RESEARCH ARTICLE

Efficient joint noise removal and multi exposure fusion

Antoni Buades[¶], Jose Luis Lisani[¶], Onofre Martorell[¶]*

Institute of Applied Computing and Community Code (IAC3) and with the Dept. of Mathematics and Computer Science, Universitat de les Illes Balears, Palma, Spain

¶ The authors acknowledge the Ministerio de Ciencia, Innovación y Universidades (MCIU), the Agencia Estatal de Investigación (AEI) and the European Regional Development Funds (ERDF) for its support to the project TIN2017-85572-P.
* o.martorell@uib.cat

Abstract

Multi-exposure fusion (MEF) is a technique that combines different snapshots of the same scene, captured with different exposure times, into a single image. This combination process (also known as fusion) is performed in such a way that the parts with better exposure of each input image have a stronger influence. Therefore, in the result image all areas are well exposed. In this paper, we propose a new method that performs MEF and noise removal. Rather than denoising each input image individually and then fusing the obtained results, the proposed strategy jointly performs fusion and denoising in the Discrete Cosinus Transform (DCT) domain, which leads to a very efficient algorithm. The method takes advantage of spatio-temporal patch selection and collaborative 3D thresholding. Several experiments show that the obtained results are significantly superior to the existing state of the art.

1 Introduction

Multi-exposure fusion (MEF) methods combine different pictures of the same scene, captured with different exposure times, into a single image. By keeping the best exposed parts of each image, it is possible to reconstruct a result where all the details of the scene are well rendered. Compared to High Dynamic Range (HDR), MEF doesn't require the estimation of the camera response function and the tone mapping of the processed values to the standard 8-bit [0, 255] range.

The differences between the existing MEF algorithms in the literature lie in the way the information of the input images is blended. The classical method by Mertens et al. [1] computes weighted averages of the values of the input images at the same pixel location, the weights depending on exposure, saturation and contrast. These averages can also be computed in the Discrete Cosinus Transform (DCT) domain [2]. Other authors [3, 4] use gradient information, and obtain the final result by solving a Poisson partial differential equation. Moreover, to prevent the appearance of visual artifacts in the resulting images, the blending is not applied

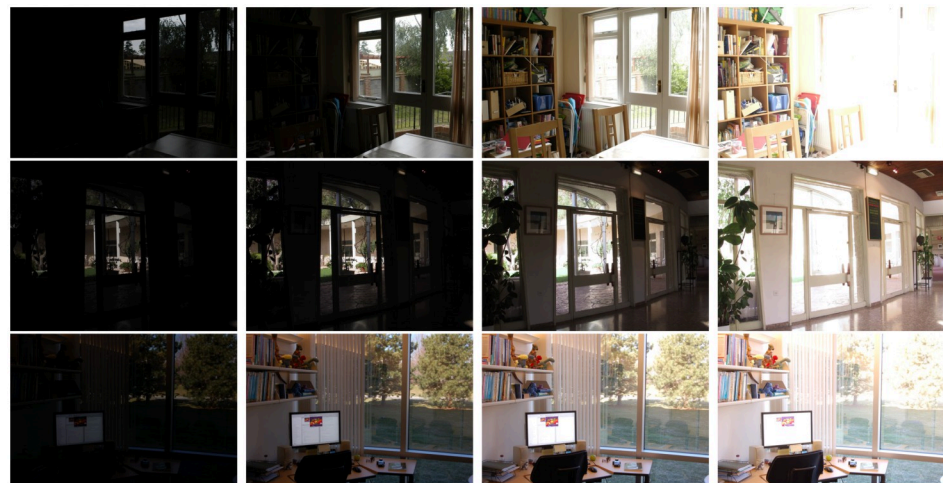on a per-pixel basis. Instead, a patch-based approach or a multiresolution blending strategy are used [5–7].

The use of image fusion techniques is not limited to multi-exposure images. Indeed, the combination of several images permits to improve their quality, removing for example noise [8], compression artifacts [9], haze [10–12], blur [13] or shaking blur from hand held video [14, 15].

In this paper we tackle the problem of fusing noisy multi-exposed images. A naive approach would consist in independently denoising the multi-exposed images and then applying a fusion technique to the result. We address both problems jointly for the first time in the literature, to the best of our knowledge. We develop a technique that draws inspiration from the well known BM3D denoising method [16] and the multi-exposure fusion method described in [2].

BM3D denoising takes advantage of the redundancy of similar image patches (i.e. small image square blocks). Each image patch is denoised by grouping together similar patches in a local neighborhood and stacking them in a 3D structure to which a 3D transform is applied. In practice, the 3D transform is applied in a separable way, i.e. 2D DCT transforms followed by a 1D DCT or Walsh-Hadamard transform, as detailed in [16]. Denoising is achieved by applying a shrinkage operator to the coefficients in the transformed domain. This denoising technique is known as collaborative filtering.

We adapt BM3D to denoise the whole set of multi-exposure images by using the spatio-temporal patch selection strategy proposed in [8]. The fusion method in [2], on the other hand, merges the 2D DCT coefficients of differently exposed images. This method is not able to merge the coefficients describing the average of the patch and copies this value from the Mertens et al. [6] solution. In this paper we modify the algorithm to get rid of the dependence on Mertens et al.'s result, and include an additional step to make it robust to noise. The fact that both BM3D and the proposed fusion method work on the 2D DCT domain permits to combine them to obtain an efficient algorithm that simultaneously denoises and fuses the multi-exposed images.

The performed experiments in the sequences of Fig 1 show the superior performance of the proposed technique over the naive approach, both in terms of quality of the results and computational efficiency. The paper is organized as follows: Section 2 describes the existing literature on MEF methods. The fusion method is described in Section 3. Section 4 presents the complete proposed technique, including the BM3D-inspired denoising stage and the fusion

Fig 1. Multi exposure data sets used for comparison.

https://doi.org/10.1371/journal.pone.0265464.g001

method. In Section 5, we discuss the implementation of the method and compare with state of the art algorithms. Finally we draw some conclusions in Section 6.

2 Related work

The literature on MEF is very extensive. Mertens et al. [1] proposed to combine the images by averaging, choosing for each pixel a different weight depending on saturation, contrast and well-exposedness. Since such an average produces ghosting effect if the images are not well aligned, several methods were proposed to take motion into account, such as An et al. [17], Liu et al. [18], Hessel et al. [19] Li et al. [20], Ocampo et al. [21] and Hayat et al. [22].

Instead of averaging directly pixel values, several methods prefer to fuse gradient information, and then obtain the final image by solving a Poisson equation [3]. Some authors [4, 23–25] propose to merge all gradient values, while others (Kuk et al. [26]) choose the gradient corresponding to the better exposed image.

Deciding independently for each pixel which is the correct combination is not a robust strategy and leads to visual artifacts. Common approaches to improve the results involve the use of pyramidal image representations or patches. Several methods [6, 7, 27, 28] adopted the Laplacian pyramid [29]. Patch-based methods make the fusion more robust by involving all the pixels in a small window [5, 30–32]. The method in [2] uses a DCT transform. The 2D DCT coefficients of the patches at the same spatial location and different exposure are combined depending on its magnitude. This combination is not valid to set the patch illumination, which is obtained using the Mertens et al. [6] algorithm.

The use of alternative color systems to the standard RGB is often proposed, specially YCbCR [33] which separates the luminance from the chromatic components. Since the chromatic components contain few high frequency information, a simpler strategy can be used to fuse them. Published methods differ depending on which one of the previously exposed techniques is applied to the Y component [2, 34, 35].

For non static sequences, the fusion creates ghosting effects near the boundary of moving objects. A possible strategy to address this problem is to weight the fusion considering patch correlation, thus discarding dissimilar patches which might belong to a different object. Another common choice is to register the whole sequence into a reference exposure [2, 36–38]. A straightforward strategy is to select the better exposed image as reference and generate a new sequence where the color of each image is equalized with respect to this reference. This new sequence can be fused using a MEF method, as proposed in [39–42]. Such a strategy avoids the creation of ghosting effects due to motion, but the fusion does not permit to get rid of noise.

