

except the reference, we find the ghost maps for the region of $b(p) = 0$ and $b(p) = 1$ separately. The ghost map for the well-contrast region ($b(p) = 1$) will be denoted as \mathbf{g}_w and the ghost map for the low-contrast region ($b(p) = 0$) as \mathbf{g}_l in the rest of this paper. After finding these ghost maps for an input frame, the overall ghost map for the frame is constructed as $\mathbf{g} = \mathbf{g}_w \cup \mathbf{g}_l$. In this paper, finding \mathbf{g}_w and \mathbf{g}_l are posed as binary labeling problems, i.e., as the energy minimization problems that are solved by graph cuts [31].

Construction of \mathbf{g}_w

The construction of \mathbf{g}_w is based on the ZNCC measure, from the observation that the ZNCC of the region containing a moving object is low when compared with the ZNCC in the static region. The energy function for this binary labeling problem is defined as

$$E_W(\mathbf{g}_w) = \sum_{p \in P_W} E_{W_D}(\mathbf{g}_w(p)) + \gamma_W \times \sum_{(p,q) \in N_W} E_{W_S}(\mathbf{g}_w(p), \mathbf{g}_w(q)) \quad (2)$$

where $\mathbf{g}_w(p)$ is the pixel value (1 or 0) at a pixel p , P_W is the set of pixels in the well-contrasted region (all the p 's with $b(p) = 1$), N_W is the set of all unordered pair of neighboring pixels over the areas of $b(p) = 1$, and γ_W is a weighting factor for balancing the data cost E_{W_D} and the smoothness cost E_{W_S} .

The data cost E_{W_D}

The ZNCC of a region R centered at a pixel p , with respect to the corresponding region of the reference image, is defined as

$$Z(p) = \frac{\sum_{p \in R} D_{\text{ref}}(p) \times D(p)}{\sqrt{\sum_{p \in R} D_{\text{ref}}(p)^2 \times \sum_{p \in R} D(p)^2}} \quad (3a)$$

where

$$D_{\text{ref}}(p) = I_{\text{ref}}(p) - \bar{I}_{\text{ref}}(p) \quad (3b)$$

$$D(p) = I(p) - \bar{I}(p) \quad (3c)$$

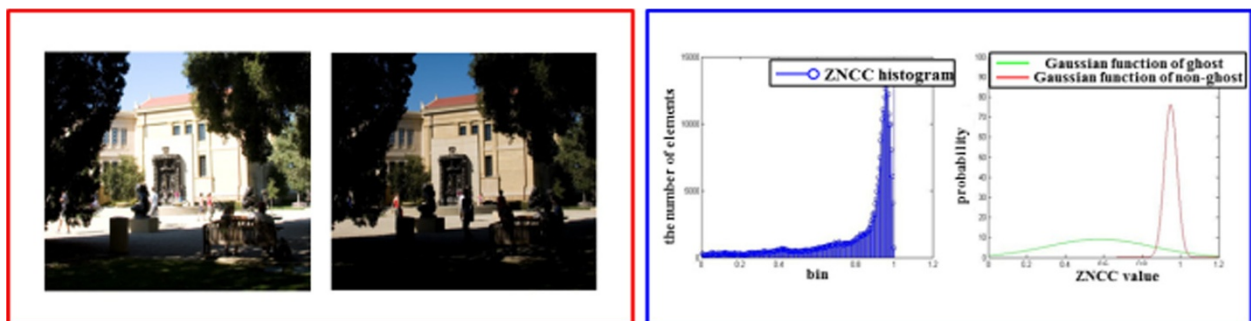


Figure 3 A pair of dynamic images with moving people (left) and the histogram of ZNCC for this set of images (right).

where I_{ref} and I mean the reference and a given frame respectively, and \bar{I}_{ref} and \bar{I} represent the mean values of I_{ref} and I in the region R , respectively. When there is no moving object in the scene, the histogram of ZNCC usually appears like Figure 2 because the ZNCC is close to 1 at most pixels. Hence, the histogram of ZNCC can be modeled as a left-sided normal distribution with the mean close to 1. On the other hand, Figure 3 shows another pair of images that capture a scene with some moving objects. Since the ZNCC becomes very small at the pixels of moving objects, the ZNCC distribution becomes multi-modal, which can be considered a mixture model consisting of two or more Gaussian distributions. Since the state of a pixel is just two (moving pixel or not) in our problem, we model the distribution as a sum of two Gaussian function as

$$p_r(\mathbf{x}) = \sum_{i=1}^2 p_r(\mathbf{x}|i)P_r(i) \quad (4)$$

where i is the state and input data \mathbf{x} is the ZNCC values. From the learning by EM algorithm, we find the parameters of two Gaussian density functions such as mean μ_i , variance σ_i , and weight $P_r(i)$.

With these models and parameters, the data term E_{W_D} is designed to give penalty to the mis-labeled pixels (e.g., labeled 1 while it is close to ghost). Specifically, the data cost is constructed as a negative of log-likelihoods of two Gaussian density function as

$$E_{W_D}(\text{'ghost'}) = -\ln Pr(x_p|G) \quad (5a)$$

$$E_{W_D}(\text{'non-ghost'}) = -\ln Pr(x_p|NG) \quad (5b)$$

where

$$Pr(x_p|G) = \frac{1}{\sqrt{2\pi}\sigma_1} \exp\left(-\frac{(x_p - \mu_1)^2}{2\sigma_1^2}\right) \quad (5c)$$

$$Pr(x_p|NG) = \frac{1}{\sqrt{2\pi}\sigma_2} \exp\left(-\frac{(x_p - \mu_2)^2}{2\sigma_2^2}\right). \quad (5d)$$

The smoothness cost E_{W_S}

The smoothness cost E_{W_S} is designed by Potts model [32] as

$$E_{W_S}(g_w(p), g_w(q)) = B(p, q) \times \delta(g_w(p), g_w(q)) \quad (6a)$$

where

$$\delta(g_w(p), g_w(q)) = \begin{cases} 1, & \text{if } g_1(p) \neq g_1(q) \\ 0, & \text{otherwise} \end{cases} \quad (6b)$$

$$B(p, q) = \exp\left(-\frac{(I(p) - I(q))^2}{2\sigma^2}\right) \quad (6c)$$

where $B(p, q)$ means the edge cue which represents pixel intensity difference. When the adjacent pixels are



Figure 4 A comparison of ghost map with hard thresholding method [26]. (a) Reference image. (b) A differently exposed image where people appear on the left region. (c) Ghost map by hard thresholding method [26]. (d) Ghost map by the proposed method.

bordering the edge, the ‘smoothness cost’ is diminished by $B(p, q)$.

