Selectively Detail-Enhanced Fusion of Differently Exposed Images With Moving Objects

Zhengguo Li, *Senior Member, IEEE,* Jinghong Zheng, *Member, IEEE,* Zijian Zhu, *Member, IEEE,* and Shiqian Wu, *Senior Member, IEEE*

Abstract—In this paper, we introduce an exposure fusion scheme for differently exposed images with moving objects. The proposed scheme comprises a ghost removal algorithm in low dynamic range domain and a selectively detail-enhanced exposure fusion algorithm. The proposed ghost removal algorithm includes a bidirectional normalization based method for the detection of non-consistent pixels and a two-round hybrid method for the correction of non-consistent pixels. Our detail-enhanced exposure fusion algorithm includes a content adaptive bilateral filter which extracts fine details from all the corrected images simultaneously in gradient domain. The final image is synthesized by selectively adding the extracted fine details to an intermediate image that is generated by fusing all the corrected images via an existing multi-scale algorithm. The proposed exposure fusion algorithm allows fine details to be exaggerated while existing exposure fusion algorithms do not provide such an option. The proposed scheme usually outperforms existing exposure fusion schemes when there are moving objects in real scenes. In addition, the proposed ghost removal algorithm is simpler than existing ghost removal algorithms and is suitable for mobile devices with limited computational resource.

Index Terms—Differently exposed images, Exposure fusion, Ghost removal, Gradient domain bilateral filter

1 Introduction

As digital camera functions penetrate into mobile phones, the fierce competition among mobile phone manufacturers pushes state-of-the-art image processing to a level that provides not just convenience but also high quality and natural output images. One of the challenges of digital image processing is the display of a natural scene which exhibits high dynamic range (HDR) on a conventional low dynamic range (LDR) display [1], [2]. To address this challenge, multiple differently exposed images are captured and fused together to produce a more detailed image [3], [4].

The challenge of exposure fusion is how to seamlessly merge the information of all input images together. As objects in the input images have obvious intensity gap, it is necessary for the exposure fusion algorithm to find a way to make the objects' intensity changes smoothly in the output images. Several interesting methods were proposed to address this problem. All input images were scaled into several down-sampled layers in [3] by using the Laplacian pyramid [5]. A novel weighting factor that was calculated by taking the luminance, contrast and color saturation of a pixel into account was introduced to blend each layer of the input images. The exposure fusion scheme in [6] was based on an observation that gradient magnitude becomes larger when a pixel gets a state of better exposed and it decreases gradually as the pixel approaches over- or under-exposure. The bilateral filter is adopted to decompose each input image into base and detail layers. The detail layers of all input LDR images are adopted to compute weighting maps for differently exposed LDR images. The complexity of exposure fusion scheme in [6] could be an issue because each image is decomposed individually via the bilateral filter. Although these methods can provide visually pleasing results, the output images often lack of detail information, as the smoothing effect of the Laplacian pyramid could result in the loss of fine details [7]. Two automatically video enhancement approaches were proposed in [8], [9] by fusing a stream of randomly-distorted images of a still scene. A gradient domain method was provided in [10] to first extract fine details from multiple differently exposed images simultaneously, and then add all the extracted details to an intermediate image that is fused by the algorithm in [3]. This method produced perceptually very pleasing results as more details are preserved in fused images. However, the complexity of the method in [10] could be an issue because it is a global optimization based method. Therefore, it is highly desirable to develop a simpler exposure fusion scheme that can exaggerates fine details as the local tone mapping algorithms in [11], [12].

The exposure fusion algorithms in [3], [6], [10] focused on static scenes. Ghosting artifacts which are generated by fusing foreground and background pixels together would be produced if they were directly applied to HDR scenes with moving objects. Many interesting ghost removal algorithms were proposed in LDR domain. An entropy based method was proposed in [13] by using the feature that local entropy is usually not changed much with respect to the exposure times. The local entropy is computed from the histograms constructed from pixels falling within a window with size 5×5. All pixels with local entropy variation larger than a threshold are marked as non-consistent pixels. As realized by the authors, this method could fail if two image regions share the same structure but with different intensity. The concept of median threshold bitmap in [13] was adopted in [14] to convert all

Zhengguo Li, Jinghong Zheng, Zijian Zhu and Shiqian Wu are with Department of Signal Processing, Institute for Infocomm Research, 1 Fusionopolis Way, 21-01 Connexis, Agency for Science, Technology and Research, Singapore 138632 ({ezgli,jzheng,zhuzj,shiqian}@i2r.a-star.edu.sg).

differently exposed LDR images into binary images. All pixels in a position are detected as non-consistent pixels if the sum of all pixels at the same position is neither 0 nor the total number of input images. This method is very simple and usually works well. However, when a moving object is always under/over exposed in all images, this method cannot produce satisfactory results. Another algorithm was proposed in [4] to estimate the probability of a pixel belonging to a moving object. The probability is incorporated in the weighting function to strongly attenuate the contribution of non-consistent pixels. Same as the algorithm in [15], the ghost removal algorithm in [4] is based on an assumption that stationary parts of an HDR scene are predominant in the sequence. Moreover, both schemes intend to remove entire moving objects from the final image. Recently, two attractive patch-match based methods were proposed in [16], [17]. The major advantage of these two methods is that they do not require input images to be well aligned or captured by a tripod camera as required by most existing ghost removal algorithms [4], [13], [15], [18], [19]. However, their complexity might be an issue for real time applications, especially for those applications on mobile devices with limited computational resource. For example, it was reported in [16] that their algorithm takes about 3 minutes to process 7 images with the spatial resolution as 1300×900. It is desirable to design an efficient ghost removal algorithm.

In this paper, we introduce an exposure fusion scheme which is composed of a ghost removal algorithm in LDR domain and a selectively detail-enhanced exposure fusion algorithm for differently exposed images with moving objects. As with the popular HDR acquisition approach in [1], [2], all the input images are assumed to be perfectly aligned, possibly using the registration algorithm in [13] or Photoshop CS 5. The proposed ghost removal algorithm consists of a detection module and a correction module. Because of different exposures, there are possibly large intensity changes between a corrected image and its reference image. Similar to our preliminary ghost removal algorithm in [20], differently exposed images are normalized in the correction module by utilizing a bidirectional method with the help of intensity mapping functions (IMFs) between them [21]. Our new detection module is simpler and more robust than the algorithm in [20]. The proposed correction module is a two-round hybrid correction approach. In the first round, the non-consistent pixels are corrected by using the IMF based correction method in [20]. In the second round, the spatial correlation in the corrected images is used to further correct the non-consistent pixels. The advantage of the proposed two-round correction is that more similar pixels in the corrected image are available to correct the non-consistent pixels in the second round. As a result, the quality of the corrected image can be improved. The corrected image serves as the reference image of its subsequent image. Moving objects are preserved and they become static in the corrected images as shown in Fig. 1. The corrected images can be regarded as differently exposed images of a static HDR scene. This is different from the schemes in [4], [15] which intend to remove all moving objects from the final image. Furthermore, due to the simplicity of the proposed ghost removal algorithm, it is friendly to mobile devices with limited computational resource.

A selectively detail-enhanced exposure fusion algorithm is then introduced to fuse all the corrected images. A unique feature of the proposed algorithm is to extract fine details from all the corrected images and selectively add them to an intermediate image. A key component of the detail extraction module is a content adaptive bilateral filter in gradient domain. Both spatial range parameter and range similarity parameter are spatially varying instead of being fixed in [22]. The proposed filter can preserve edges and smooth flat areas better than the filter in [22]. It extracts fine details from all the corrected images simultaneously while existing bilateral filters in [7], [23] can only extract fine details from a single input image. Instead of adding all the extracted fine details to the intermediate image as in the detail-enhanced exposure fusion algorithm in [10], the extracted fine details are selectively added to the intermediate image. With the proposed detail enhancement algorithm, fine details in all regions except flat ones can be amplified and added to the intermediate image. The proposed scheme is applicable to both dynamic and static HDR scenes. With our scheme, HDR photographers can take photos of dynamic HDR scenes without taking many pictures of the same scene and selecting frames which do not show moving objects as suggested by [24].

The rest of this paper is organized as follows. A new ghost removal algorithm is designed to detect and correct non-consistent pixels of differently exposed images in Section 2. A selectively detail-enhanced exposure fusion algorithm is provided in Section 3 to fuse all the corrected images into a more detailed image. Experimental results are presented in Section 4 to verify the efficiency of the proposed scheme. Finally, the conclusion remarks are listed in Section 5.