Several methods (Z. G. Li et al. [43, 44], Singh et al. [45], Raman et al. [46], Li et al. [47]) divide the images into low and high frequency components using the Bilateral [48] or the Guided filter [49]. This separation permits an additional enhancement of the image details.

Multi-exposure fusion might be accomplished by the use of variational techniques [50–52]. The proposed methods favor the geometry of the short exposure views and the chromaticity of the ones with longer exposure. Other methods are defined in order to maximize particular quality measures as [53–56].

Recently, neural networks have been proposed for multi exposure fusion. They can be divided into two categories, depending on whether they require the ground truth to be correctly exposed in all image areas to define the loss function (supervised methods) or not (unsupervised or self-supervised method). For HDR Kalantari et al. [57] and Wu et al. [58] proposed supervised methods. For MEF, Xu et al. [59] proposed a Generative Adversarial Network (GAN) strategy to fuse pairs of images with different exposure time. In the group of the self-supervised methods, Prabhakar et al. [60, 61] designs a metric which evaluates the quality of

the multi-exposure fusion image. Xu et al. [62] trained a self-supervised neural network to preserve the similarity between the fusion result and the source exposures.

Li et al. (CNNFEAT) [63] combines the several exposures depending on a series of descriptors learnt by a neural network. Zhang et al. [64] and Zhang et al. [65] propose unified deep learning frameworks for several fusion tasks, including MEF.

There is very few literature dealing with noise removal during multi-exposure image fusion, and most published papers are focused on HDR. Akyuz et al. [66] denoise each frame before fusion, but this is performed in the radiance domain. Tico et al. [39] combine an initial fusion result with the image in the sequence with the shortest exposure. This combination is performed in the wavelet domain and coefficient attenuation is applied to the coefficients of the difference. Min et al. [67] filter the set of images with spatio-temporal motion compensated anisotropic filters prior to HDR reconstruction. Lee et al. [68] use sub-band architecture for fusion, with a weighted combination using a motion indicator function to avoid ghosting effects. The low frequency bands are filtered with a multi-resolution bilateral filter while the high frequency bands are filtered by soft thresholding. Ahmad et al. [69] identify noisy pixels and reduce their weight during image fusion.

3 MEF algorithm

We propose a novel algorithm for multi-exposure fusion which will serve as basis for the joint fusion and denoising method. It operates in the DCT domain and it is inspired by the fusion method in [2]. Compared to [2], we introduce a new strategy for fusing the average coefficient of the DCT and a completely different color management. It also includes a novel noise removal step, allowing for its application with moderately noisy images.

The basic algorithm, applied to single-channel images, is described first. Its extension to color images is presented in Section 3.2. Finally, a modification that confers denoising capability to the method is proposed in Section 3.3.

3.1 Single channel image fusion

Let's denote by Y_k, $k = 1, 2, \ldots, K$ a sequence of luminance images acquired with different exposure taking values in the range [0, 1]. This sequence is supposed to be static. We split the images Y_k into n_b partially-overlapped patches of $b \times b$ pixels, $\{B_k^l\}$, $l = 1, \ldots, n_b$, and the 2D DCT transform of each patch is computed.

At a given pixel location, some of these patches may belong to under or over exposed parts of the images, while others may be well exposed. Both under and over exposed patches will have non-zero DCT coefficients of small magnitude, due to the lack of high frequency information. Conversely, the coefficients of well exposed patches will be large. These considerations lead to the following equation that aggregates the non-zero DCT coefficients of all the patches corresponding to the same spatial location:

$$\hat{B}^l(\xi) = \sum_{k=1}^{K} w_k^l(\xi)\hat{B}_k^l(\xi), \quad \xi \neq 0, \quad l = 1, 2, \ldots, n_b, \tag{1}$$

where $\hat{B}_k^l$ denotes the DCT transform of patch l in image k, and the weights $w_k^l(\xi)$ are defined depending on the frequency ξ,

$$w_k^l(\xi) = \frac{|\hat{B}_k^l(\xi)|^p}{\sum_{n=1}^{K} |\hat{B}_n^l(\xi)|^p} \qquad \xi \neq 0. \tag{2}$$

where $p > 0$ is a parameter of the method. We observe that, for a given frequency, patches

having higher coefficient magnitudes (i.e. well exposed patches) contribute more than the others to the weighted sum in Eq (1).

This strategy does not apply to the zero frequency DCT coefficient, $\xi = 0$, i.e. the average of the patch values. Since large zero frequency coefficients correspond to over-exposed images, applying the same weighted combination would simply overexpose the fused image. We weight these coefficients depending on the average value of the patch and on the average value of the image to which it belongs:

$$\hat{B}^l(0) = \sum_{k=1}^{K} w_k^l(0)\hat{B}_k^l(0), \quad l = 1, 2, \ldots, n_b, \tag{3}$$

where

$$w_k^l(0) = \frac{1}{C} e^{-(\hat{B}_k^l(0)-0.5)^2/\sigma_l^2} \cdot e^{-(\mu_k-0.5)^2/\sigma_g^2}, \tag{4}$$

with μ_k the average of the values of image k, C a normalizing constant, and σ_l and σ_g parameters of the method. Observe that 0.5 corresponds to the center of the [0, 1] range. Well exposed images are supposed to have an average value close to 0.5. The weighting factor (4) favors patches whose average value is close to this central value, and that belong to well exposed images. For all the experiments in this paper we fixed $\sigma_g = 0.3$ and $\sigma_l = 0.5$.

Finally, the fused patches are obtained after applying the inverse DCT transform to the result of the aggregation

$$B^l(x) = \mathcal{F}^{-1}(\hat{B}^l(\xi)), \quad l = 1, \ldots, n_b. \tag{5}$$

Since patches are partially overlapped, the pixels in overlapping areas are averaged to produce the final image.

3.2 Color image fusion

We use an orthonormal version of the well known YUV color space, described by the following linear transformation.

$$\begin{pmatrix} Y \\ U \\ V \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \\ \frac{\sqrt{2}}{2} & 0 & -\frac{\sqrt{2}}{2} \\ \frac{1}{\sqrt{6}} & -\frac{2}{\sqrt{6}} & \frac{1}{\sqrt{6}} \end{pmatrix} \begin{pmatrix} R \\ G \\ B \end{pmatrix}$$

The Y channel, given by the normalized average of the RGB values, represents the luminance, while U and V contain the chromatic information. The use of an orthonormal transform is motivated by the denoising stage described below. The matrix rows are mutually orthogonal which guarantees that noise is not color correlated by the transform. Each row has an Euclidean norm equal to one which permits to maintain the same noise standard deviation of the original image. This transformation can be obtained from the classical YUV decomposition by applying an orthogonalization method [70]. This color transformation was used for example in [16].

The Y channel is processed as described in the previous section. However, for the U and V components, we apply the weighted average defined by Eqs (1) and (2) to all the coefficients,

Fig 2. Comparison between applying the fusion to each RGB channel independently and using the YUV luminance and chromatic components. The luminance of the two results is identical, while the color is enhanced by the latter strategy.

including $\xi = 0$. The reason is that aggregating the average values of the chromaticity components (i.e. the values of the zero coefficients in the transformed domain) does not increase the risk of over-exposure, but enhances the patch average chromaticity, making the result more colorful than when applying the single channel method to each of the R, G and B components. This is noticeable in Fig 2, in which both strategies are compared.

3.3 Noise removal

The most common noise model is the additive white Gaussian noise (AWGN). The observed noisy image v is related to the underlying clean image u by

$$v = u + n,$$

being n a noise image, independently and identically distributed at each pixel as a zero-mean Gaussian random variable with standard deviation σ. For other types of noise, the initial data can be modified by using variance stabilization transforms or whitening strategies, or the designed algorithm might be modified by adapting locally the parameters or applying multi-scale methods. For this reason, AWGN is the most commonly assumed model in order to design general noise removal algorithms.