For each of the input frames, the binary map \mathbf{g}_w is found by constructing and minimizing Equation 2. For example, Figure 4a,b shows the same area of multi-exposure images, where Figure 4a shows the crop of the reference frame and Figure 4b shows the crop of an under-exposed image where there appears moving people. The binary map resulting from the above equation is shown in Figure 4d, where the white area (labeled as 1) is the non-ghost area and the dark area (labeled as 0) contains motion pixels that are to be excluded. Figure 4c shows a binary map obtained with our previous method in [26], and the comparison with Figure 4d shows that the proposed method gives a more accurate ghost map.

Construction of \mathbf{g}_l

The ghost map \mathbf{g}_w for the well-contrasted region is found from the above procedure, and now we find the ghost map \mathbf{g}_l for the low-contrast region (for the regions with $b(p) = 0$). The problem with the low-contrast region is that there are too little textures to apply the feature-based methods (such as median pixel value [22], gradient [23], and ZNCC). Hence, we resort to intensity relationship between the frames for detecting the motion pixels in

these areas. The basic idea is that the static area shows the intensity changes according to the amount of exposure difference, whereas the areas with motion pixels will not follow that. In other words, the static area will have the luminance changes according to the IMF, whereas the dynamic areas will not.

Based on the above observation, we design the energy function for finding \mathbf{g}_l as

$$E_L(\mathbf{g}_l) = \sum_{p \in P_L} E_{L_D}(g_l(p)) + \gamma_L \times \sum_{(p,q) \in N_L} E_{L_S}(g_l(p), g_l(q)) \quad (7)$$

where $g_l(p)$ is the pixel value (1 or 0) at a pixel p , P_L is the set of pixels in the low-contrast region (all the p 's with $b(p) = 0$), N_L is the set of all unordered pair of neighboring pixels over the areas of $b(p) = 0$, and γ_L is a weighting factor for balancing the data cost E_{L_D} and the smoothness cost E_{L_S} . The smoothness cost E_{L_S} that prevents noisy result is defined the same as Equations 6a and 6c, except that g_w is replaced by g_l .

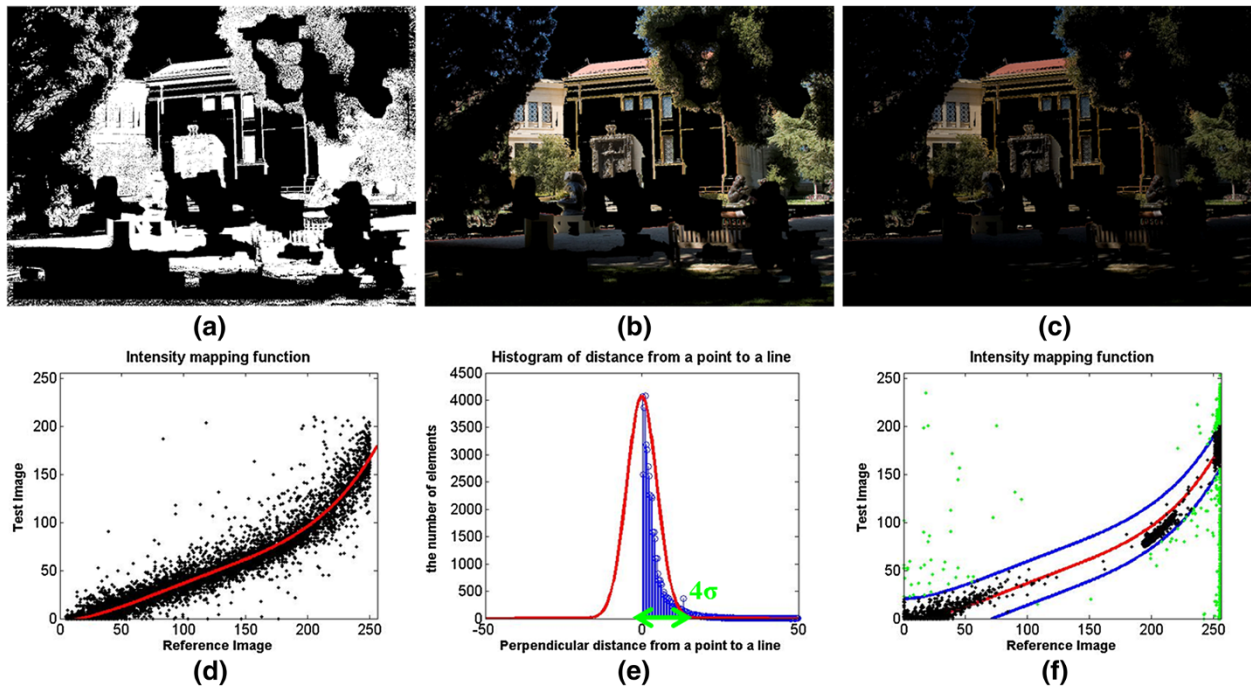


Figure 5 Estimation of IMF between a given exposure image and the reference. **(a)** A binary map where the white pixel denotes $g_w(p) = 1$ (well contrasted and static). **(b)** Overlap of the reference image with the binary map. **(c)** Overlap of comparing image with the binary map. **(d)** Each dot represents a pair of pixel values (only for the pair of colored pixels of (b) and (c)), and the red line is the estimated IMF by curve fitting. **(e)** Histogram of the distances of dots from the IMF. **(f)** The dots out of 4σ range are considered to include ghost pixels.

The data cost E_{LD}

As stated above, we use the ‘compliance of IMF’ for detecting the moving pixels in the low-contrast region, and thus, we have to estimate the IMF. In this paper, unlike the existing IMF estimation methods in [24,25] which use all the pixels without considering the pixel quality, we use the pixels only in the areas of ‘high-contrast region without moving objects,’ which correspond to the region of $g_w(p) = 1$ for a given image.

The algorithm for estimating the IMF is graphically shown in Figure 5, where Figure 5a shows the binary map $g_w(p)$. That is, the white pixels denote the ones with $g_w(p) = 1$. Figure 5b shows the overlap of this map with the reference frame, i.e., the pixel-wise multiplication of the map in Figure 5a and the reference image. Likewise,

Figure 5c shows the multiplication of the map in Figure 5a with the given input image to be compared with the reference. Finally, the IMF is estimated by comparing only the colored pixels of Figure 5b,c. Figure 5d shows the plot of the pairs of pixel values from these colored regions of Figure 5b,c, and the red line is considered the IMF, which is obtained by curve fitting the dots by the fourth-order polynomial [24,25]. Then, we define the tolerance range (upper and lower blue lines in Figure 5) of intensity variation from the IMF, which means that a dot out of this range is a pair of pixels where one of the pixels possibly belongs to a moving object. Figure 5e shows how this range is determined. Specifically, it shows the histogram of the distances of dots from the IMF curve, and a dot out of 4σ range (out of blue lines in Figure 5f) is considered the ghost pixel pair.