2 Ghost Removal of Differently Exposed Images

The proposed ghost removal algorithm is composed of two modules: a detection module and a correction module. Non-consistent pixels in all the input images are first detected by using the detection module. All the non-consistent pixels are then corrected by using the correction module. Subsequently, all pixels in all the corrected images are consistent.

2.1 Detection of Non-consistent Pixels

Let $\mathbf{Z}_k(\mathbf{p}) = [\mathbf{Z}_{k,1}(\mathbf{p}), \mathbf{Z}_{k,2}(\mathbf{p}), \mathbf{Z}_{k,3}(\mathbf{p})]^T$ denote the image intensity vector of the position $\mathbf{p}$ in the image $\mathbf{Z}_k$, i.e., $\mathbf{p}$ is a spatial position, $k(1 \leq k \leq N)$ indexes different exposures, and N is the total number of input LDR images. A middle image, denoted as $\mathbf{Z}_{k_0}$ is selected as the initial reference image according to the overall exposedness of luminance components of all input images. All pixels in the image $\mathbf{Z}_{k_0}$ are marked as consistent pixels. A pixel in the image $\mathbf{Z}_k(k \neq k_0)$ is marked as a consistent pixel if it is similar to the pixel at

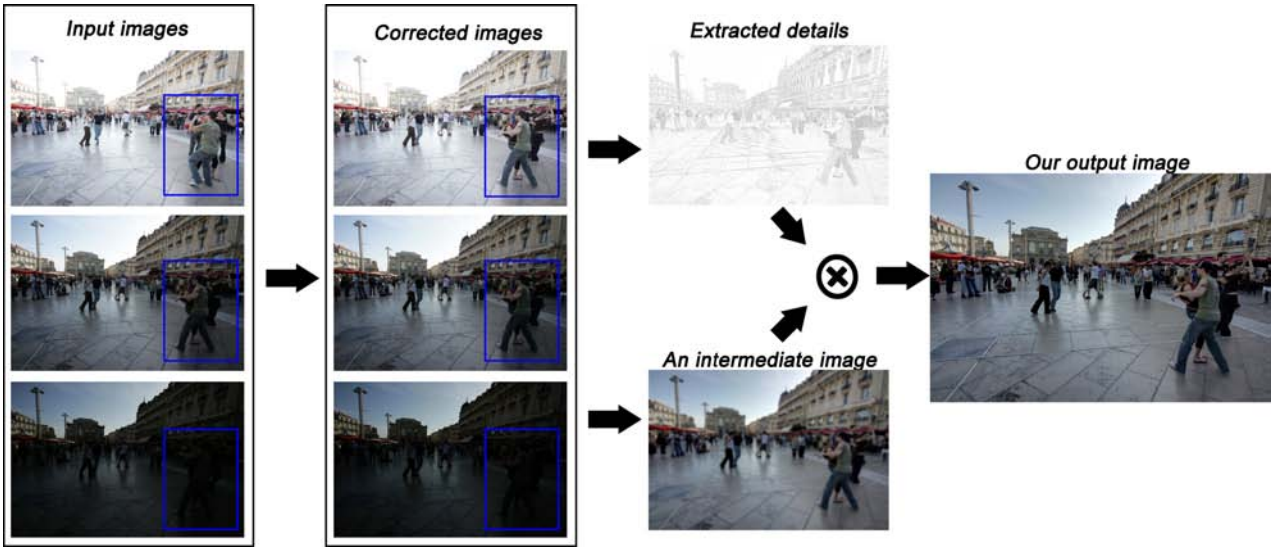


Fig. 1: *The overall structure of the proposed exposure fusion scheme. Image courtesy of Jacques Joffre.*

TABLE 1: *Notations used in this paper*

N	number of images
M	number of pixels in an image
$\mathbf{p}$	a pixel position, i.e., (x, y)
l	a color channel
m	a pyramid level
$\mathbf{p}_r$	$(x, y+1)$
$\mathbf{p}_b$	$(x+1, y)$
$\mathbf{p}'$	(x', y')
$\tilde{\mathbf{p}}$	(x', y)
$\mathbf{Z}_k$	the kth input image
$\nabla\mathbf{Z}_k(\mathbf{p})$	$(\frac{\partial\mathbf{Z}_k(\mathbf{p})}{\partial x}, \frac{\partial\mathbf{Z}_k(\mathbf{p})}{\partial y})$
$\mathbf{Y}_k$	luminance component of $\mathbf{Z}_k$
$\breve{\mathbf{Z}}, \tilde{\mathbf{Z}}$	normalized images
$\mathbf{C}_k$	consistent pixels in the image $\mathbf{Z}_k$
$\bar{\mathbf{C}}_k$	inconsistent pixels in the image $\mathbf{Z}_k$
$\mathbf{\Omega}_k$	under-exposed pixels in the image $\mathbf{Z}_k$
$\mathbf{\Lambda}$, $\hat{\mathbf{\Lambda}}$	intensity mapping functions
$\mathbf{\Pi}_k(\mathbf{p})$	indicate whether pixel $\mathbf{Z}_k(\mathbf{p})$ is consistent
$\Theta_\rho(\mathbf{p})$	a square window centered at the pixel $\mathbf{p}$ of a radius ρ
q	$q \in \{x, y\}$
$\mathbf{V}(\mathbf{p})$	vector field
$\mathbf{w}_1$, $\mathbf{w}_2$, $\mathbf{w}_g$ $\mathbf{w}_{c,k}$, $\mathbf{w}_{s,k}$, $\mathbf{w}_{e,k}$, $\mathbf{w}_{f,k}$	weighting functions
$\mathbf{L}^d$	a detail layer
$\mathbf{Z}_{int}$	an intermediate image

the same position in its reference image. For simplicity, the image $\mathbf{Z}_{k_0+1}$ is taken as an example to illustrate the details of detection and correction modules in this section. Without loss of generality, the exposure time of the image $\mathbf{Z}_{k_0+1}$ is assumed to be larger than that of the image $\mathbf{Z}_{k_0}$.

Due to different exposures between the images $\mathbf{Z}_{k_0}$ and $\mathbf{Z}_{k_0+1}$, there is possibly large intensity change between two collocated pixels $\mathbf{Z}_{k_0}(\mathbf{p})$ and $\mathbf{Z}_{k_0+1}(\mathbf{p})$. An IMF based bidirectional normalization method is adopted to model the intensity change between them. Let $\widehat{\mathbf{\Lambda}}_l(z)(0 \le z \le 255)$ and $\mathbf{\Lambda}_l(z)(0 \le z \le 255)$ be the IMFs from the image $\mathbf{Z}_{k_0,l}$ to the image $\mathbf{Z}_{k_0+1,l}$ and vice versa [21]. According to features of differently exposed images, the image $\mathbf{Z}_{k_0+1}$ includes less noise since the image $\mathbf{Z}_{k_0+1}$ is brighter than the image $\mathbf{Z}_{k_0}$ [25], [26]. It is thus desired to use the IMF $\mathbf{\Lambda}_l(z)$ to normalize the images $\mathbf{Z}_{k_0,l}$ and $\mathbf{Z}_{k_0+1,l}$ as much as possible. On the other had, the IMF $\mathbf{\Lambda}_l(z)$ is not accurate when the intensity $\mathbf{Z}_{k_0+1,l}(\mathbf{p})$ is over-exposed. Therefore, the pixels $\mathbf{Z}_{k_0}(\mathbf{p})$ and $\mathbf{Z}_{k_0+1}(\mathbf{p})$ are normalized via the following bidirectional method:

$$\breve{\mathbf{Z}}_l(\mathbf{p}) = \begin{cases} \mathbf{Z}_{k_0,l}(\mathbf{p}); & \text{if } \mathbf{Z}_{k_0+1,l}(\mathbf{p}) \le \eta \\ \mathbf{Z}_{k_0+1,l}(\mathbf{p}); & \text{otherwise} \end{cases}, \quad (1)$$

$$\tilde{\mathbf{Z}}_l(\mathbf{p}) = \begin{cases} \mathbf{\Lambda}_l(\mathbf{Z}_{k_0+1,l}(\mathbf{p})); & \text{if } \mathbf{Z}_{k_0+1,l}(\mathbf{p}) \le \eta \\ \widehat{\mathbf{\Lambda}}_l(\mathbf{Z}_{k_0,l}(\mathbf{p})); & \text{otherwise} \end{cases}, \quad (2)$$

where the value of η is empirically chosen as 216 for 8-bit images in this paper.

Clearly, the proposed bidirectional normalization method is simpler than the normalization method in [20]. The new method is also more robust than the method in [20] because the new method utilizes the feature of noise in differently exposed images: an image with a large exposure time is less noisy than an image with a small exposure time [25], [26].