Denoising can be achieved by using a thresholding estimator that projects the noisy image to an orthonormal basis and reconstructs the denoised result with the transform coefficients larger than a given threshold [71].

Following this principle, the fusion method proposed in Sections 3.1 can naturally incorporate noise removal by modifying the weight definition (2) to

$$w_k^l(\xi) = \frac{Thr_\sigma(|\hat{B}_k^l(\xi)|)^p}{\sum_{n=1}^{K} Thr_\sigma(|\hat{B}_n^l(\xi)|)^p}, \qquad \xi \neq 0 \tag{6}$$

where

$$Thr_\sigma(|\hat{B}_k^l(\xi)|) = \begin{cases} 0 & |\hat{B}_k^l(\xi)| < T \cdot \sigma \\ \\ |\hat{B}_k^l(\xi)| & \text{otherwise} \end{cases}$$

Since the modified YUV color transformation is orthonormal, the noise standard deviation is not modified by the linear transformation proposed in Section 3.2, converting from RGB to the mentioned space. Thus, the same thresholding can be applied to each channel. The

| One of the exposures | Fusion | Fusion with DCT thresh. |
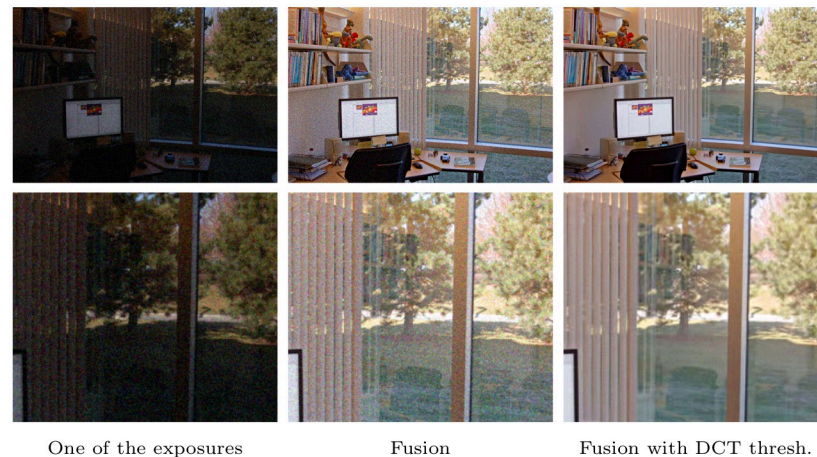
Fig 3. Noisy example with sequence office and noise with standard deviation σ = 15. Top, from left to right: one image from the multi-exposure set, fusion without DCT thresholding, fusion with DCT thresholding. Bottom: detail of the images.

threshold parameter T is set to 2.7 as usual when denoising by thresholding in an orthonormal basis [16]. Fig 3 compares the application of the fusion chain with and without this DCT thresholding stage.

When dealing with high levels of noise, the DCT thresholding method described above is not enough to provide good denoising results. The next section describes how the fusion method can be combined with a collaborative denoising technique to obtain much better results.

4 Joint noise removal and fusion procedure

We assume, as in the previous sections, an AWGN noise model with standard deviation σ and that the input images are co-registered. Many variants have been proposed to deal with the noise removal of image sequences having the same exposure and noise conditions. Such methods cannot be directly used for multi-exposed images.

We propose a joint noise removal and fusion procedure. The noise removal stage is an adaptation of the BM3D collaborative DCT thresholding technique proposed in [16]. The use of such technique permits a natural integration with the fusion method proposed in Section 3, since both denoising and fusion are performed in the DCT domain.

BM3D is a patch-based image denoising method, which means that first the image is split into overlapping patches which are processed independently. Each patch is denoised by finding groups of similar patches in a local neighborhood and stacking them in a 3D structure. A separable 3D transform is applied to this structure (that is, 2D DCT transforms of each patch followed by a 1D transform in the third dimension of the stack) and denoising is achieved by setting to zero all the coefficients below a fixed threshold that depends on the standard deviation of the noise (which is assumed to be known). After computing the inverse 3D transform of the thresholded coefficients a denoised stack of patches is obtained. Since patches are partially overlapped the final denoised image is obtained after averaging, at each pixel position, the contributions of each denoised patch, in a process known as aggregation. This process should not be confused with the aggregation in the DCT domain described in Section 3.1 and modeled by Eqs (1) and (3).

In the original BM3D implementation this process is repeated two times, first applied to the noisy image, and then using as input both the noisy image and the first denoising result

(known as oracle). In this second step the thresholding operation in the DCT domain is replaced by a Wiener filtering guided by the oracle image. The authors show that this two steps process restores more details and improves the denoising performance.

A naive solution to our problem would consist in denoising each multi-exposed image with the previous algorithm and then apply the fusion method described in Section 3 to the denoised set. However, we propose a more efficient method, which produces better results.

Instead of processing each image independently, we use information of the whole multi-exposed set to denoise each image. In particular, the group of similar patches used to create each 3D stack is searched using the patch-selection procedure proposed in [8], which has proved to reduce the dependence on noise in the patch comparison, improving the robustness of the denoising method and reducing the usual artifacts of collaborative filtering.

Given the multi-exposed set of images and an initial patch location x, we associate to x a 3D block composed by the 2D patches from the set located at the same spatial position. Then, we search for similar 3D blocks associated to other spatial locations y. The distance between different 3D blocks is computed as

$$d_{3D}(\boldsymbol{x}, \boldsymbol{y}) = \sum_{\text{image } i \text{ in set}} ||P_i(\boldsymbol{x}) - P_i(\boldsymbol{y})|| \tag{7}$$

where $P_i(\boldsymbol{x})$ and $P_i(\boldsymbol{y})$ denote the 2D patches referenced by x and y in image i (each 2D patch is referenced by its top-left vertex). We extract from the selected 3D blocks, the set of 2D patches belonging to each one of the images. Each set of 2D patches is denoised independently applying the collaborative strategy as proposed in [16].

Since the fusion method described in Section 3 operates in the DCT domain, it can be integrated with the BM3D algorithm. For each group of 2D patches belonging to a particular exposure, a collaborative 2D+1D transform and threshold is applied. After applying the inverse 1D transform, and prior to the computation of the inverse 2D-DCT transforms, we group the 2D DCT coefficients of the patches belonging to the same selected spatial location and having different exposures. The multi-exposure fusion described by Eqs (1) and (3) is performed on each of these groups. After this fusion step, only one DCT transformed patch is obtained at each selected location, to which the inverse DCT is applied. Fig 4 illustrates the process.

By repeating the operation on each patch location and applying aggregation in the image domain, we obtain the final denoised and fused result. In our implementation no second iteration of the algorithm is performed.
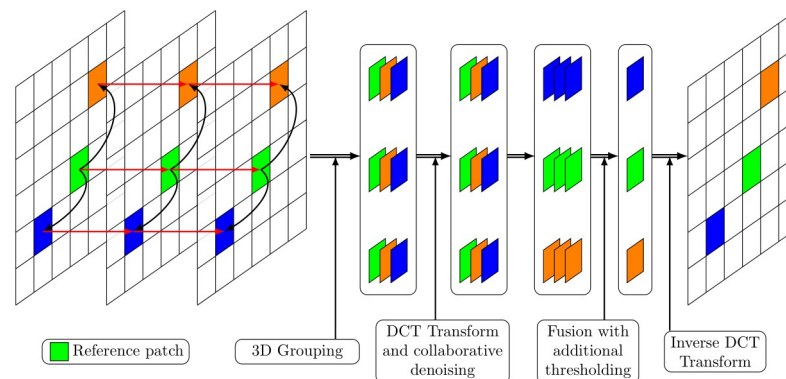
Fig 4. Processing scheme of a specific reference patch.