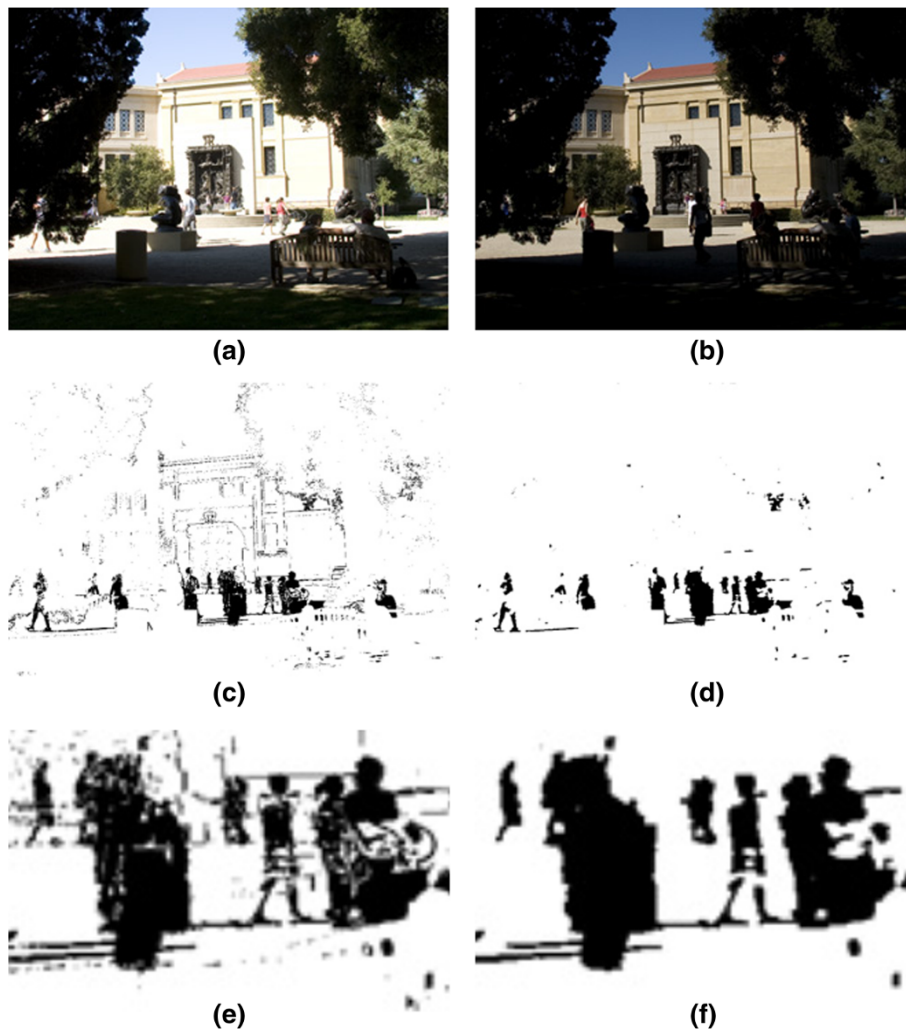


Figure 6 Ghost maps obtained by the IMF-based methods. (a) Reference image. (b) Differently exposed image. (c) Ghost map by hard thresholding method [29]. (d) Ghost map by the proposed method. (e) Magnification of (c). (f) Magnification of (d).

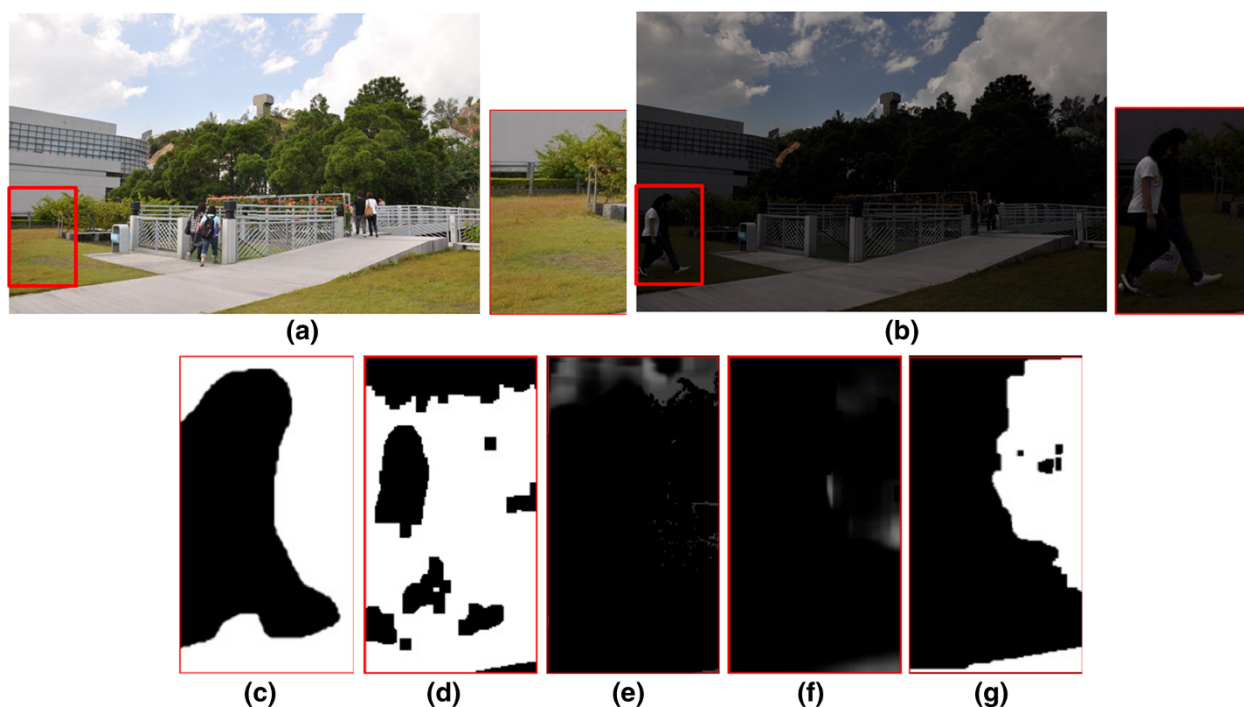


Figure 7 Comparison of the overall ghost maps. (a) Reference image and the magnification of the red box region. (b) Under-exposed image and the magnification of the red box region. (c) Ground truth ghost map for the red box region. (d) Motion map by [22]. (e) Consistency map by [23]. (f) Ghost map by [26]. (g) Ghost map by the proposed method.

