Signal Processing 178 (2021) 107791

Contents lists available at ScienceDirect

Signal Processing

journal homepage: www.elsevier.com/locate/sigpro

# Multi-exposure image fusion based on linear embeddings and watershed masking

Oguzhan Ulucan, Diclehan Karakaya, Mehmet Turkan*

*Department of Electrical and Electronics Engineering, Izmir University of Economics, Izmir, Turkey*

ARTICLE INFO

*Article history:*
Received 1 May 2020
Revised 18 June 2020
Accepted 28 August 2020
Available online 9 September 2020

*Keywords:*
Multi-exposure fusion
Linear embedding
Watershed masking
High dynamic range imaging

ABSTRACT

High dynamic range imaging (HDRI) is a challenging technology but yet demanding for modern imaging applications. Low-cost image sensors have limited dynamic range, and it is not always possible to capture and display natural scenes with high contrast and information loss in any exposure is inevitable. Three solutions for HDRI are using expensive high dynamic range (HDR) cameras with HDR-compatible displays, tone mapping operators for low dynamic range (LDR) screens, and capturing and fusing multiple exposures of the same LDR scene via image fusion algorithms. Companies that produce user grade devices prefer multi-exposure fusion (MEF) approaches to obtain HDR-like images for LDR screens due to its low cost. Hence, merging a stack of images containing different exposures of the same scene into a single informative image is an attractive research field. In this study, a novel, simple yet effective method is proposed for static image exposure fusion. The developed technique is based on weight map extraction via linear embeddings and watershed masking. The main advantage lies in watershed masking-based adjustment for obtaining accurate weights for image fusion. The comprehensive experimental comparisons demonstrate very strong visual and statistical results, and this approach should facilitate future MEF studies.



## 1. Introduction

Capturing visual content through cameras with limited dynamic range may cause undesirable outcomes, i.e., faulty exposure leads to unwanted bright or dark regions in images, hence information loss is inevitable for details in highlights or shadows. The information lost in such regions, however, might be present in a set of distinctly exposed images. The fundamental aim of MEF is to retain the most informative parts of different exposed images (e.g., under, normal and over) and blend them into a single informative image by using pixel based or region (block) based approaches, e.g., Ma and Wang [1], Mertens et al. [2].

In order to capture and visualize HDR content, users should employ an HDR camera with an HDR compatible screen. In contrast, most of the display devices today project images in 8 bits per pixel for each color channel, known as LDR. For LDR compatible displays, tone mapping operators need to be applied to HDR content or multiple exposures of the same LDR scene can be fused via image fusion algorithms referred to as MEF. MEF is commonly preferred in most consumer grade devices because of its ease of implementation and its lower cost compared to HDR cameras and HDR compatible display devices.

In recent decades, many successful MEF studies have been proposed in literature, e.g., Ma and Wang [1], Mertens et al. [2], Gu et al. [3], Li et al. [4], Li and Kang [5], Li et al. [6], Li and Zhang [7], Paul et al. [8], Ma et al. [9], Lee et al. [10], Hayat and Imran [11]. Each study mainly differs in the method of determining fusion weight maps of different exposures. The measurement of fusion weights is extremely challenging and, for both static and dynamic contents, the careful use of weight maps can eliminate the jitter effect in between exposures [12] as well as artifacts such as halos in the sharp changeovers in the reconstructed images. For instance, a pixel based approach in [2] measures contrast, saturation and well-exposedness features as fusion maps. The reconstructed image is obtained as a linear weighted combination of images in the exposure stack. This method is one of the algorithms selected for comparing the proposed method in this paper, as one of the most effective approaches for MEF, despite its relative age. In [13], image fusion is carried out based on an edge preserving filter, which is known as the bilateral filter. The main goal of this study is to blend under- and over-exposed parts of the same scene, while preserving fine details in these images. However, this method provides statistically weak results when the image sequence contains multiple over-exposed images. An approach used in [3], which is

* Corresponding author.
*E-mail address:* mehmet.turkan@iue.edu.tr (M. Turkan).