https://doi.org/10.1371/journal.pone.0265464.g004

5 Discussion and experimental results

In this section we compare the proposed method with state of the art algorithms for exposure fusion. We compare with Mertens et al. [1], Ma et al. [5], Li et al. [44], Kou et al. [28], Ma et al. [55], Hessel et al, (EEF) [19], Xu et al. [35], Zhang et al. (IFCNN) [64], Li et al. (CNNFEAT) [63], Hayat et al. (MEF-Sift) [22] and Martorell et al. [2]. The results from Ma et al. [5] and Ma et al. [55] were computed with the software downloaded from the corresponding author's webpage. The results of Mertens et al. [1] were obtained from the dataset provided in [56, 72]. The code for Hessel et al. (EEF) [19, 73], Xu et al. (FusionDN) [35, 74], Zhang et al.(IFCNN) [64, 75], Li et al. (CNNFEAT) [63, 76] and Hayat et al. (MEF-Sift) [22, 77] were obtained from the corresponding GitHub webpages. The code by Xu et al., originally proposed to fuse only pairs of images, has been adapted to fuse any sequence. We fuse the first two images of the set, this output is then combined with the third input image, and so on until all the images in the sequence have been fused. The results from Li et al. [44], Kou et al. [28] and Martorell et al. [2] were computed with the code provided by the authors. In all cases, default parameter settings are adopted.

Our results were computed using the same parameters for all the tests in this section. We use patches of size 8 × 8 pixels which is the standard size used by patch based denoising algorithms, as for example the BM3D [16]. The threshold for the collaborative filtering is set to the standard value 2.7σ. For the fusion stage, we fixed $p = 7$ as the power exponent of the coefficient magnitudes in Eq (2), $\sigma_g = 0.3$ and $\sigma_l = 0.5$ for the combination of the $\xi = 0$ coefficients. These latter parameters were set experimentally.

A sliding window approach is applied for the DCT based denoising/fusion. Once it is processed, the window is moved along both directions with a displacement step of $N_{step} = 2$. The fact that the whole window is fused permits the processing of all the pixels in the image.

5.1 Noise-free sequences

In this case, we just compare the ability of fusing the different exposure images. We use for our method just the fusion algorithm described in section 3, but without applying any thresholding of the DCT. Note that none of the compared methods include this noise filtering step.

Fig 5 displays the results of all the methods on the "Belgium House" data set. Most methods produce a result with good global illumination. However, looking closer at the images in Fig 6, we observe that many outdoor details in Mertens et al. [1] and Ma et al. [5] are overexposed. Li et al. [44] and Kou et al. [28] are not able to maintain the letters on the blackboard on the right side of Fig 6 and Ma et al. [55] is not able to preserve the details on the tree at the top of the image. Li et al. [63] the fusion result is over-smoothed, while the result by Hayat et al. [22] is over-saturated at the bright parts. The results of [2], Hessel et al [19], Zhang et al. [64] and ours are quite similar.

5.2 Noisy sequences with white uniform Gaussian noise

We compare in Fig 7 the results of all the algorithms on a noisy multi exposure sequence obtained after adding noise with standard deviation $\sigma = 15$ to the clean images (see Fig 8). We apply all the algorithms with their default parameters. It is clear from this figure, that none of the methods (except ours) is well adapted to the presence of noise. See Fig 9 for details of the results.

In the next experiment we add noise of standard deviation 25 to the clean multi-exposed images (see Fig 8) and apply the original BM3D algorithm [16] to denoise each one of them. We then apply the different multi exposure fusion methods to the denoised data (except our method, which is applied directly to the noisy sequence) and display the results in Fig 10. An
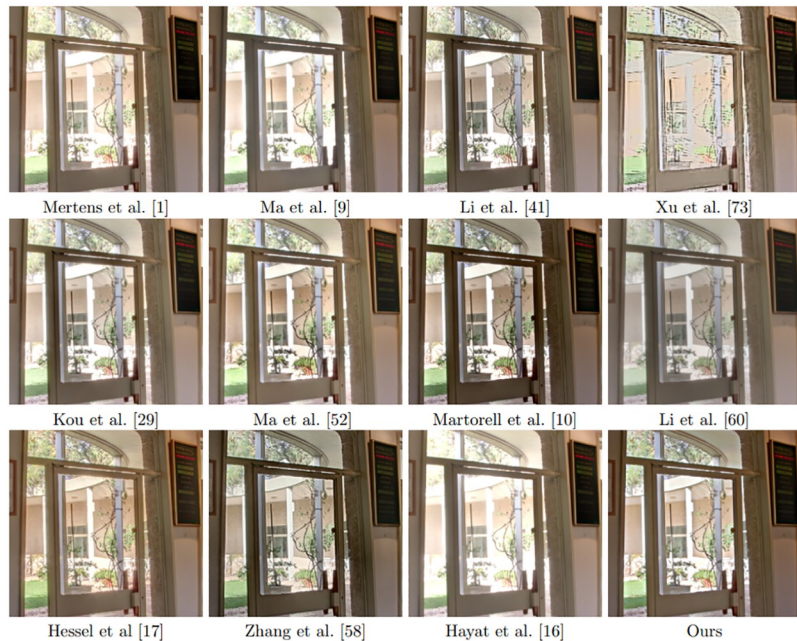
Fig 5. Results of fusion of noise-free multi-exposure images with different methods.

excerpt of the results is zoomed in and displayed in Fig 11. We observe that our method is the only one able to denoise and fuse the multi exposure sequence without producing noticeable artifacts.

5.2.1 Numerical evaluation. We propose a quantitative evaluation of the results by adapting the MEF-SSIM index proposed by Ma et al. [56] to take into account that the initial images are noisy.

Fig 6. Excerpt of the results shown in Fig 5.

Fig 7. Results of fusion of noisy multi-exposure images with different methods. The noise standard deviation of each input image is 15.

We begin by briefly describing this measure. Ma et al. decompose each patch $\mathbf{x}_k$, $k = 1, 2, \ldots, K$ of the set of input images as

$$\mathbf{x}_k = ||\mathbf{x}_k - \mu_{\mathbf{x}_k}|| \cdot \frac{\mathbf{x}_k - \mu_{\mathbf{x}_k}}{||\mathbf{x}_k - \mu_{\mathbf{x}_k}||} + \mu_{\mathbf{x}_k} = c_k \cdot \mathbf{s}_k + l_k \tag{8}$$

where $c_k = ||\mathbf{x}_k - \mu_{\mathbf{x}_k}||$, $l_k = \mu_{\mathbf{x}_k}$ and $\mathbf{s}_k = (\mathbf{x}_k - \mu_{\mathbf{x}_k})/||\mathbf{x}_k - \mu_{\mathbf{x}_k}||$ roughly represent the contrast, luminance and structure components of a patch $\mathbf{x}_k$. Using the previous decomposition, they compute the desired contrast and structure of the output image, respectively, as

$$\hat{c} = \max_{k=1,\cdots,K} c_k = \max_{k=1,\cdots,K} ||\mathbf{x}_k - \mu_{\mathbf{x}_k}||. \tag{9}$$
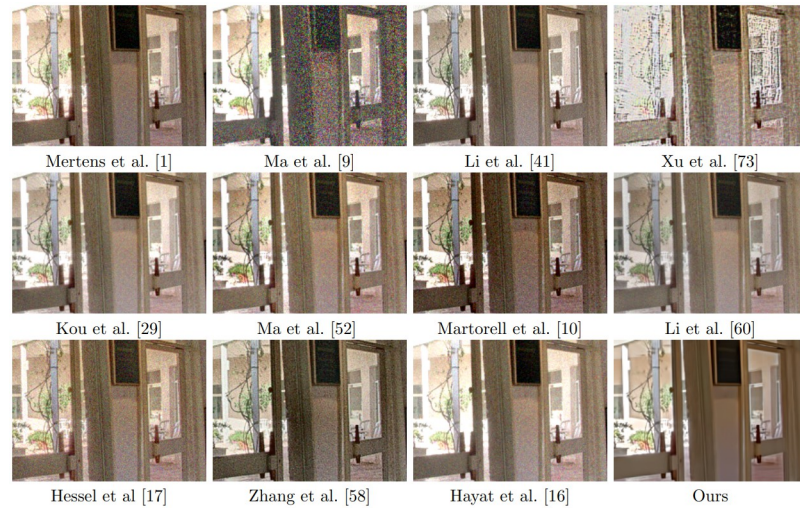
and

$$\bar{\mathbf{s}} = \frac{\sum_{k=1}^{K} w(\mathbf{x}_k - \mu_{\mathbf{x}_k})\mathbf{s}_k}{\sum_{k=1}^{K} w(\mathbf{x}_k - \mu_{\mathbf{x}_k})}, \qquad \hat{s} = \frac{\bar{s}}{||\bar{s}||}, \tag{10}$$

Fig 8. Noisy multi-exposure data sets used for comparison. On the first row, the noise standard deviation of each input image is 15. On the second row the standard deviation is 25.