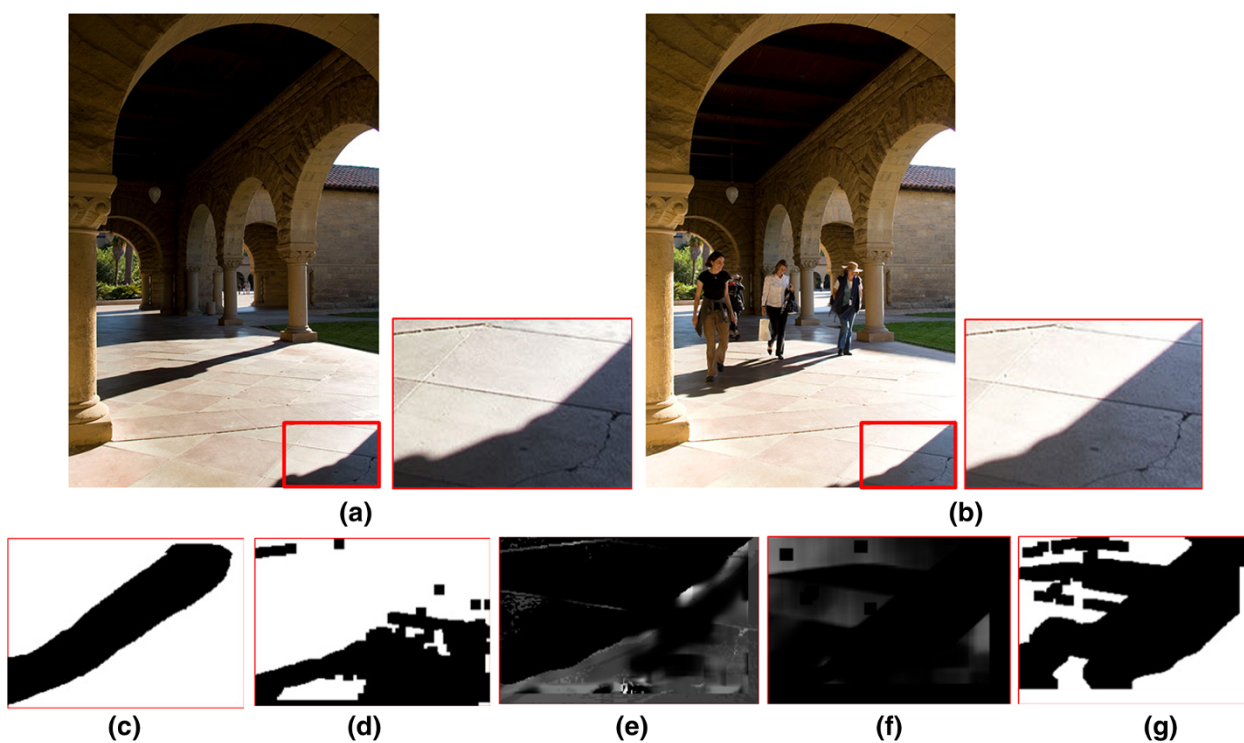


Figure 8 Another comparison of ghost maps. (a) Reference image and the magnification of red box region. (b) Over-exposed image and its magnification in the red box. (c) Ground truth ghost map for the red box region. (d) Motion map by [22]. (e) Consistency map by [23]. (f) Ghost map by [26]. (g) Ghost map by our method.

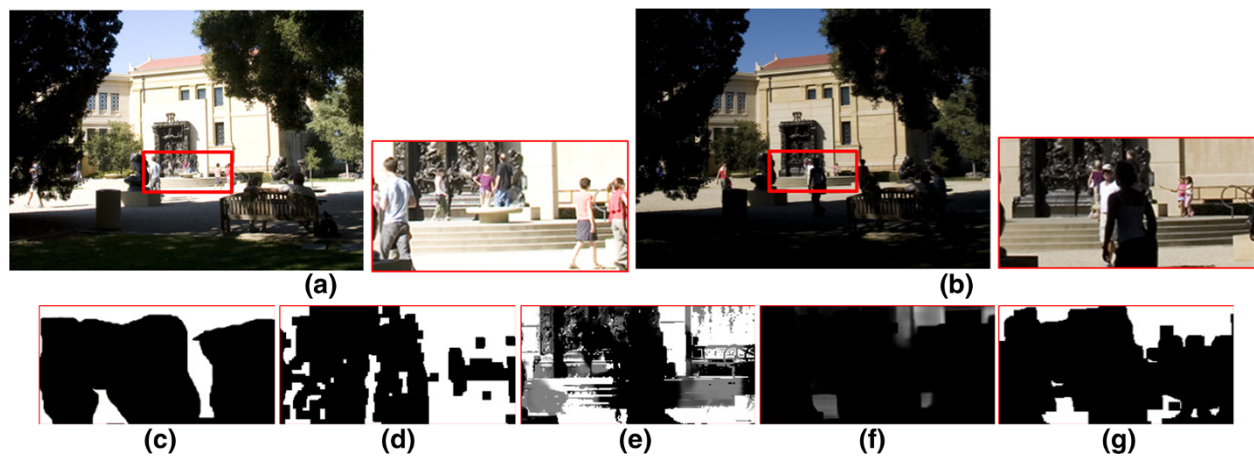


Figure 9 Comparison of ghost maps for other set of images. (a) Reference image and the magnification of red box region. **(b)** Under-exposed image and its magnification in the red box. **(c)** Ground truth ghost map for the red box region. **(d)** Motion map by [22]. **(e)** Consistency map by [23]. **(f)** Ghost map by [26]. **(g)** Ghost map by the proposed method.