The set of consistent pixels in the image $\mathbf{Z}_{k_0+1}$ is denoted as $\mathbf{C}_{k_0+1}$, and it is defined as:

$$\mathbf{C}_{k_0+1} = \{\mathbf{p} || \breve{\mathbf{Z}}_l(\mathbf{p}) - \tilde{\mathbf{Z}}_l(\mathbf{p})| < \text{Thr}(\mathbf{p})\breve{\mathbf{Z}}_l(\mathbf{p}), l = 1, 2, 3\}, \quad (3)$$

and $\bar{\mathbf{C}}_{k_0+1}$ is a set of non-consistent pixels in the image $\mathbf{Z}_k$. The value of $\text{Thr}(\mathbf{p})$ can be computed using the Richards curve [27], the Weber ratio [28], and the concept of just-noticeable-distortion [29]. More conservative thresholds may be required in some situations to prevent subtle ghosting from appearing in the final image. Our new detection method can be more accurate than the detection method in [20] because the Weber ratio [28] and the concept of just-noticeable-distortion [29] are applied to each color component independently.

Notice that only small portions of differently exposed images belong to non-consistent pixels. Instead of detecting non-consistent pixels by checking all pixels as in [20], a sub-

sampling based method can be adopted to detect moving objects by only checking part of pixels. If a pixel is detected as a consistent one, its neighboring pixels are skipped. Otherwise, its neighboring pixels are further checked by using a sub-sampling method. An example with a sub-sampling factor being selected as $4(= 2^{3-1})$ is adopted to illustrate the proposed fast movement detection scheme. The example is shown in Fig. 2. The pixels that are labeled with 1 are checked at the first round, and they are marked by the blue color if they are invalid. The neighboring pixels of a blue pixel labeled with 1 are then checked by using a sub-sampling method. In other words, all neighboring pixels that are labeled by 2 are checked at the second round. They are also marked by the blue color if they are invalid. Finally, the eight neighboring pixels of a blue pixel labeled by 2, i.e., those pixels are labeled by 3, are checked in the final round. Only 1/16 of all pixels are detected at the first round. Since only a small portion of pixels in an LDR image belongs to moving objects, the second and third rounds of detections are only conducted for a small amount of pixels in the LDR image. As such, the complexity of the detection module can be significantly reduced. Meanwhile, since many pixels are in the neighborhoods of two pixels, a flag is attached to each pixel so as to indicate whether it has been detected. With the flag, each pixel will only be detected once.

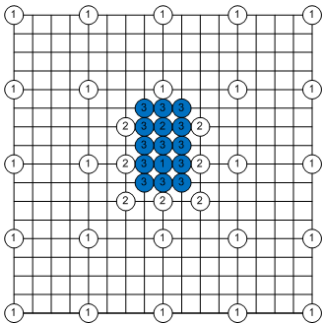


Fig. 2: *A sub-sampling based movement detection scheme.*

The detection results are used to initialize a weighting matrix $\mathbf{\Pi}_{k_0+1}$ and the value of $\mathbf{\Pi}_{k_0+1}(\mathbf{p})$ is 1 if the pixel $\mathbf{Z}_{k_0+1}(\mathbf{p})$ is a consistent pixel and 0 otherwise. To reduce the effects of noise, the value of $\mathbf{\Pi}_{k_0+1}(\mathbf{p})$ is filtered over a window of 3×3.

2.2 Correction of Non-consistent Pixels

After detecting all the non-consistent pixels in the image $\mathbf{Z}_{k_0+1}$, all the consistent pixels in the set $\mathbf{C}_{k_0+1}$ are kept and all the non-consistent pixels in the set $\bar{\mathbf{C}}_{k_0+1}$ are corrected by using a two-round hybrid correlation method.

Since the exposure time of the image $\mathbf{Z}_{k_0+1}$ is larger than that of the image $\mathbf{Z}_{k_0}$, the IMF $\widehat{\mathbf{\Lambda}}_l$ is accurate if the pixel $\mathbf{Z}_{k_0}(\mathbf{p})$ is not under-exposed. A reliable set is then defined as:

$$\mathbf{\Omega}_{k_0} = \{\mathbf{p} | \mathbf{Z}_{k_0,l}(\mathbf{p}) > 5, l = 1, 2, 3\}. \quad (4)$$

In the first round, all the non-consistent pixels in the set $(\mathbf{\Omega}_{k_0} \cap \bar{\mathbf{C}}_{k_0+1})$ are corrected by using the temporal correlation between the images $\mathbf{Z}_{k_0+1}$ and $\mathbf{Z}_{k_0}$. For a pixel $p \in (\mathbf{\Omega}_{k_0} \cap \bar{\mathbf{C}}_{k_0+1})$, the pixel $\mathbf{Z}_{k_0+1}(\mathbf{p})$ is replaced by the following synthesized pixel:

$$[\widehat{\mathbf{\Lambda}}_1(\mathbf{Z}_{k_0,1}(\mathbf{p})), \widehat{\mathbf{\Lambda}}_2(\mathbf{Z}_{k_0,2}(\mathbf{p})), \widehat{\mathbf{\Lambda}}_3(\mathbf{Z}_{k_0,3}(\mathbf{p}))]. \quad (5)$$

This is exactly the same as the correction module in [20].

In the second round, all the non-consistent pixels in the set $(\bar{\mathbf{C}}_{k_0+1} - \mathbf{\Omega}_{k_0} \cap \bar{\mathbf{C}}_{k_0+1})$ are corrected by using the spatial correlation in the image $\mathbf{Z}_{k_0+1}$ rather than using the IMF based method in [20]. This is because the IMFs from the image $\mathbf{Z}_{k_0}$ to the image $\mathbf{Z}_{k_0+1}$ are not accurate for all the non-consistent pixels in the set $(\bar{\mathbf{C}}_{k_0+1} - \mathbf{\Omega}_{k_0} \cap \bar{\mathbf{C}}_{k_0+1})$. Due to this, it is much more challenging to correct the non-consistent pixels in the set $(\bar{\mathbf{C}}_{k_0+1} - \mathbf{\Omega}_{k_0} \cap \bar{\mathbf{C}}_{k_0+1})$ than the non-consistent pixels in the set $(\mathbf{\Omega}_{k_0} \cap \bar{\mathbf{C}}_{k_0+1})$.

Given a pixel $\mathbf{p}$ in the set $(\bar{\mathbf{C}}_{k_0+1} - \mathbf{\Omega}_{k_0} \cap \bar{\mathbf{C}}_{k_0+1})$, a corrected or consistent pixel $\mathbf{Z}_{k_0+1}(\mathbf{p}')$ is said to be similar to the pixel $\mathbf{Z}_{k_0+1}(\mathbf{p})$ if the value of $\mathbf{Z}_{k_0,l}(\mathbf{p}')$ is exactly the same as that of $\mathbf{Z}_{k_0,l}(\mathbf{p})$ for all l's. Let $\Theta_\rho(\mathbf{p})$ is a square window centered at the pixel $\mathbf{p}$ of a radius ρ. The value of ρ is selected as 4 provided that its value is specified in this paper. The value of $\mathbf{Z}_{k_0+1,l}(\mathbf{p})$ is corrected as the average value of all similar $\mathbf{Z}_{k_0+1,l}(\mathbf{p}')$ in the set $\Theta_\rho(\mathbf{p})$.

The objective to correct all the non-consistent pixels in two rounds is to guarantee that there are more candidates in the second round to correct the non-consistent pixels in the set $(\bar{\mathbf{C}}_{k_0+1} - \mathbf{\Omega}_{k_0} \cap \bar{\mathbf{C}}_{k_0+1})$. As a result, the non-consistent pixels in the set $(\bar{\mathbf{C}}_{k_0+1} - \mathbf{\Omega}_{k_0} \cap \bar{\mathbf{C}}_{k_0+1})$ are better corrected.