Fig 9. Excerpt of the results shown in Fig 8. It is clear from this figure, that our method is the only one that takes noise into account.

where $w(\cdot)$ is a weighting function that determines the contribution of each source image patch in the structure of the fused image patch (see Ma et al. [56] for more details on these weights.) With that, the desired output patch result is

$$\hat{\mathbf{x}} = \hat{c} \cdot \hat{\mathbf{s}}. \tag{11}$$

Finally, the value that measures the structural similarity between a set of input patches $\{\mathbf{x}_k\}$, $k = 1, 2, \ldots, K$ from a sequence of multi-exposure images and the corresponding patch of the fused image $\mathbf{y}$ is given by

$$S(\{\mathbf{x}_k\}, \mathbf{y}) = \frac{2\sigma_{\hat{\mathbf{x}}\mathbf{y}} + C}{\sigma_{\hat{\mathbf{x}}}^2 + \sigma_{\mathbf{y}^2} + C}. \tag{12}$$
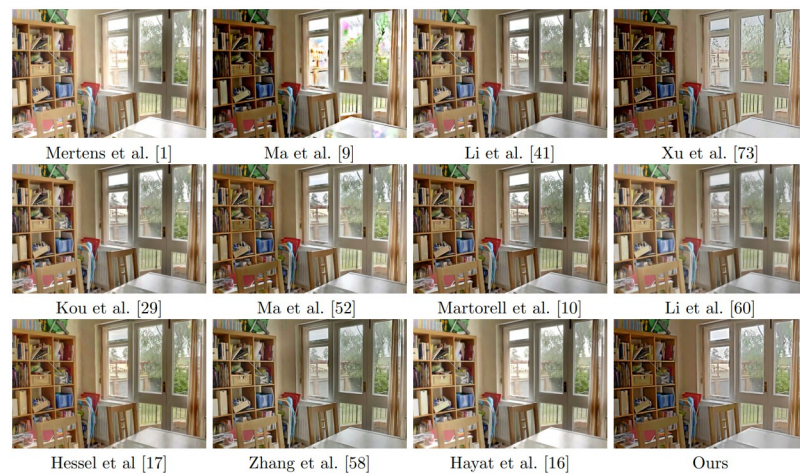
Fig 10. Results of fusion and denoising. Our method is the only one applied directly to the noisymulti-exposure images. The rest of methods fuse denoised versions of the images obtained using the BM3D algorithm. The noise standard deviation of each input image is 25.
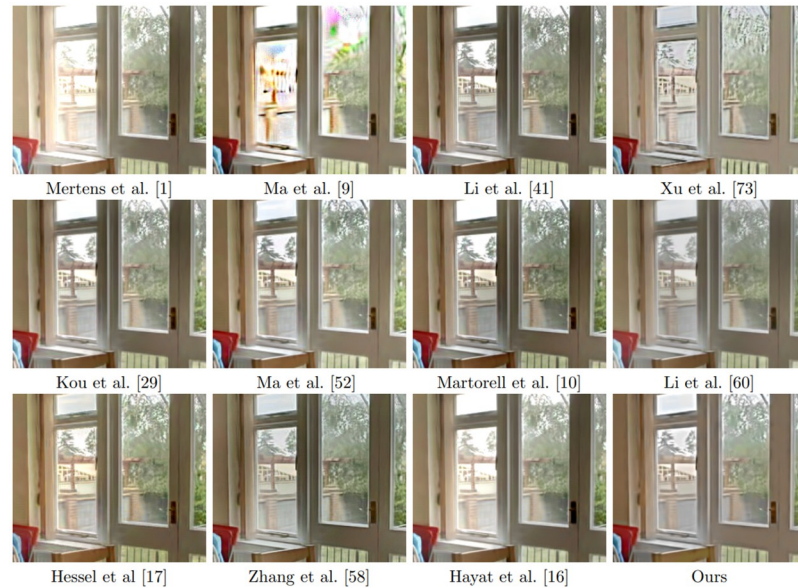
Fig 11. Excerpt of the results shown in Fig 10.

https://doi.org/10.1371/journal.pone.0265464.g011

The final measure is given by the mean of $S(\{\mathbf{x}_k\}, \mathbf{y})$ with $\{\mathbf{x}_k\}$ centered at each pixel of the image

$$Q(\mathbf{Y}) = \frac{1}{M} \sum_{j=1}^{M} S(\{\mathbf{x}_k\}(j), \mathbf{y}(j)). \quad (13)$$

The defined strength of the signal $\hat{c}$ needs to be modified in order to subtract the noise energy. Indeed, the strength of the noisy patches writes as the sum of the signal and noise energies. In order to evaluate the noise removal in the fused image $\mathbf{y}$, we redefine

$$\hat{c} = \sqrt{\max(\hat{c}^2 - \sigma^2, 0)}. \quad (14)$$

We denote this modified measure as MEF-SSIM$_n$. Larger values indicate a better performance of the method.

The values of MEF-SSIM$_n$ for the examples of Figs 7 and 10 are displayed in Table 1. As it can be seen, our method has the best MEF-SSIM$_n$ score.

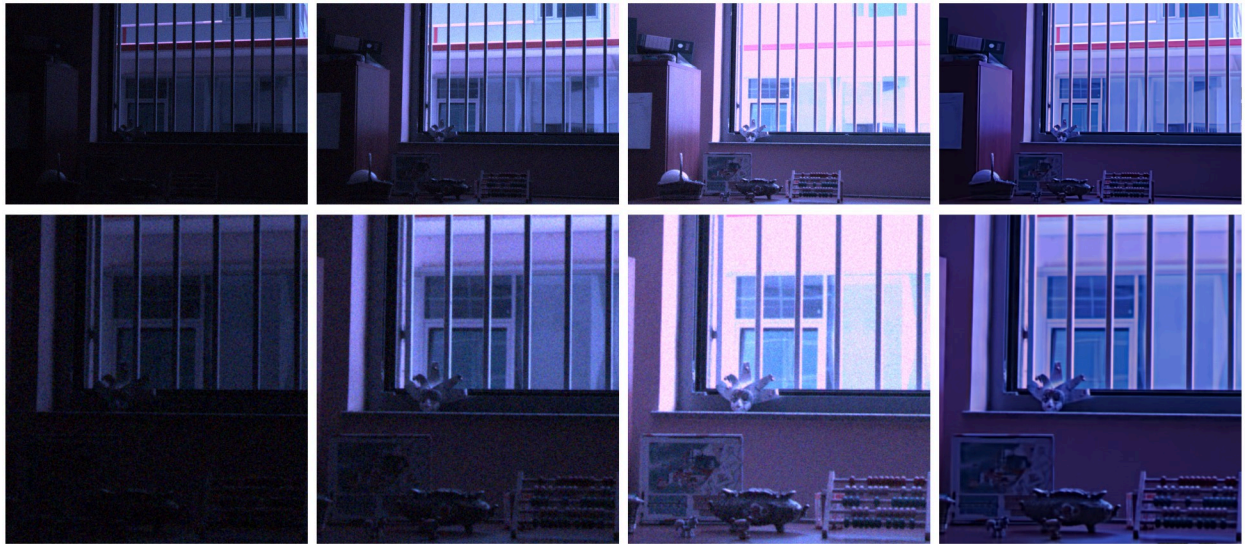## 5.3 Realistic multi exposure noisy images

The AWGN model does not hold in practice for real photographs. The noise is approximately white and additive at the camera sensor. However, it is signal dependent, meaning that the noise standard deviation at each pixel depends on its noise-free value. The noise characteristics

Table 1. Values of MEF-SSIM$_n$ for the examples of Figs 7 and 10.