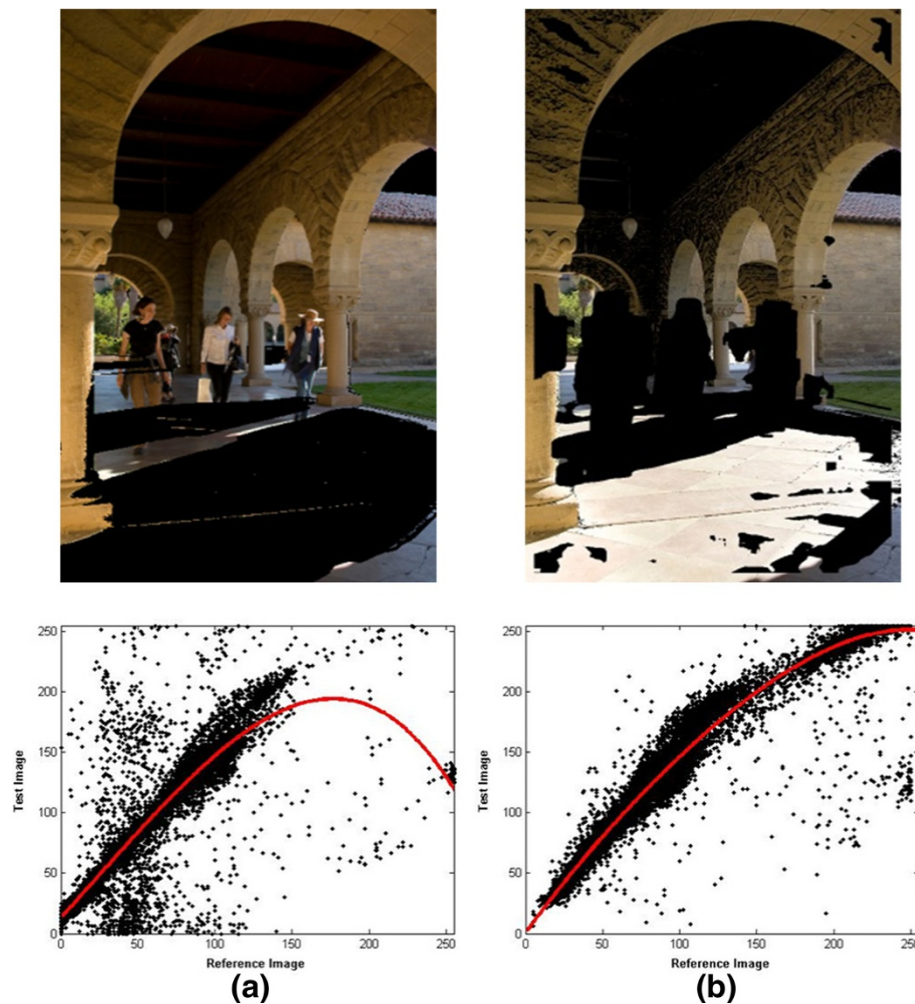


Figure 10 Comparison with Raman and Chaudhuri [25]. (a) (top) Detected static region of over-exposed source image in Figure 8b by [25] and (bottom) estimation of IMF (the red line which is obtained by curve fitting the dots by the fourth-order polynomial) using the static region. **(b)** The result by the proposed method.

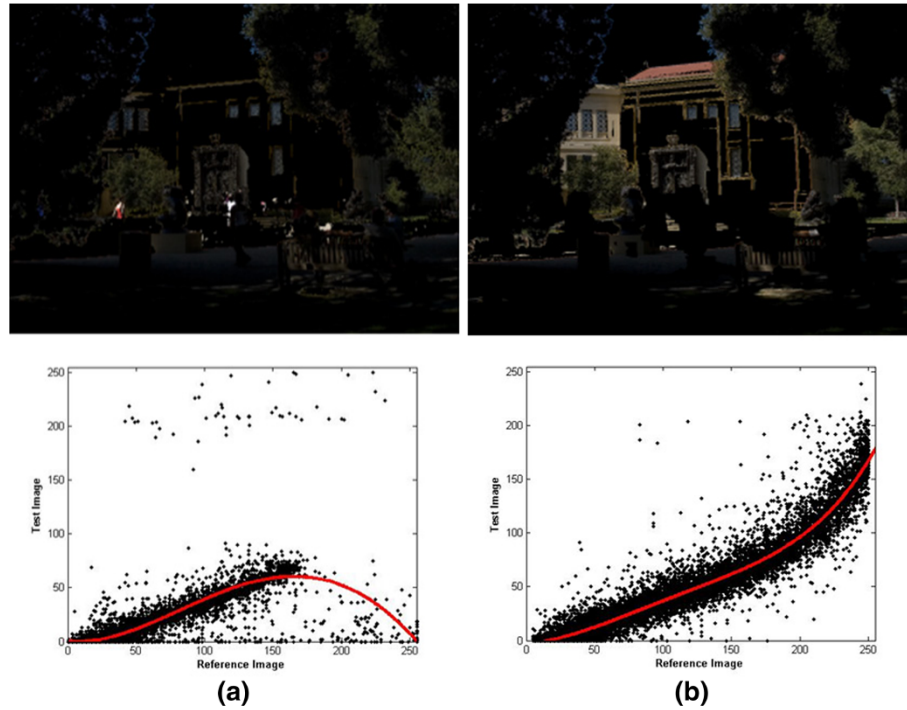


Figure 11 Another comparison with Raman and Chaudhuri [25]. **(a)** Detected static region of under-exposed source image in Figure 9b by [25] and (bottom) estimation of IMF (the red line which is obtained by curve fitting the dots by the fourth-order polynomial) using the static region. **(b)** Result by the proposed method.

Based on the estimated IMF, we design the data cost of energy function as

$$E_{LD}(g_I(p)) = \begin{cases} 0, & \text{if } \{g_I(p) = 0 \cap G(p) > 4\sigma\} \\ & \text{or } \{g_I(p) = 1 \cap G(p) \leq 4\sigma\} \\ 1, & \text{otherwise} \end{cases} \quad (8a)$$

where

$$G(p) = |I(p) - \text{IMF}(I_{\text{ref}}(p))| \quad (8b)$$

where $I(p)$ is the intensity of pixel p of the given frame and $\text{IMF}(I_{\text{ref}}(p))$ is the mapping of $I_{\text{ref}}(p)$ according to the IMF. Then, minimizing the total energy by graph cuts gives a binary map \mathbf{g}_I for the given input frame. Figure 6 compares the binary maps that represent ghost pixels as black, which are generated by hard thresholding method [29] and the proposed optimization method. Figure 6a,b shows the reference image and a differently exposed image, respectively, and Figure 6c,d shows the ghost maps of the thresholding and probabilistic methods, respectively. Also, Figure 6e,f shows the crops of the above images for better comparison.

It can be observed that the optimization leads to less map.

Experimental results

In the previous section, we have seen that each of the steps provides a less noisy and/or more plausible ghost map than the hard thresholding methods [26,29]. In this section, we compare the accuracy of the overall ghost map $\mathbf{g} = \mathbf{g}_w \cup \mathbf{g}_I$ with those of the existing methods. First, we compare the accuracy of the ghost map with [22,23,26], for the pair of the reference image in Figure 7a and a differently exposed image in Figure 7b. It can be seen that people appear in Figure 7b in the red box area. Hence, the ground truth ghost map for this area should be like in Figure 7c, and the ghost maps produced by [22,23,26] and the proposed method appear in Figure 7d,e,f,g, respectively. As can be observed in the figures, the median threshold bitmap approach [22] fails to detect the moving object, and the gradient domain method [23] determines the non-ghost region as a ghost region. The reason for this seems that the brightness/gradient difference between the moving object and the background is too small. Our previous work [26] used weight factor for ZNCC measure to extremely diminish ghost effect, so this method also regards the non-ghost pixels as ghost. On the other hand,

the ghost map produced by the proposed algorithm is closest to the ground truth. Figures 8 and 9 show another results that the proposed method provides better ghost map.