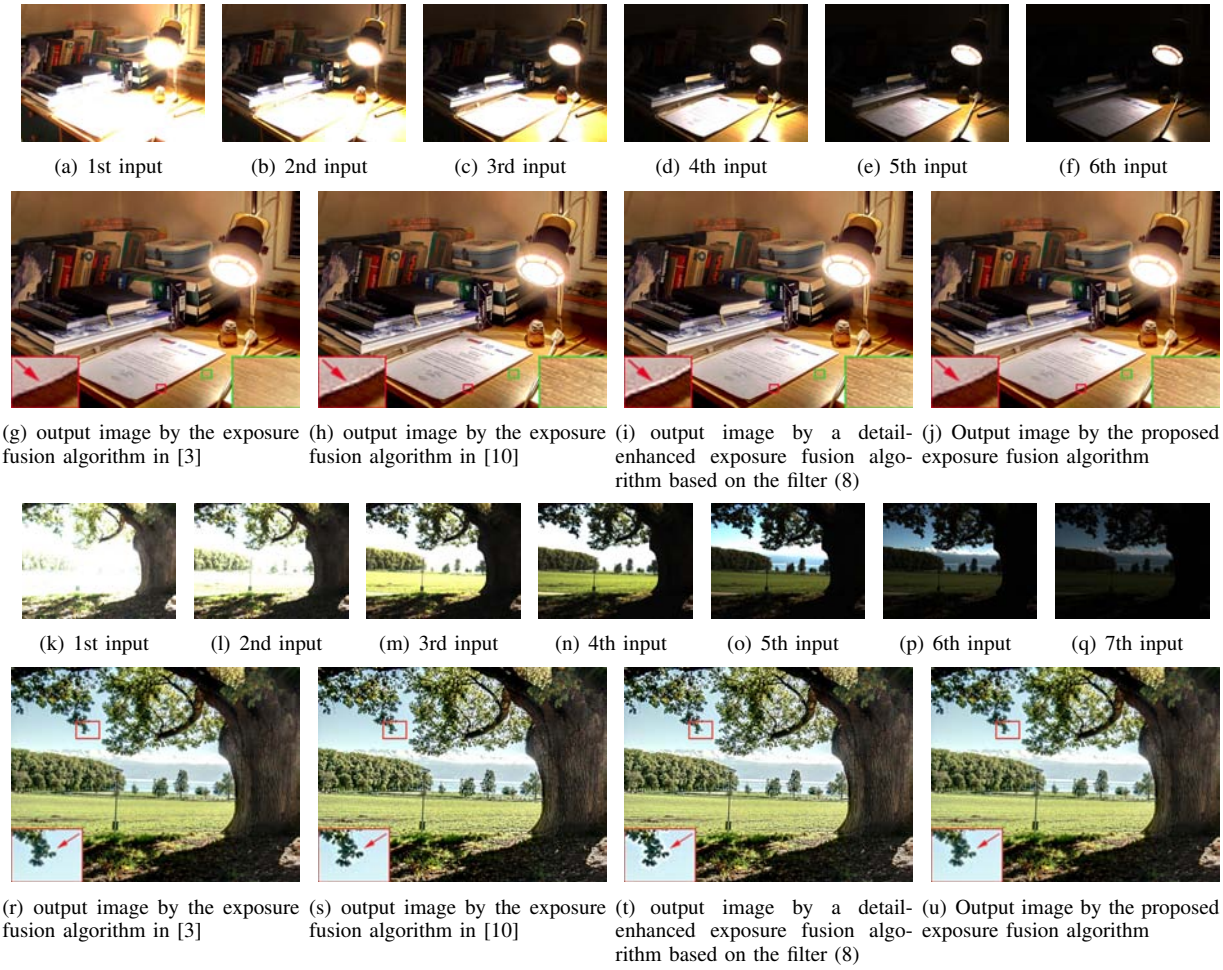
After correcting all the non-consistent regions of the image $\mathbf{Z}_{k_0+1}$, the corrected image serves as the reference image of the image $\mathbf{Z}_{k_0+2}$. All images are detected and corrected in an order of $(k_0 - 1)$, $\cdots$, 1, $(k_0 + 1)$, $\cdots$, $(N - 1)$, and N. All images are involved in the correction of non-consistent regions. As a result, the quality of moving objects is usually high in the final image. Moving objects are preserved and synchronized according to the moving objects in the initial reference image as shown in Fig. 1. The corrected images can be regarded as differently exposed images of a static HDR scene.

3 Selectively Detail-Enhanced Fusion of The Corrected Images

The corrected images $\mathbf{Z}_k (1 \leq k \leq N)$ can be regarded as differently exposed images of a static HDR scene. They are fused by using a new exposure fusion algorithm which includes a unique detail extraction module.

3.1 A Gradient Domain Content Adaptive Bilateral Filter

The bilateral filter in [23] is extended to the gradient domain in this subsection. For the image $\mathbf{Z}_k$, the bilateral filter in [23]



(a) 1st input (b) 2nd input (c) 3rd input (d) 4th input (e) 5th input (f) 6th input

(g) output image by the exposure fusion algorithm in [3] (h) output image by the exposure fusion algorithm in [10] (i) output image by a detail-enhanced exposure fusion algorithm based on the filter (8) (j) Output image by the proposed exposure fusion algorithm

(k) 1st input (l) 2nd input (m) 3rd input (n) 4th input (o) 5th input (p) 6th input (q) 7th input

(r) output image by the exposure fusion algorithm in [3] (s) output image by the exposure fusion algorithm in [10] (t) output image by a detail-enhanced exposure fusion algorithm based on the filter (8) (u) Output image by the proposed exposure fusion algorithm

Fig. 3: *Comparison of different exposure fusion algorithms. Image courtesy of Laurance Meylan.*

is

$$\mathbf{Z}_k^b(\mathbf{p}) = \frac{\sum\limits_{\mathbf{p}' \in \Theta_\rho(\mathbf{p})} \mathbf{w}_1(\mathbf{Z}_k, \mathbf{p}, \mathbf{p}')\mathbf{Z}_k(\mathbf{p})}{\sum\limits_{\mathbf{p}' \in \Theta_\rho(\mathbf{p})} \mathbf{w}_1(\mathbf{Z}_k, \mathbf{p}, \mathbf{p}')}, \tag{6}$$

where the value of ρ is empirically chosen as 4 to obtain a good tradeoff between the complexity and the quality of final images. $\mathbf{w}_1(\mathbf{Z}_k, \mathbf{p}, \mathbf{p}')$ is the product of $\exp(-\frac{(x-x')^2+(y-y')^2}{\sigma_s^2})$ and $\exp(-\frac{(\mathbf{Z}_k(\mathbf{p})-\mathbf{Z}_k(\mathbf{p}'))^2}{\sigma_r^2})$. σ_s and σ_r are two parameters which adjust the sensitive of the spatial similarity and the range similarity, respectively. Their values are empirically selected as 2.

Let $\mathbf{p}_r$ and $\mathbf{p}_b$ be the right and bottom pixels of the pixel $\mathbf{p}$. Defining the gradient field of image $\mathbf{Z}_k$, $\nabla\mathbf{Z}_k(p)$ as

$$\begin{aligned}\frac{\partial \mathbf{Z}_k(\mathbf{p})}{\partial x} &= \mathbf{Z}_k(\mathbf{p}_r) - \mathbf{Z}_k(\mathbf{p}),\\ \frac{\partial \mathbf{Z}_k(\mathbf{p})}{\partial y} &= \mathbf{Z}_k(\mathbf{p}_b) - \mathbf{Z}_k(\mathbf{p}),\end{aligned}$$

a vector field $\mathbf{G}(\nabla\mathbf{Z}_k, \mathbf{p}, \mathbf{p}')$ is then computed as:

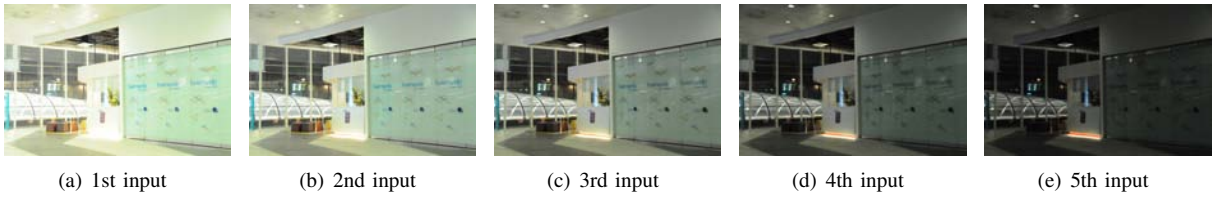
$$\mathbf{G}_x(\nabla\mathbf{Z}_k, \mathbf{p}, \mathbf{p}') = \begin{cases} \sum\limits_{r=x'}^{x-1} \frac{\partial \mathbf{Z}_k(r,y)}{\partial x}; & \text{if } x \geq x' \\ -\sum\limits_{r=x}^{x'-1} \frac{\partial \mathbf{Z}_k(r,y)}{\partial x}; & \text{otherwise} \end{cases},$$

$$\mathbf{G}_y(\nabla\mathbf{Z}_k, \mathbf{p}, \mathbf{p}') = \begin{cases} \sum\limits_{r=y'}^{y-1} \frac{\partial \mathbf{Z}_k(x',r)}{\partial y}; & \text{if } y \geq y' \\ -\sum\limits_{r=y}^{y'-1} \frac{\partial \mathbf{Z}_k(x',r)}{\partial y}; & \text{otherwise} \end{cases}.$$