	Mertens et al. [1]	Ma et al. [5]	Li et al. [44]	Xu et al. [35]	Kou et al. [28]	Ma et al. [55]	Martorell et al. [2]	Li et al. [63]	Hessel et al [19]	Zhang et al. [64]	Hayat et al. [22]	Ours
Fig 7	0.607	0.349	0.611	0.567	0.666	0.599	0.570	0.695	0.553	0.572	0.634	0.841
Fig 10	0.728	0.673	0.727	0.689	0.740	0.730	0.738	0.755	0.709	0.709	0.742	0.758

https://doi.org/10.1371/journal.pone.0265464.t001

Fig 12. Realistic noise simulation. From left to right: three noisy images with different exposures and the fusion and noise removal result. Below an excerpt of each image.

are then modified by the camera processing pipeline, consisting of demosaicking, color processing, gamma correction and compression [78].

In order to test our fusion and noise removal method with a 'realistic' noise case, we use RAW images which contain the acquired data at the sensor of the camera. This data can be obtained by selecting the RAW as output format when using professional reflex cameras. We add signal dependent white noise to a set of multi exposure RAW images and simulate a typical camera processing pipeline, obtaining their final color version in a common graphics format (PNG, JPEG, etc.).

The results in Fig 12 show that, even in this situation, the proposed method is able to denoise and fuse the images without introducing artifacts.

5.4 Computational analysis

The time complexity of the proposed algorithm is $\mathcal{O}(|Y|)$, where $|Y|$ denotes the size of each image in the multi-exposure sequence.

Assuming that the 3D transforms used for the collaborative filtering are performed in a separable way (i.e. 2D transforms followed by 1D transforms, as detailed in [16]), the overall number of operations of the algorithm, per pixel, is approximately

$$\mathcal{C}_{\mathcal{T}_{2D}} + 2Kb^2 N_S^2 + 2Kb^2 \mathcal{C}_{\mathcal{T}_{1D}} + 2Kkb^2 + k\mathcal{C}'_{\mathcal{T}_{2D}} + kb^2 \tag{15}$$

where:

- $\mathcal{C}_{\mathcal{T}_{2D}}$ denotes the number of operations required to compute the 2D DCTof the stack of patches similar to the one centered at the considered pixel. If we consider a neighborhood of size $N_S \times N_S$ around the pixel, this implies the computation of KN_S^2 2D DCTs, where K is the number of images in the sequence. The time complexity can be reduced by pre-computing the transforms in each block of size $K \times N_S \times N_S$ and reusing them in overlapping blocks, similarly to what is proposed in [16].

- The second term accounts for the 3D block matching step. This implies the exhaustive search, in a $N_S \times N_S$ neighborhood of the pixel, of 3D blocks of size $K \times b \times b$.

- The third term counts the number of operations for the computation of the 1D DCT transforms (direct and inverse) of the k nearest neighbors of each patch, in each frame. $\mathcal{C}_{T_{1D}}$ denotes the cost of computing a 1D DCT transform (direct or inverse) of a vector of size k.

- The fourth term accounts for the fusion step, which involves k 3D blocks of patches of size $K \times b \times b$.

- The fifth term accounts for the number of operations needed to compute the k inverse 2D DCT transforms of the fused patches. $\mathcal{C}'_{T_{2D}}$ denotes the cost of computing the inverse 2D DCT transform of a patch of $b \times b$ pixels.

- Finally, the last term counts the number of operations involved in the image aggregation step.

Observe that the number of denoising operations per pixel, for each image, is smaller than that for the original BM3D algorithm, since only one step of the collaborative filtering is applied. In addition, the inverse DCT is applied only to the fused patches, since it is not necessary to denoise each individual image of the multi-exposure set.

Moreover, the previous estimation assumes that an exhaustive-search algorithm has been used for block matching. The costs $\mathcal{C}_T$ of the DCT transforms depends on the availability of fast algorithms. By using predictive search techniques and fast separable transforms the complexity of the algorithm could be significantly reduced. Moreover, the overall number of operations can be further reduced by processing only one out of each $N_{\text{step}} < b$ pixels in both the horizontal and vertical directions. Due to the overlapping of the patches, the aggregation step used in the final step of the algorithm guarantees that all the pixels are correctly processed. In this case, the overall complexity of the method is reduced by a factor N_{step}^2.

6 Conclusions

In this paper we propose a patch-based method for the simultaneous denoising and fusion of a sequence of multi-exposed images. Both tasks are performed in the DCT domain and take advantage of a collaborative 3D thesholding approach similar to BM3D [16] for denoising, and the proposed fusion technique. For the collaborative denoising, a spatio-temporal criterion is used to select similar patches along the sequence, following the approach in [8]. The overall strategy permits to denoise and fuse the set of images without the need of recovering each denoised image image in the multi-exposure set, leading to a very efficient procedure.

Several experiments show that the proposed method permits to obtain state-of-the-art fusion results even when the input images are noisy. As future work, we plan to extend the current approach to multi-exposed video sequences.

Acknowledgments

Author Contributions

Methodology: Antoni Buades, Jose Luis Lisani.

Resources: Antoni Buades.

Software: Antoni Buades, Jose Luis Lisani, Onofre Martorell.

Validation: Antoni Buades, Jose Luis Lisani.

Visualization: Antoni Buades, Jose Luis Lisani, Onofre Martorell.

Writing – original draft: Antoni Buades, Jose Luis Lisani.

Writing – review & editing: Antoni Buades, Jose Luis Lisani, Onofre Martorell.

References

1. Mertens T, Kautz J, Van Reeth F. Exposure Fusion: A Simple and Practical Alternative to High Dynamic Range Photography. Computer Graphics Forum. 2009. https://doi.org/10.1111/j.1467-8659.2008.01171.x

2. Martorell O, Sbert C, Buades A. Ghosting-free DCT based multi-exposure image fusion. Signal Processing: Image Communication. 2019; 78:409–425.

3. Pérez P, Gangnet M, Blake A. Poisson image editing. In: ACM Transactions on graphics (TOG). vol. 22. ACM; 2003. p. 313–318.

4. Raskar R, Ilie A, Yu J. Image fusion for context enhancement and video surrealism. In: ACM SIG-GRAPH 2005 Courses. ACM; 2005. p. 4.

5. Ma K, Li H, Yong H, Wang Z, Meng D, Zhang L. Robust Multi-Exposure Image Fusion: A Structural Patch Decomposition Approach. IEEE Trans Image Processing. 2017; 26(5):2519–2532. https://doi.org/10.1109/TIP.2017.2671921

6. Mertens T, Kautz J, Reeth FV. Exposure Fusion. In: Proceedings of the 15th Pacific Conference on Computer Graphics and Applications. PG'07. Washington, DC, USA: IEEE Computer Society; 2007. p. 382–390. Available from: http://dx.doi.org/10.1109/PG.2007.23.