Figures 10 and 11 show the comparison of [25] and the proposed method for detecting the static region and estimating the IMF between the reference and a given frame. Figure 10a shows the static region detected by [25] (the dynamic region is denoted as black and only the static region remains) for the pair of images in Figure 8a,b and also the estimated IMF from this result. It can be seen that the moving people (dynamic region) are not removed and conversely the floor is detected as moving pixels. In the case of the result of our method (Figure 10b), it can be seen that dynamic regions are successfully detected, and hence, we can obtain more plausible IMF that should be a monotonically increasing function [15]. Figure 11

shows a similar result for the multi-exposure images in Figure 9a,b.

Figures 12,13,14 show the comparison of final fusion results for the eight existing methods [16,17,20-23,26,30] and the proposed method. Specifically, Figure 12a shows the sequence of multi-exposure images, and Figure 12b,c,d shows the comparisons of [16,22] and [26] with the proposed method. The first row of these figures shows the area for the comparison in the red boxes, the second row are the results of these compared methods in the order [16,22,26], and the bottom row shows our results for the corresponding areas. In these figures, it can be seen that the compared methods show some noticeable ghosts, whereas the proposed method does not. Figure 12f shows the overall area of fusion result of our method, and Figure 12e shows the result by [20]. Comparison shows that our method yields comparable

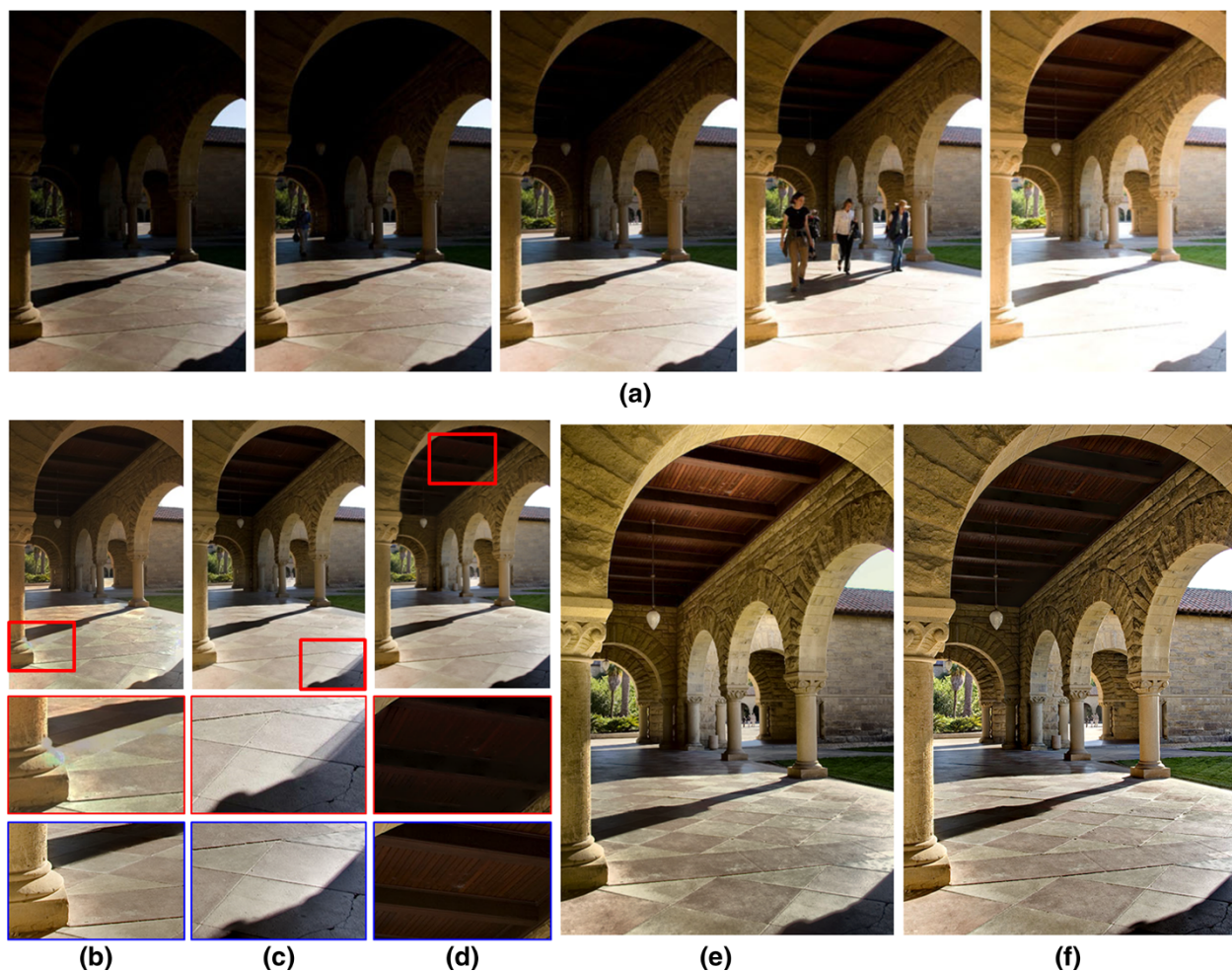


Figure 12 Comparison of results for a set of multi-exposure images. (a) Multi-exposure images. **(b)** Result by the patch-based algorithm based on the HDRI method [16]. **(c)** Result by the median threshold bitmap approach [22]. **(d)** Result by our previous method [26]. **(e)** Result by the low-rank matrix-based approach [20]. **(f)** Result by the proposed method.



Figure 13 Another comparison of results for a set of multi-exposure images. **(a)** Multi-exposure images. **(b)** Result by the patch-based algorithm based on the HDRI method [16]. **(c)** Gradient domain approach [23]. **(d)** PatchMatch-based method [30]. **(e)** Low-rank approach in [20]. **(f)** Low-rank matrix-based approaches in [21]. **(g)** Hybrid patch-based approach [17]. **(h)** The proposed method.

output as the HDRI approach by Oh et al. [20], which also yields almost no noticeable artifacts for the given image sequence. Likewise, Figure 13 shows the comparison of [16,20,21,23,30] and [17] with the proposed method. The photos of Figure 13a are the input images, and the second rows of Figure 13b,c,d,e show the results by [16,23,30] and [20], respectively, where the images in the third row are the results of the proposed method at the same area. It can be observed that the method of Gallo et al. [16] removes ghost successfully; however, when the difference of brightness among the neighboring patch is large, this causes some visible seam as shown in Figure 13b. In Figure 13c,d,e, we can see ghost artifacts in the existing methods, whereas the artifact is not noticeable in the case

of our algorithm (third row). Also, Figure 13f,g,h shows the results of [17,21] and the proposed method for this set of images, where the ghost artifact is not noticeable in the overall area. Figure 14 shows another comparison of [22,23,30] and [21] with the proposed method. Figure 14a shows the set of input images, and the second row of Figure 14b,c,d,e shows the result of [22,23,30] and [21], respectively, in the red box area of images in the first row. The images in the third row are the results by the proposed method in the same area, and Figure 14f shows the result of our method in the overall area. It can be observed that the proposed method shows almost no ghost artifacts, while others have a little noticeable artifacts as in the second row.