Denoting $(\mathbf{Z}_k(\mathbf{p}) - \mathbf{Z}_k^b(\mathbf{p}))$ as $\mathbf{Z}_k^d(\mathbf{p})$, it can be derived that the value of $\mathbf{Z}_k^d(\mathbf{p})$ is:

$$\mathbf{Z}_k^d(\mathbf{p}) = \frac{\sum\limits_{\mathbf{p}' \in \Omega_r(\mathbf{p})} \left[\mathbf{w}_2(\nabla\mathbf{Z}_k, \mathbf{p}, \mathbf{p}', 1) \sum\limits_{q \in \{x,y\}} \mathbf{G}_q(\nabla\mathbf{Z}_k, \mathbf{p}, \mathbf{p}') \right]}{\sum\limits_{\mathbf{p}' \in \Omega_r(\mathbf{p})} \mathbf{w}_2(\nabla\mathbf{Z}_k, \mathbf{p}, \mathbf{p}', 1)}, \tag{7}$$

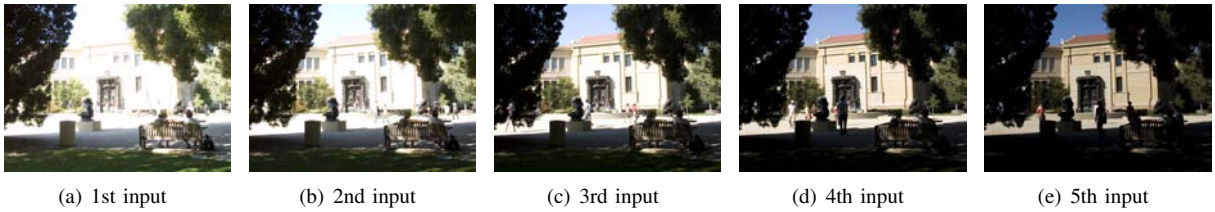
where $\mathbf{w}_2(\nabla\mathbf{Z}_k, \mathbf{p}, \mathbf{p}', \varepsilon)$ is the product of



(a) 1st input (b) 2nd input (c) 3rd input (d) 4th input (e) 5th input

(f) output image by the exposure fusion algorithm in [3] (g) output image by the exposure fusion algorithm in [10] (h) output image by a detail-enhanced exposure fusion algorithm based on the filter (8) (i) Output image by the proposed exposure fusion algorithm

Fig. 4: *Comparison of different exposure fusion algorithms for differently exposed images with noise.*



(a) 1st input (b) 2nd input (c) 3rd input (d) 4th input (e) 5th input

(f) output image by the correction module in [20] with the 3rd input as the initial reference image (g) output image by the proposed correction module with the 3rd input as the initial reference image (h) output image by the proposed scheme with the 5th input as the initial reference image

Fig. 5: *Comparison of the proposed detection module with the detection module in [20]. The correction module is fixed as the proposed one. Image courtesy of Orazio Gallo.*

$\exp(-\frac{(x-x')^2+(y-y')^2}{\sigma_s^2/\varepsilon})$ and $\exp(-\frac{(\sum_{q\in\{x,y\}}\mathbf{G}_q(\nabla\mathbf{Z}_k,\mathbf{p},\mathbf{p}'))^2}{\sigma_r^2/\varepsilon})$.

Replacing the gradient field $\nabla\mathbf{Z}_k$ by a vector field $\mathbf{V}(\mathbf{p}) = (\mathbf{V}_x(\mathbf{p}), \mathbf{V}_y(\mathbf{p}))$, a gradient domain bilateral filter is obtained as follows [22]:

$$\mathbf{L}^d(\mathbf{p}) = \frac{\sum_{\mathbf{p}'\in\Theta_\rho(\mathbf{p})}\left[\mathbf{w}_2(\mathbf{V},\mathbf{p},\mathbf{p}',1)\sum_{q\in\{x,y\}}\mathbf{G}_q(\mathbf{V},\mathbf{p},\mathbf{p}')\right]}{\sum_{\mathbf{p}'\in\Theta_\rho(\mathbf{p})}\mathbf{w}_2(\mathbf{V},\mathbf{p},\mathbf{p}',1)}. \tag{8}$$

Same as the filter in Equation (6), the filter (8) could produce halo artifacts due to unwanted smoothing of sharp edges. To overcome this problem, a content adaptive weighting $\mathbf{w}_g(\mathbf{p})$ is introduced and incorporated in the proposed filter (8). The value of $\mathbf{w}_g(\mathbf{p})$ is larger than 1 if the pixel $\mathbf{p}$ is at an edge and smaller than 1 if it is in a smooth area. This results in the following gradient domain content adaptive bilateral filter:

$$\mathbf{L}^d(\mathbf{p}) = \frac{\sum_{\mathbf{p}'\in\Theta_\rho(\mathbf{p})}\left[\mathbf{w}_2(\mathbf{V},\mathbf{p},\mathbf{p}',\mathbf{w}_g(\mathbf{p}))\sum_{q\in\{x,y\}}\mathbf{G}_q(\mathbf{V},\mathbf{p},\mathbf{p}')\right]}{\sum_{\mathbf{p}'\in\Theta_\rho(\mathbf{p})}\mathbf{w}_2(\mathbf{V},\mathbf{p},\mathbf{p}',\mathbf{w}_g(\mathbf{p}))}. \tag{9}$$

It is worth noting that two key components in the proposed bilateral filter (9) are the vector field $\mathbf{V} = (\mathbf{V}_x, \mathbf{V}_y)$ and the weighting function $\mathbf{w}_g(\mathbf{p})$. The details are provided in the next section by taking exposure fusion as an example.

(a) 1st input (b) 2nd input (c) 3rd input (d) output image by the correction module in [20] (e) output image by the proposed correction module

Fig. 6: *Comparison of the proposed correction module with the correction module in [20]. The detection module is fixed as the proposed one and the 2nd input is selected as the initial reference image. Image courtesy of Jacques Joffre.*

(a) 1st input (b) 2nd input (c) 3rd input

(d) output image by the algorithm in [4] (e) output image by the algorithm in [13] (f) output image by the algorithm in [14] (g) output image by the proposed scheme

Fig. 7: *Comparison of different ghost removal algorithms. Image courtesy of Jacques Joffre.*

3.2 Selectively Detail-enhanced Exposure Fusion

The first step of the proposed exposure fusion algorithm is to build up the vector field $\mathbf{V}$ which includes fine details of all the corrected images. The details of the image $\mathbf{Z}_k$ are represented by the variations of the luminance component $\mathbf{Y}_k$ in log domain. Normally, the gradient of a pixel with the largest absolute value along different exposures corresponds to the most desirable detail at the position. However, there is a high likelihood that the maximum gradient is noisy, especially in dark regions of an HDR scene. The vector field is thus built up by using the exposedness level of gradients over all exposures. As indicated in [3], [10], a well exposed pixel includes more reliable information than an under/over-exposed pixel. Therefore, weighting factor of a well exposed pixel is larger than that of an under/over-exposed pixel. Two weighting factors of a gradient vector $\nabla\mathbf{Y}_k(\mathbf{p})$ are computed as:

$$\begin{aligned}\mathbf{\Gamma}_{k,x}(\mathbf{p}) &= \gamma(\mathbf{Y}_k(\mathbf{p}))\gamma(\mathbf{Y}_k(\mathbf{p}_r)),\\ \mathbf{\Gamma}_{k,y}(\mathbf{p}) &= \gamma(\mathbf{Y}_k(\mathbf{p}))\gamma(\mathbf{Y}_k(\mathbf{p}_b)),\end{aligned}$$

where the weighting function $\gamma(z)$ is defined as:

$$\gamma(z) = \begin{cases} z+1; & \text{if } z \le 127 \\ 256 - z; & \text{otherwise} \end{cases}.$$