7. Ancuti CO, Ancuti C, De Vleeschouwer C, Bovik AC. Single-scale fusion: An effective approach to merging images. IEEE Transactions on Image Processing. 2017; 26(1):65–78. PMID: 27810821

8. Buades A, Lisani JL, Miladinović M. Patch-based video denoising with optical flow estimation. IEEE Transactions on Image Processing. 2016; 25(6):2573–2586. https://doi.org/10.1109/TIP.2016.2551639 PMID: 27071173

9. Haro G, Buades A, Morel JM. Photographing paintings by image fusion. SIAM Journal on Imaging Sciences. 2012; 5(3):1055–1087. https://doi.org/10.1137/120873923

10. Joshi N, Cohen MF. Seeing Mt. Rainier: Lucky imaging for multi-image denoising, sharpening, and haze removal. In: Computational Photography (ICCP), 2010 IEEE International Conference on. IEEE; 2010. p. 1–8.

11. Zheng M, Qi G, Zhu Z, Li Y, Wei H, Liu Y. Image dehazing by an artificial image fusion method based on adaptive structure decomposition. IEEE Sensors Journal. 2020; 20(14):8062–8072. https://doi.org/10.1109/JSEN.2020.2981719

12. Zhu Z, Wei H, Hu G, Li Y, Qi G, Mazur N. A novel fast single image dehazing algorithm based on artificial multiexposure image fusion. IEEE Transactions on Instrumentation and Measurement. 2020; 70:1–23.

13. Rav-Acha A, Peleg S. Two motion-blurred images are better than one. Pattern recognition letters. 2005; 26(3):311–317. https://doi.org/10.1016/j.patrec.2004.10.017

14. Delbracio M, Sapiro G. Removing Camera Shake via Weighted Fourier Burst Accumulation. Image Processing, IEEE Transactions on. 2015; 24(11):3293–3307. https://doi.org/10.1109/TIP.2015.2442914 PMID: 26068313

15. Delbracio M, Sapiro G. Hand-Held Video Deblurring Via Efficient Fourier Aggregation. Computational Imaging, IEEE Transactions on. 2015; 1(4):270–283. https://doi.org/10.1109/TCI.2015.2501245

16. Dabov K, Foi A, Katkovnik V, Egiazarian K. Image denoising by sparse 3-D transform-domain collaborative filtering. IEEE Transactions on image processing. 2007; 16(8):2080–2095. https://doi.org/10.1109/TIP.2007.901238 PMID: 17688213

17. An J, Lee SH, Kuk JG, Cho NI. A multi-exposure image fusion algorithm without ghost effect. In: Acoustics, Speech and Signal Processing (ICASSP), 2011 IEEE International Conference on. IEEE; 2011. p. 1565–1568.

18. Liu Y, Wang Z. Dense SIFT for ghost-free multi-exposure fusion. Journal of Visual Communication and Image Representation. 2015; 31:208–224. https://doi.org/10.1016/j.jvcir.2015.06.021

19. Hessel C. Extended Exposure Fusion. Image Processing On Line. 2019; 9:453–468. https://doi.org/10.5201/ipol.2019.278

20. Li S, Kang X. Fast multi-exposure image fusion with median filter and recursive filter. IEEE Transactions on Consumer Electronics. 2012; 58(2). https://doi.org/10.1109/TCE.2012.6227469

21. Ocampo-Blandon C, Gousseau Y. Non-local Exposure Fusion. In: Iberoamerican Congress on Pattern Recognition. Springer; 2016. p. 484–492.

22. Hayat N, Imran M. Ghost-free multi exposure image fusion technique using dense SIFT descriptor and guided filter. Journal of Visual Communication and Image Representation. 2019; 62:295–308. https://doi.org/10.1016/j.jvcir.2019.06.002

23. Zhang W, Cham WK. Gradient-directed multiexposure composition. IEEE Transactions on Image Processing. 2012; 21(4):2318–2323. https://doi.org/10.1109/TIP.2011.2170079 PMID: 21965210

24. Gu B, Li W, Wong J, Zhu M, Wang M. Gradient field multi-exposure images fusion for high dynamic range image visualization. Journal of Visual Communication and Image Representation. 2012; 23 (4):604—610. https://doi.org/10.1016/j.jvcir.2012.02.009

25. Sun J, Zhu H, Xu Z, Han C. Poisson image fusion based on Markov random field fusion model. Information fusion. 2013; 14(3):241–254. https://doi.org/10.1016/j.inffus.2012.07.003

26. Kuk JG, Cho NI, Lee SU. High dynamic range (HDR) imaging by gradient domain fusion. In: Acoustics, Speech and Signal Processing (ICASSP), 2011 IEEE International Conference on. IEEE; 2011. p. 1461–1464.

27. Kou F, Li Z, Wen C, Chen W. Edge-preserving smoothing pyramid based multi-scale exposure fusion. Journal of Visual Communication and Image Representation. 2018; 53:235–244. https://doi.org/10.1016/j.jvcir.2018.03.020

28. Kou F, Li Z, Wen C, Chen W. Multi-scale exposure fusion via gradient domain guided image filtering. In: 2017 IEEE International Conference on Multimedia and Expo (ICME). IEEE; 2017. p. 1105–1110.

29. Burt PJ, Adelson EH. The Laplacian Pyramid as a Compact Image Code. IEEE TRANSACTIONS ON COMMUNICATIONS. 1983; 31:532–540. https://doi.org/10.1109/TCOM.1983.1095851

30. Goshtasby AA. Fusion of Multi-exposure Images. Image Vision Comput. 2005; 23(6):611–618. https://doi.org/10.1016/j.imavis.2005.02.004

31. Zhang W, Hu S, Liu K. Patch-based correlation for deghosting in exposure fusion. Information Sciences. 2017; 415:19–27. https://doi.org/10.1016/j.ins.2017.05.019

32. Zhang W, Hu S, Liu K, Yao J. Motion-free exposure fusion based on inter-consistency and intra-consistency. Information Sciences. 2017; 376:190–201. https://doi.org/10.1016/j.ins.2016.10.020

33. Gonzalez RC, Wood RE. Digital image processing, 2nd Edtn.

34. Paul S, Sevcenco IS, Agathoklis P. Multi-exposure and multi-focus image fusion in gradient domain. Journal of Circuits, Systems and Computers. 2016; 25(10):1650123. https://doi.org/10.1142/S0218126616501231

35. Xu H, Ma J, Le Z, Jiang J, Guo X. Fusiondn: A unified densely connected network for image fusion. In: Proceedings of the AAAI Conference on Artificial Intelligence. vol. 34; 2020. p. 12484–12491.

36. Hu J, Gallo O, Pulli K, Sun X. HDR deghosting: How to deal with saturation? In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition; 2013. p. 1163–1170.

37. Li Z, Zheng J, Zhu Z, Wu S. Selectively detail-enhanced fusion of differently exposed images with moving objects. IEEE Transactions on Image Processing. 2014; 23(10):4372–4382. https://doi.org/10.1109/TIP.2014.2349432 PMID: 25148667

38. Zheng J, Li Z. Superpixel based patch match for differently exposed images with moving objects and camera movements. In: 2015 IEEE International Conference on Image Processing (ICIP). IEEE; 2015. p. 4516–4520.

39. Tico M, Gelfand N, Pulli K. Motion-blur-free exposure fusion. In: Image Processing (ICIP), 2010 17th IEEE International Conference on. IEEE; 2010. p. 3321–3324.

40. Prabhakar KR, Babu RV. Ghosting-free multi-exposure image fusion in gradient domain. In: Acoustics, Speech and Signal Processing (ICASSP), 2016 IEEE International Conference on. IEEE; 2016. p. 1766–1770.

41. Sie WR, Hsu CT. Alignment-free exposure fusion of image pairs. In: Image Processing (ICIP), 2014 IEEE International Conference on. IEEE; 2014. p. 1802–1806.