Figure 14 Another comparison of results for a set of multi-exposure images. **(a)** Multi-exposure images. **(b)** Result by the median threshold bitmap approach [22]. **(c)** Gradient domain approach [23]. **(d)** PatchMatch-based method [30]. **(e)** Low-rank matrix-based approach [21]. **(f)** The proposed method.

It is noted that our method has some limitations when the selected reference has ‘saturated and moving’ foreground object. For example, Figure 15a shows a set of multi-exposure images where the third image is selected as the reference frame because it has the largest area of well-contrast region in the background. In this case, the proposed algorithm yields the fusion result as shown in Figure 15c because the foreground object is moving and hence excluded from the fusion process. On the other hand, since the algorithm in [17] sets the first image as the reference and tries to track the inconsistent pixels, it keeps the foreground object very well as shown in Figure 15b. When we wish to keep the contrast of foreground object, we have to select the reference manually in this case. If we also select the reference as the first frame, then we obtain the result shown in Figure 15d. Finally, it is worth to comment that the fusion results can be enhanced by any conventional histogram equalization or edge-preserving

enhancement method such as [33], like the HDRI performs tone mapping process for the optimal display of HDR on the LDR display devices. The executables for our algorithm and full resolution results with this post-processing are available at <http://ispl.snu.ac.kr/~jhahn/deghost/>, which are also available as Additional files 1 and 2.

Conclusions

We have proposed an HDR image fusion algorithm with reduced ghost artifacts, by detecting inconsistent pixels in the high-contrast and low-contrast regions separately. To detect inconsistent pixels in high-contrast areas, a ZNCC measure is used based on the observation that the ZNCC histogram displays a unimodal distribution in static regions, whereas it has a multimodal shape in dynamic regions. A cost function based on the parameters of these probability distributions is designed, whose

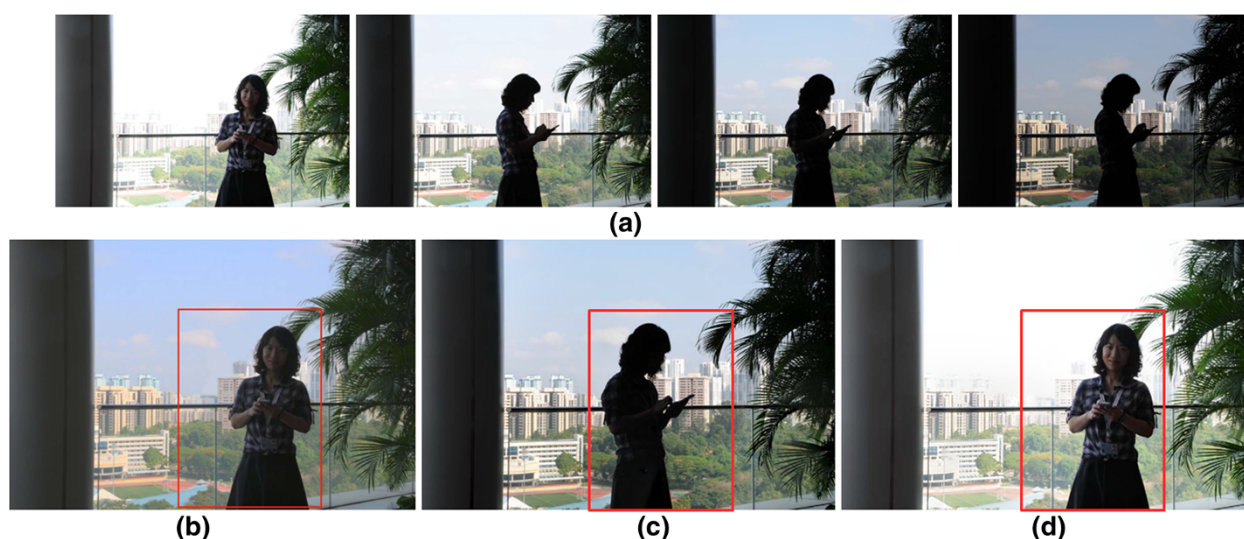


Figure 15 Another comparison of results. **(a)** Multi-exposure images. **(b)** Result by the hybrid patch-based approach [17]. **(c)** The proposed method. **(d)** The proposed method with the manual selection of the reference frame.

minimization yields the ghost map for the highly contrasted region. To detect the ghost map in the low-contrast region, the IMF is first estimated using pixels from the high-contrast regions having no moving objects. Next, a cost function that encodes the IMF compliance of the pixel pairs is designed, whose minimization gives the ghost map for the low-contrast areas. The overall ghost map is defined as the logical operation of these two maps, and the ghost pixels are excluded from the fusion process. Since the proposed algorithm can find faint moving objects in areas where the pixel values are about to be saturated due to over- and under-exposure, it provides satisfactory HDR outputs with no noticeable ghost artifacts. However, the proposed method has limitations in correcting moving foreground object when it is saturated in the reference frame (Figure 15), because they are simply excluded from the fusion process. In this case, we have to manually select a reference frame that has well-exposed foreground objects, which can degrade the fusion results due to the narrower well-exposed background region than in the reference. Otherwise, we need to correct the inconsistent pixels instead of simply excluding them, which is a very challenging problem, especially when the moving foreground object is not consistently detected in each frame due to saturation, noise, or non-rigid motion.

Additional files

Additional file 1: Results. This file contains the final full-size fusion results for several sets of multi-exposure images.

Additional file 2: Codes. This file contains MatLab codes for running the proposed algorithm.

Competing interests

The authors declare that they have no competing interests.

Acknowledgements

This research was supported by the Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Science, ICT and Future Planning (2009-0083495).

Author details

¹Department of Electrical and Computer Engineering, INMC, Seoul National University, Seoul, Korea. ²Samsung SDS, Seoul, Korea.