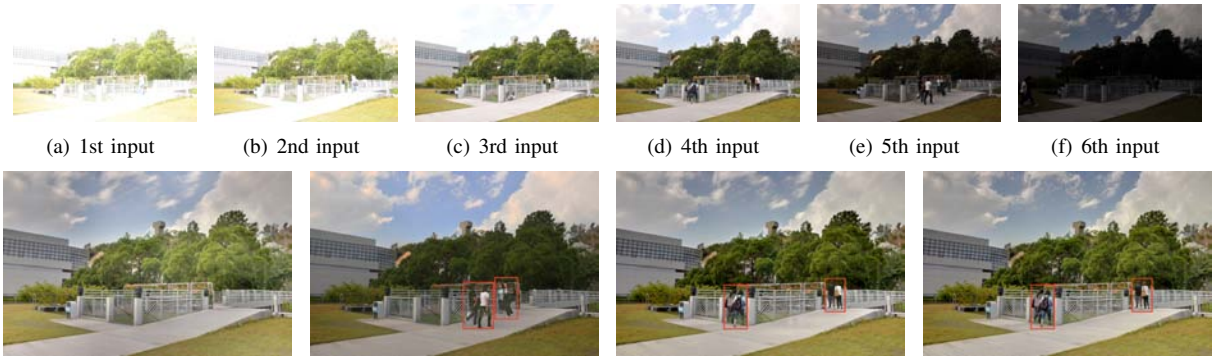
As the logarithmic conversion allows us to measure local contrast by using spatial difference [7], the desired vector field $\mathbf{V}$ is thus computed as:

$$\mathbf{V}_q(\mathbf{p}) = \frac{\partial \log(\mathbf{Y}_{\nu_q(\mathbf{p})}(\mathbf{p}))}{\partial q}\ ;\ q \in \{x, y\}, \tag{10}$$

where $\nu_q(\mathbf{p})$ is defined as

$$\nu_q(\mathbf{p}) = \arg\max_k \{\mathbf{\Gamma}_{k,q}(\mathbf{p})\}. \tag{11}$$

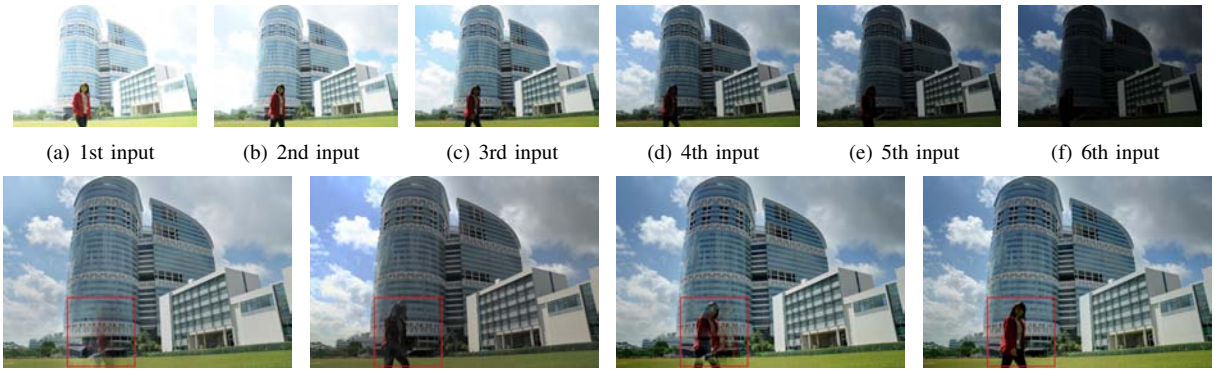
The value of $\mathbf{w}_g(\mathbf{p})$ is computed by using all luminance components $\mathbf{Y}_k (1 \le k \le N)$ in log domain. Let $\sigma^2_{\log(\mathbf{Y}_k)}(\mathbf{p})$ be the local variance of $\log(\mathbf{Y}_k)$ in a 3×3 square window centered at the pixel $\mathbf{p}$. Due to different exposures, a well exposed pixel in one input image could be under/over-exposed



(a) 1st input (b) 2nd input (c) 3rd input (d) 4th input (e) 5th input (f) 6th input

(g) output image by the algorithm in [4] (h) output image by the algorithm in [13] (i) output image by the algorithm in [14] (j) output image by the proposed scheme

Fig. 8: *Comparison of different ghost removal algorithms. Image courtesy of Wei Zhang.*



(a) 1st input (b) 2nd input (c) 3rd input (d) 4th input (e) 5th input (f) 6th input

(g) output image by the algorithm in [4] (h) output image by the algorithm in [13] (i) output image by the algorithm in [14] (j) output image by the proposed scheme

Fig. 9: *Comparison of different ghost removal algorithms.*

in another image. This implies that the value of $\sigma^2_{\log(\mathbf{Y}_k)}(\mathbf{p})$ is different for different k. On the other hand, gradient magnitude becomes larger when a pixel get better exposed, and it decreases as the pixel becomes under/over-exposed. Therefore, the largest value of $\sigma^2_{\log(\mathbf{Y}_k)}(\mathbf{p})$ along all k's is chosen to represent the overall local variance of pixel $\mathbf{p}$. The value of $\mathbf{w}_g(\mathbf{p})$ is then given as:

$$\mathbf{w}_g(\mathbf{p}) = \frac{1}{M}\sum_{\mathbf{p}'=1}^{M}\frac{\max\limits_{1\le k\le N}\{\sigma^2_{\log(\mathbf{Y}_k)}(\mathbf{p})\}+0.001}{\max\limits_{1\le k\le N}\{\sigma^2_{\log(\mathbf{Y}_k)}(\mathbf{p}')\}+0.001}, \tag{12}$$

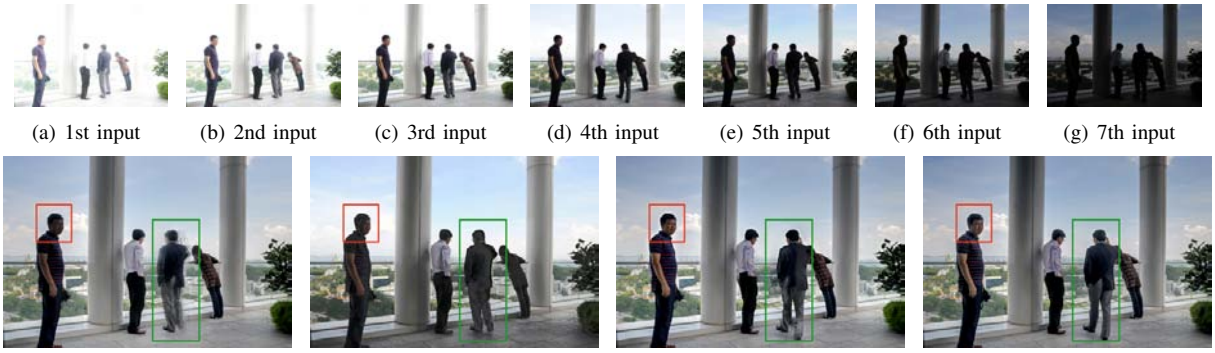
where M is the total number of pixels in the image $\mathbf{Z}_k$.

The value of $\mathbf{w}_g(\mathbf{p})$ is larger than 1 if the pixel $\mathbf{p}$ is at an edge and smaller than 1 if the pixel $\mathbf{p}$ is in a smooth area. Clearly, larger weights are assigned to pixels at edges than those pixels in flat areas by using the proposed weight $\mathbf{w}_g(\mathbf{p})$ in Equation (12). Applying this content-aware weighting, it is observed that the halo artifacts are greatly reduced. However, there might be blocking artifacts in final images. To prevent possible blocking artifacts from appearing in the final image, the value of $\mathbf{w}_g(\mathbf{p})$ is smoothed by using a Gaussian filter.