42. Hu J, Gallo O, Pulli K. Exposure stacks of live scenes with hand-held cameras. Computer Vision–ECCV 2012. 2012; p. 499–512. https://doi.org/10.1007/978-3-642-33718-5_36

43. Li ZG, Zheng JH, Rahardja S. Detail-enhanced exposure fusion. IEEE Trans Image Processing. 2012; 21(11):4672–4676. https://doi.org/10.1109/TIP.2012.2207396 PMID: 22801512

44. Li Z, Wei Z, Wen C, Zheng J. Detail-Enhanced Multi-Scale Exposure Fusion. IEEE Transactions on Image Processing. 2017; 26(3):1243–1252. https://doi.org/10.1109/TIP.2017.2651366 PMID: 28092537

45. Singh H, Kumar V, Bhooshan S. A novel approach for detail-enhanced exposure fusion using guided filter. The Scientific World Journal. 2014; 2014(659217):8. https://doi.org/10.1155/2014/659217 PMID: 24683353

46. Raman S, Chaudhuri S. Bilateral Filter Based Compositing for Variable Exposure Photography. In: Alliez P, Magnor M, editors. Eurographics 2009—Short Papers. The Eurographics Association; 2009.

47. Li S, Kang X, Hu J. Image fusion with guided filtering. IEEE Trans Image Processing. 2013; 22 (7):2864–2875. https://doi.org/10.1109/TIP.2013.2244222 PMID: 23372084

48. Tomasi C, Manduchi R. Bilateral filtering for gray and color images. In: Computer Vision, 1998. Sixth International Conference on. IEEE; 1998. p. 839–846.

49. He K, Sun J, Tang X. Guided image filtering. IEEE transactions on pattern analysis and machine intelligence. 2013; 35(6):1397–1409. https://doi.org/10.1109/TPAMI.2012.213 PMID: 23599054

50. Piella G. Image fusion for enhanced visualization: A variational approach. International Journal of Computer Vision. 2009; 83(1):1–11. https://doi.org/10.1007/s11263-009-0206-4

51. Bertalmio M, Levine S. Variational approach for the fusion of exposure bracketed pairs. IEEE Transactions on Image Processing. 2013; 22(2):712–723. https://doi.org/10.1109/TIP.2012.2221730 PMID: 23047876

52. Hafner D, Weickert J. Variational exposure fusion with optimal local contrast. In: International Conference on Scale Space and Variational Methods in Computer Vision. Springer; 2015. p. 425–436.

53. Ma K, Yeganeh H, Zeng K, Wang Z. High dynamic range image compression by optimizing tone mapped image quality index. IEEE Transactions on Image Processing. 2015; 24(10):3086–3097. https://doi.org/10.1109/TIP.2015.2436340 PMID: 26011881

54. Laparra V, Berardino A, Ballé J, Simoncelli EP. Perceptually optimized image rendering. JOSA A. 2017; 34(9):1511–1525. https://doi.org/10.1364/JOSAA.34.001511 PMID: 29036154

55. Ma K, Duanmu Z, Yeganeh H, Wang Z. Multi-exposure image fusion by optimizing a structural similarity index. IEEE Transactions on Computational Imaging. 2018; 4(1):60–72. https://doi.org/10.1109/TCI. 2017.2786138

56. Ma K, Zeng K, Wang Z. Perceptual quality assessment for multi-exposure image fusion. IEEE Transactions on Image Processing. 2015; 24(11):3345–3356. https://doi.org/10.1109/TIP.2015.2442920 PMID: 26068317

57. Kalantari NK, Ramamoorthi R. Deep high dynamic range imaging of dynamic scenes. ACM Trans Graph. 2017; 36(4):144–1. https://doi.org/10.1145/3072959.3073609

58. Wu S, Xu J, Tai YW, Tang CK. Deep High Dynamic Range Imaging with Large Foreground Motions. In: Proceedings of the European Conference on Computer Vision (ECCV); 2018. p. 117–132.

59. Xu H, Ma J, Zhang XP. MEF-GAN: Multi-exposure image fusion via generative adversarial networks. IEEE Transactions on Image Processing. 2020; 29:7203–7216. https://doi.org/10.1109/TIP.2020. 2999855

60. Prabhakar KR, Srikar VS, Babu RV. DeepFuse: A Deep Unsupervised Approach for Exposure Fusion with Extreme Exposure Image Pairs. In: ICCV; 2017. p. 4724–4732.

61. Prabhakar KR, Arora R, Singh KP, Babu RV. A fast, scalable and reliable deghosting method for extreme exposure fusion. In: 2019 IEEE International Conference on Computational Photography (ICCP). IEEE; 2019.

62. Xu H, Ma J, Jiang J, Guo X, Ling H. U2Fusion: A Unified Unsupervised Image Fusion Network. IEEE Transactions on Pattern Analysis and Machine Intelligence. 2020.

63. Li H, Zhang L. Multi-Exposure Fusion with CNN Features. In: 2018 25th IEEE International Conference on Image Processing (ICIP); 2018. p. 1723–1727.

64. Zhang Y, Liu Y, Sun P, Yan H, Zhao X, Zhang L. IFCNN: A general image fusion framework based on convolutional neural network. Information Fusion. 2020; 54:99–118. https://doi.org/10.1016/j.inffus. 2019.07.011

65. Zhang H, Xu H, Xiao Y, Guo X, Ma J. Rethinking the image fusion: A fast unified image fusion network based on proportional maintenance of gradient and intensity. In: Proceedings of the AAAI Conference on Artificial Intelligence. vol. 34; 2020. p. 12797–12804.

66. Akyüz AO, Reinhard E. Noise reduction in high dynamic range imaging. Journal of Visual Communication and Image Representation. 2007; 18(5):366–376. https://doi.org/10.1016/j.jvcir.2007.04.001

67. Min TH, Park RH, Chang S. Noise reduction in high dynamic range images. Signal, Image and Video Processing. 2011; 5(3):315–328. https://doi.org/10.1007/s11760-010-0203-7

68. Lee DK, Park RH, Chang S. Ghost and noise removal in exposure fusion for high dynamic range imaging. International Journal of Computer Graphics & Animation. 2014; 4(4):1. https://doi.org/10.5121/ijcga.2014.4401

69. Ahmad A, Riaz MM, Ghafoor A, Zaidi T. Noise resistant fusion for multi-exposure sensors. IEEE Sensors Journal. 2016; 16(13):5123–5124. https://doi.org/10.1109/JSEN.2016.2556715

70. Roman, Steven and Axler, S and Gehring, FW Advanced linear algebra Springer, 2005.

71. Yu Guoshen and Sapiro Guillermo DCT Image Denoising: a Simple and Effective Image Denoising Algorithm, Image Processing On Line. 2011; 1:292–296. https://doi.org/10.5201/ipol.2011.ys-dct

72. Zeng K, Ma K, Hassen R, Wang Z. Perceptual evaluation of multi-exposure image fusion algorithms. In: 2014 Sixth International Workshop on Quality of Multimedia Experience (QoMEX); 2014. p. 7–12.

73. Hessel C. EEF source code. http://www.ipol.im/pub/art/2019/278/?utm_source=doi, 2019. [Online; accessed 7-March-2022].

74. Xu H, Ma J, Le Z, Jiang J, Guo X. Fusiondn source code. https://github.com/hanna-xu/FusionDN, 2020. [Online; accessed 7-March-2022].

75. Zhang Y, Liu Y, Sun P, Yan H, Zhao X, Zhang L. IFCNN source code. https://github.com/uzeful/IFCNN, 2019. [Online; accessed 7-March-2022].

76. Li H, Zhang L. CNNFEAT source code. https://github.com/xiaohuiben/MEF-CNN-feature, 2018. [Online; accessed 7-March-2022].

77. Hayat N, Imran M. MEF-Sift source code. https://github.com/ImranNust/Source-Code, 2019. [Online; accessed 7-March-2022].

78. Ramanath R., Snyder W. E., Yoo Y., & Drew M. S. Color image processing pipeline. IEEE Signal Processing Magazine, 22(1), 34–43, 2005. https://doi.org/10.1109/MSP.2005.1407713