Received: 15 April 2014 Accepted: 07 August 2014

Published: 21 August 2014

References

1. PE Debevec, J Malik, Recovering high dynamic range radiance maps from photographs, in *Proceedings of the 24th Annual Conference on Computer Graphics and Interactive Techniques (SIGGRAPH'97)* (Los Angeles, Aug 1997), pp. 369–378
2. E Reinhard, G Ward, S Pattanaik, P Debevec, *High Dynamic Range Imaging: Acquisition, Display, and Image-Based Lighting (The Morgan Kaufmann Series in Computer Graphics)* (Morgan Kaufmann, San Francisco, 2005)
3. S Mann, R Picard, S Mann, RW Picard, On being 'undigital' with digital cameras: extending dynamic range by combining differently exposed pictures, in *Proceedings of the 48th IS&T's Annual Conference* (Washington, DC, May 1995), pp. 442–448
4. K Devlin, *A review of tone reproduction techniques. Technical report CSTR-02-005*. (Department of Computer Science, University of Bristol, 2002)
5. AA Goshtasby, Fusion of multi-exposure images. *Image Vis. Comput.* **23**(6), 611–618 (2005)
6. T Mertens, J Kautz, FV Reeth, Exposure fusion: a simple and practical alternative to high dynamic range photography. *Comput. Graph Forum.* **28**(1), 161–171 (2009)
7. S Raman, S Chaudhuri, A matte-less, variational approach to automatic scene compositing, in *IEEE International 11th Conference on Computer Vision* (Los Alamitos, Oct 2007), pp. 1–6
8. MH Malik, S Asif, M Gilani, Wavelet based exposure fusion, in *Proceedings of the World Congress on Engineering* (London, July 2008), pp. 688–693

9. S Raman, S Chaudhuri, Bilateral filter based compositing for variable exposure photography, in *Proceeding of Eurographics Short Papers* (Munich, Mar 2009), pp. 1–4
10. J Shen, Y Zhao, Y He, Detail-preserving exposure fusion using subband architecture. *Vis. Comput.* **28**(5), 463–473 (2012)
11. K Jacobs, C Loscos, G Ward, Automatic high-dynamic range image generation for dynamic scenes. *IEEE Comput. Graph. Appl.* **28**(2), 84–93 (2008)
12. E Khan, A Akyuz, E Reinhard, Robust generation of high dynamic range images, in *Proceedings of the IEEE International Conference on Image Processing* (Atlanta, Oct 2006), pp. 2005–2008
13. Z Li, S Rahardja, Z Zhu, S Xie, S Wu, Movement detection for the synthesis of high dynamic range images, in *Proceedings of the IEEE International Conference on Image Processing* (Hong Kong, Sept 2010), pp. 3133–3136
14. S Wu, S Xie, S Rahardja, Z Li, A robust and fast anti-ghosting algorithm for high dynamic range imaging, in *Proceedings of the IEEE International Conference on Image Processing* (Hong Kong, Sept 2010), pp. 397–400
15. MD Grossberg, SK Nayar, Determining the camera response from images: what is knowable. *IEEE Trans. Pattern Anal. Mach. Intell.* **25**(11), 1455–1467 (2003)
16. O Gallo, N Gelfand, W Chen, M Tico, K Pulli, Artifact-free high dynamic range imaging, in *IEEE International Conference on Computational Photography* (San Francisco, Apr 2009), pp. 1–7
17. J Zheng, Z Li, Z Zhu, S Wu, S Rahardja, Hybrid patching for a sequence of differently exposed images with moving objects. *IEEE Trans. Image Process.* **22**(12), 5190–5201 (2013)
18. H Zimmer, A Bruhn, J Weickert, Freehand HDR imaging of moving scenes with simultaneous resolution enhancement. *Comput. Graph. Forum.* **30**(2), 405–414 (2011)
19. J Hu, O Gallo, K Pulli, Exposure stacks of live scene with hand-held cameras, in *12th European Conference on Computer Vision* (Firenze, Oct 2012), pp. 499–512
20. T-H Oh, J-Y Lee, IS Kweon, High dynamic range imaging by a rank-1 constraint, in *IEEE International Conference on Image Processing* (Melbourne, Sept 2013), pp. 790–794
21. C Lee, Y Li, V Monga, Ghost-free high dynamic range imaging via rank minimization. *IEEE Signal Process. Lett.* **21**(9), 1045–1049 (2014)
22. F Pece, J Kautz, Bitmap movement detection: HDR for dynamic scenes, in *The 11th European Conference on Visual Media Production* (London, Nov 2010), pp. 1–8
23. W Zhang, W-K Cham, Gradient-directed composition of multi-exposure images, in *IEEE Conference on Computer Vision and Pattern Recognition* (San Francisco, June 2010), pp. 530–536
24. S Raman, S Chaudhuri, 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* (Chennai, Dec 2010), pp. 56–63
25. S Raman, S Chaudhuri, Reconstruction of high contrast images for dynamic scenes. *Vis. Comput.* **27**(12), 1099–1114 (2011)
26. J An, SH Lee, JG Kuk, NI Cho, A multi-exposure image fusion algorithm without ghost effect, in *IEEE International Conference on Acoustics, Speech, and Signal Processing* (Prague, May 2011), pp. 1565–1568
27. J Malik, P Perona, Preattentive texture discrimination with early vision mechanisms. *J. Opt. Soc. Am.* **7**(5), 923–932 (1990)
28. P Burt, E Adelson, The Laplacian pyramid as a compact image code. *IEEE Trans. Comm.* **31**(4), 532–540 (1983)
29. J An, SJ Ha, JG Kuk, NI Cho, Reduction of ghost effect in exposure fusion by detecting the ghost pixels in saturated and non-saturated regions, in *IEEE International Conference on Acoustics, Speech, and Signal Processing* (Kyoto, Mar 2012), pp. 1101–1104
30. J Hu, O Gallo, K Pulli, X Sun, HDR deghosting: how to deal with saturation?, in *IEEE Conference on Computer Vision and Pattern Recognition* (Portland, June 2013), pp. 1163–1170
31. YY Boykov, MP Jolly, Interactive graph cuts for optimal boundary & region segmentation of objects in n-d images, in *Proceedings of International Conference on Computer Vision* (Vancouver, July 2001), pp. 105–112
32. Y Boykov, O Veksler, R Zabih, Markov random fields with efficient approximations, in *IEEE Conference on Computer Vision and Pattern Recognition* (Santa Barbara, June 1998), pp. 648–655
33. K He, J Sun, X Tang, Guided image filtering. *IEEE Trans. Pattern Anal. Mach. Intell.* **35**(6), 1397–1409 (2013)

doi:10.1186/1687-5281-2014-42

Cite this article as: An et al.: Probabilistic motion pixel detection for the reduction of ghost artifacts in high dynamic range images from multiple exposures. *EURASIP Journal on Image and Video Processing* 2014 **2014**:42.

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Immediate publication on acceptance
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► springeropen.com