With the vector field $\mathbf{V}_q$ in Equation (10) and the weighting function $\mathbf{w}_g(\mathbf{p})$ in Equation (12), fine details are extracted from all the corrected images $\mathbf{Z}_k(1 \le k \le L)$ simultaneously by using the proposed filter (9). Let an intermediate image that is fused via the multi-scale exposure fusion algorithm in [3] be denoted as $\mathbf{Z}_{int}$ and it is obtained by a weighted blending of all the corrected images. $\mathbf{w}_{c,k}(\mathbf{p})$, $\mathbf{w}_{s,k}(\mathbf{p})$, and $\mathbf{w}_{e,k}(\mathbf{p})$ measure contrast, color saturation, and well-exposedness of pixel $\mathbf{Z}_k(\mathbf{p})$, respectively. $\mathbf{w}_{c,k}(\mathbf{p})$ is obtained by applying a Laplacian filter to the gray-scale version of each image. $\mathbf{w}_{s,k}(\mathbf{p})$ is computed as the standard deviation within the R, G and B channel. $\mathbf{w}_{e,k}(\mathbf{p})$ is yielded by applying a Gauss curve to each channel separately and multiplying the results. Their product is denoted as $\mathbf{w}_{f,k}(\mathbf{p})$. $\mathbf{\Psi}\{\mathbf{Z}_k(\mathbf{p})\}^m$ and $\mathbf{\Phi}\{\mathbf{w}_{f,k}(\mathbf{p})\}^m$ are Laplacian pyramid of image $\mathbf{Z}_k(\mathbf{p})$ and Gaussian pyramid of weight map $\mathbf{w}_{f,k}(\mathbf{p})$, respectively. Pixel intensities in the different pyramid levels are blended as:

$$\mathbf{\Psi}\{\tilde{Z}(\mathbf{p})\}^m = \sum_{k=1}^{N}\left[\mathbf{\Psi}\{\mathbf{Z}_k(\mathbf{p})\}^l\mathbf{\Phi}\{\mathbf{w}_{f,k}(\mathbf{p})\}^l\right]. \tag{13}$$

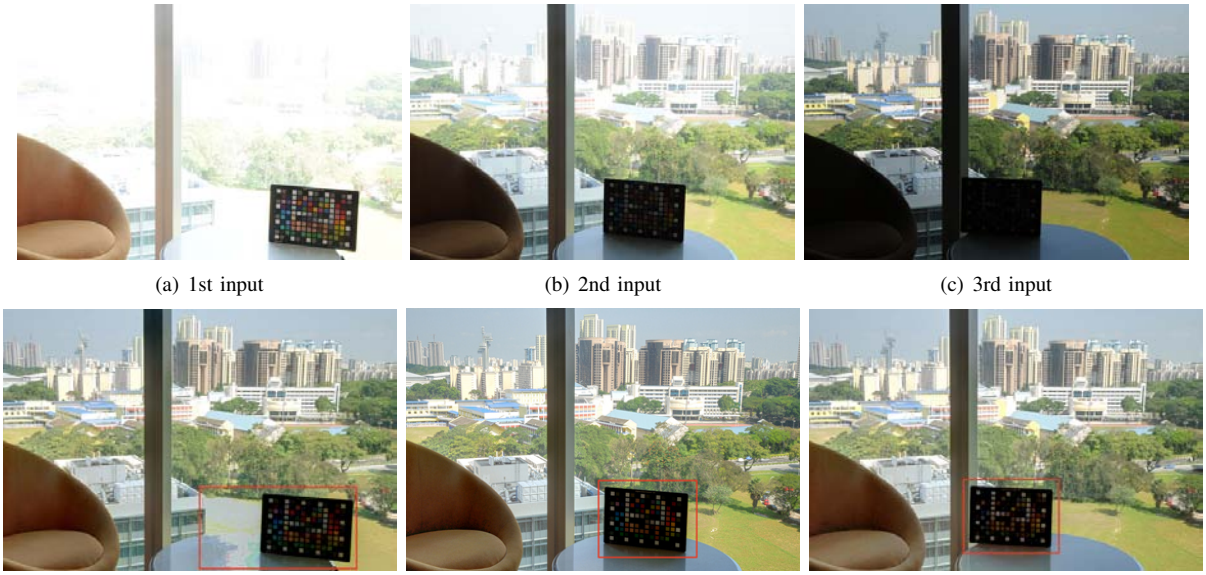
The pyramid $\mathbf{\Psi}\{\tilde{Z}(\mathbf{p})\}^m$ is collapsed to produce an intermediate image $\mathbf{Z}_{int}(\mathbf{p})$. It is worth noting that all the extracted fine details are added to the intermediate image $\mathbf{Z}_{int}$ by the detail-enhanced exposure fusion algorithm in [10]. Unfortunately, noise is also amplified when fine details are enhanced. The human visual system (HVS) can tolerate amplified noise in complex regions but is particularly sensitive to amplified noise



(a) 1st input (b) 2nd input (c) 3rd input (d) 4th input (e) 5th input (f) 6th input (g) 7th input

(h) output image by the algorithm in [4] (i) output image by the algorithm in [13] (j) output image by the algorithm in [14] (k) output image by the proposed scheme

Fig. 10: *Comparison of different ghost removal algorithms.*



(a) 1st input (b) 2nd input (c) 3rd input

(d) output image by selecting the 1st input as the initial reference image (e) output image by selecting the 2nd input as the initial reference image (f) output image by selecting the 3rd input as the initial reference image

Fig. 11: *Comparison of different selection of initial reference image.*

in flat areas. In addition, it is very difficult to separate noise from fine details. To reduce amplification of noise which is inherent in the fully detail-enhanced exposure fusion algorithm [10], a selectively detail-enhanced exposure fusion algorithm is introduced to enhance fine details in all regions except those in flat ones. Mathematically, the proposed selectively detail-enhanced exposure fusion algorithm is represented as:

$$\mathbf{Z}_f(\mathbf{p}) = \mathbf{Z}_{int}(\mathbf{p}) \exp^{\mathbf{L}^d(\mathbf{p})\eta(\mathbf{p})}, \tag{14}$$

where the value of $\eta(\mathbf{p})$ is computed by using $\mathbf{w}_g(\mathbf{p})$ in Equation (12). Its value is almost 0 if the pixel $\mathbf{p}$ is in a flat region and 1 otherwise.

4 Simulation Results

It is very difficult to provide a rigorous comparison with all available algorithms due to their implementation specific differences and parameter settings. We therefore focus on comparing the proposed scheme with the algorithms in [3], [4], [10], [14], [20] in this section. Experimental results are either generated by using the original implementations from authors or provided by authors themselves for a better comparison. They are evaluated by testing ten sets of differently exposed images. Seven of them are dynamic scenes and three of them are static scenes. Readers are invited to view to the electronic version of the full-size figures and zoom in these figures in order to better appreciate the differences among images.

4.1 Comparison Among Different Exposure Fusion Algorithms

The proposed exposure fusion algorithm is compared with the exposure fusion algorithm in [3], the detail-enhanced exposure fusion algorithm in [10], and a detail-enhanced exposure fusion algorithm based on the filter (8) by testing two sets of differently exposed images. It is shown in Fig. 3(g) that the

output image by the exposure fusion algorithm in [3] causes the loss of texture of the desk due to the smoothing effect of the Laplacian pyramid. This problem can be overcome by the detail-enhanced fusion algorithm in [10] and the proposed detail-enhanced exposure fusion algorithm. It is also shown that the detail-enhanced exposure fusion algorithm based on the filter (8) produces halo artifacts while the halo artifacts are removed by using the proposed content adaptive weighting $\mathbf{w}_g(\mathbf{p})$ in Equation (12). It is shown in Fig. 4 that noise in flat regions is amplified using both the detail-enhanced exposure fusion algorithm in [10] and the detail-enhanced exposure fusion algorithm based on the filter (8) while this problem is overcome using the proposed selectively detail-enhanced exposure fusion algorithm. Therefore, the visual quality of final images is further improved via the proposed selectively detail-enhanced exposure fusion algorithm. On the other hand, it should be pointed out that noise in complex regions could be amplified by the proposed selectively detail-enhanced exposure fusion algorithm. Fortunately, our HVS is not sensitive to the noise in the complex regions.

4.2 Comparison Among Different Ghost Removal Algorithms

The proposed ghost removal algorithm is first compared with the algorithm in [20]. Since there are detection and correction modules in both algorithms, the two modules are compared respectively by allowing only one module being changed. It is shown in Figs. 5(f) and 5(g) that there are visible ghosting artifacts by using the detection module in [20] while the ghost artifacts are prevented from appearing in the final image via the proposed detection module. It is demonstrated in the sky regions of Figs. 6(d) and 6(e) that the proposed correction module also outperforms the correction module in [20].

The proposed ghost removal algorithm is then compared with the ghost removal algorithms in [4], [13], [14] by testing four sets of differently exposed images with moving objects as illustrated in Figs. 7-10, respectively. The ghost removal algorithms in both [13] and [14] work fairly well on the sequence in Fig. 8 while there are visible ghosting artifacts for the image sequences in Figs. 7-10. The ghost removal algorithm in [4] requires that moving objects appear in a relative smaller number of images. The image sequence in Fig. 8 satisfies the requirement. Thus, the algorithm in [4] removes ghosting artifacts very well as shown in Figs. 8(g). Owing to this, the algorithm in [4] is useful when a photographer wants to capture a popular scenery site without any foreground object. Unfortunately, image sequences that are captured by using exposure bracketing usually do not satisfy this requirement. Consequently, ghosting artifacts cannot be completely removed by using the algorithm in [4] as in Figs. 7(d), 9(g) and 10(h).

4.3 Limitations of the Proposed Scheme

The selection of the initial reference image plays an important role in the proposed ghost removal algorithm. The image with the overall best exposure, typically the middle one, usually prevents ghosting artifacts from appearing in the final images. However, the quality of the final image needs to be improved if 1) there are over/under-exposed pixels in the initial reference image; 2) their pixels at the same positions in other images are non-consistent; and 3) similar pixels cannot be found in their neighborhoods. One example is given in Fig. 11. There is an over-exposed background region when the initial reference image is selected as the brightest input in Fig. 11(a). Their corresponding regions are occluded by the moving object in other images. It is shown in Fig. 11(d) that the quality of the final image, especially the surface of the table, needs to be improved. Another example is given in Fig. 5. If the initial reference image is selected as the darkest one, the quality of moving human subjects in the final image needs to be improved as shown in Fig. 5(h). If running time is not a concern, the methods in [30], [31] could be adopted to address this problem.

It is illustrated in Figs. 3-11 that color saturation is sometimes reduced by using existing exposure fusion schemes including the proposed one even though a color saturation measure is explicitly used in these exposure fusion schemes. This problem could be addressed by enhancing the colorfulness of the final image in future research.

5 Conclusion and Discussion

In this paper, a selectively detail-enhanced exposure fusion scheme is proposed to fuse multiple differently exposed low dynamic range (LDR) images into a more detailed and natural LDR image. The proposed scheme is applicable to both static and dynamic high dynamic range (HDR) scenes. It is also applicable to multiple exposed images of the same scene regardless of their lighting conditions. Therefore, it could be an ideal solution for real time exposure fusion over mobile devices.

Similar to the content adaptive bilateral filters in [32], the acceleration of the proposed filter could be an issue. Fortunately, the idea in [33] might be borrowed to accelerate the proposed filter. This problem will be studied in our future research.

References

[1] P. E. Debevec and J. Malik, "Rendering high dynamic range radiance maps from photographs," In *Proceedings of SIGGRAPH 1997*, pp. 369–378, May 1997.

[2] S. Mann, "Comparametric equations with practical applications in quantigraphic image processing," IEEE Trans. on Image Processing, vol. 9, no. 8, pp. 1389-1406, Aug. 2000.

[3] T. Mertens, J. Kautz, and F. V. Reeth, "Exposure fusion: a simple and practical alternative to high dynamic range photography," Computer Graphics Forum, vol. 28, no. 1, pp. 161-171, Jan. 2009.

[4] W. Zhang and W. K. Cham, "Gradient-Directed Multiexposure Composition," IEEE Trans. on Image Processing, vol. 21, no. 4, pp. 2318-2323, Apr. 2012.

[5] P. Burt and T. Adelson, "The laplacian pyramid as a compact image code," IEEE Trans. on Communication, vol. 31, no. 4, pp. 532–540, Apr. 1983.

[6] S. Raman and S. Chaudhuri, "Bilateral filter based compositing for variable exposure photography," In *EUROGRAPHICS*, pp. 1-4, Munich, Germany, Apr. 2009.

[7] R. Fattal, M. Agrawala, and S. Rusinkiewicz, "Multiscale shape and details enhancement for multi-light image collections," ACM Trans. on Graphics, pp. 51:1-51:10, Aug. 2007.

[8] M. Aubailly, M. A. Vorontsov, G. W. Carhart, and M. T. Valley, "Automated video enhancement from a stream of atmospherically-distorted images: the lucky-region fusion approach," In *Proc. SPIE 7463 (2009)*.

[9] D. D. Baumgartner and B. J. Schachter, "Improving FLIR ATR performance in a turbulent atmosphere with a moving platform," In *Proc. SPIE 8391 (2012)*.

[10] Z. G. Li, J. H. Zheng, and S. Rahardja, "Detail-enhanced exposure fusion," IEEE Trans. on Image Processing, vol. 21, no. 11, pp.4672-4676, Nov 2012.

[11] E. Reinhard, G. Ward, S. Pattanaik, and P. E. Debevec, High dynamic range imaging: acquisition, display and image-based lighting. Morgan Kaufmann, 2005.

[12] Z. Farbman, R. Fattal, D. Lischiski, and R. Szeliski, "Edge-preserving decompositions for multi-scale tone and detail manipulation," ACM Trans. on Graphics, Aug. 2008.

[13] K. Jacobs, C. Loscos, and G. Ward, "Automatic high dynamic range image generation for dynamic scenes," IEEE Computer Graphics and Applications, vol. 28, no. 2, pp. 84-93, Apr. 2008.

[14] F. Pece and J. Kautz, "Bitmap movement detection: HDR for dynamic scenes," In *Conference on Visual Media Production*, pp. 1-8, London, United Kingdom, Oct. 2010.

[15] E. A. Khan, A. O. Akyuz, and R. Reinhard, "Ghost removal in high dynamic range images," In *IEEE International Conference on Image Processing*, pp. 2005-2008, 2006.

[16] P. Sen, N. Khademi Kalantari, M. Yaesoubi, S. Darabi, D. Goldman, and E. Shechtman, "Robust patch-based HDR reconstruction of dynamic scenes," In *SIGGRAPH ASIA 2012*.

[17] J. Hu, O. Gallo, and K. Pulli, "Exposure stacks of live scene with hand-held cameras," In *ECCV 2012*.

[18] S. Kang, M. Uttendaele, S. Winder, and R. Szeliski, "High dynamic range video," ACM Trans. On Graphics, pp. 319-325, Aug. 2003.

[19] O. Gallo, N. Gelfand, W. Chen, M. Tico, and K. Pulli, "Artifact-free high dynamic range imaging," In *IEEE International Conference on Computational Photography (ICCP)*, pp. 1-7, 2009.

[20] Z. G. Li, S. Rahardja, Z. J. Zhu, S. L. Xie, and S. Q. Wu, "Movement detection for the synthesis of high dynamic range images," In *IEEE International Conference on Image Processing*, pp. 3133-3136, Hongkong, Oct. 2010.

[21] M. Grossberg and S. Nayar, "Determining the camera response from images: what is knowable?" IEEE Trans. on Pattern Analysis and Machine Intelligence, vol. 25, no. 11, pp. 1455–1467, Nov. 2003.

[22] Z. G. Li, J. H. Zheng, Z. J. Zhu, S. Q. Wu, and S. Rahardja, "A bilateral filter in gradient domain," In *2012 International Conference on Acoustics, Speech, and Signal Processing*, pp. 1113-1116, Kyoto, Japan, Mar. 2012.

[23] C. Tomasi and R. Manduchi, "Bilateral filtering for gray and color image," In *Proc. IEEE International Conference on Computer Vision 1998*, pp. 836–846, Bombay, India, Jan. 1998.

[24] U. Steinmueller, "Capture more light: getting the most out of your HDR images," http://www.learn.usa.canon.com/resources/articles/2011/capture_more_light_hdr_images_article.shtml.

[25] E. Martinec, "Noise, dynamic range and bit depth in digital SLRs," *http://theory.uchicago.edu/ ejm/pix/20d/tests/noise*, 2008.

[26] C. Liu, R. Szeliski, S. B. Kang, C. L. Zitnick, and W. T. Freeman, "Automatic estimation and removal of noise from a single image," IEEE Trans. on Pattern Analysis and Machine Intelligence, vol. 30, no. 2, pp. 299-314, Jun. 2008.

[27] F. J. Richards, "A flexible growth function for empirical use," Journal of Experimental Botany, pp. 290–301, 1959.

[28] R. C. Gonzalez and R. E. Woods, Digital image processing. Prentice Hall, 2002.

[29] C. H. Chou and Y. C. Li, "A perceptually tunes subband image coder based on the measure of just-noticeable-distortion profile," IEEE Trans. on Circuits and Systems for Video Technology, vol. 5, no. 6, pp.467–476, Dec. 1995.

[30] J. Hu, O. Gallo, K. Pulli, and X. Sun, "HDR deghosting: how to deal with Saturation?," In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Jun. 2013.

[31] J. H. Zheng, Z. G. Li, Z. J. Zhu, S. Q. Wu, and S. Rahardja, "Hybrid patching for a sequence of differently exposed images with moving objects," IEEE Trans. on Image Processing, vol. 22, no. 12, pp. 5190-5201, Dec. 2013.

[32] B. Y. Zhang and J. P. Allebach, "Adaptive bilateral filter for sharpness enhancement and noise removal," IEEE Trans. on Image Processing, vol. 17, no. 5, pp. 664-678, May 2008.

[33] J. Baek and David E. Jacobs, "Accelerating spatially varying Gaussian filters," In *SIGGRAPH Asia 2010*, Seoul, South Korea, Dec. 2010.