https://doi.org/10.1016/j.sigpro.2020.107791
0165-1684/

tolerable for slight movements usually resulting in ghosting effects in the case of dynamic images, is proposed to fuse static images. The proposed method is based on the gradient field modification and the distance of intensities in feature space is calculated via Euclidean metric. Although this study demonstrates promising results in fusing LDR images, its performance could be improved by choosing a more effective metric, based on human visual system models. In [4], MEF is performed with fine details enhancement. A quadratic optimization based method is used for extracting fine details from each image in the stack. Previously existing methods are used to obtain an initial fused image, which is enhanced with the vector of fine details. This technique can significantly increase the overall quality of HDR-like images combined with the existing MEF algorithms. The study in [5] achieves the fusion of both static and dynamic image sequences. The proposed method extracts the following quality measurements: local contrast and brightness for static images, and color dissimilarity weight for dynamic images. While recursive filter is used for smoothing the weight maps in both static and dynamic images, a novel histogram equalization approach and median filter are used to detect the motion in dynamic images to overcome the greatest challenge in dynamic MEF, i.e., ghosting effects. In another successful work [6], a novel method based on guided filtering is designed for not only MEF, but also for multi-spectral, multi-focus and multi-modal fusion. A two-scale image reconstruction is carried out for obtaining the final fused image. The base and detail layers of each image are fused separately in the first step, and then the final image is obtained through a combination of the fused base layer and the fused detail layer. In [8], an algorithm is designed for both MEF and multi-focus fusion purposes. Since the human visual system is more responsive to luminance than chrominance channels, fusion of these channels is carried out independently. While the fusion of chrominance channels is performed via simply taking a weighted sum of the input chrominance values, input luminance channels are fused in the gradient domain via a wavelet-based gradient blending algorithm. A current MEF study [10] extracts adaptive weights from pixel intensities and gradient information. In order to determine the pixel quality in each exposure, two weight maps are characterized via two functions. The first weight map is based on an adapted version of the well-exposedness feature introduced in [2]. The second map is extracted from the gradient information in each exposure. Finally, the fused image is obtained via the pyramidal image decomposition approach as in [2].

A recent successful development in MEF is described in [9]. In this study, signal strength, signal structure and mean intensity are calculated to obtain fusion weight maps. The fused images obtained via these weight maps prove that this technique gives greatly improved results, not only for static cases but also for dynamic image sequences. According to experimental results, the designed method generates sharply-detailed, high quality color MEF images with minimum ghosting effects. In [11], a successful MEF approach based on guided filtering and dense-SIFT descriptor is proposed. The method utilizes brightness, local contrast and color dissimilarity features as weights. The local contrast feature is obtained through dense-SIFT, while the color dissimilarity feature is acquired via histogram equalization and median filtering as in [5]. A guided filter is then applied to these weights to eliminate the noise and discontinuities, followed by an adopted pyramid decomposition for fusion.

After presenting their potential with several image processing and machine learning applications, convolutional neural networks (CNNs) are employed in many MEF studies, e.g., Li and Zhang [7], Que and Lee [14], Liu and Leung [15], Hu et al. [16], Li et al. [17]. For instance, several pretrained classification, super-resolution and denoising networks are compared in [7] in terms of their feature extraction capacity for MEF applications. According to the reported results, rather than using the third layer of a denoising or a super-resolution network, the first layer of a classification network should be employed since it is more efficient for feature extraction and it has a lower computational cost. After feature extraction, local visibility and temporal consistency parameters are calculated to determine the feature maps for the general weighted fusion operation. Additionally in [15], a CNNs-based method applicable to both decolorization and MEF is proposed. The local gradient information in different exposures are used as inputs for the network, resulting in satisfactory quantitative and qualitative results. Nevertheless, a drawback of this study is that the proposed fusion network called "FusionNet" functions on image stacks consisting of only three exposures.

This paper proposes a simple yet effective algorithm based on linear embeddings [18] and watershed masking [19] to fuse a stack of static images with different exposures. While initial fusion maps are extracted through linear embeddings of image pixel/patch spaces, a watershed masking procedure is used for adjusting these maps to refine informative parts of images for the final fusion step. The performance of the proposed approach is compared with the well-known MEF algorithms using the perceptual quality assessment method introduced in [20]. This study presents promising visual and statistical experimental results, and will guide future work and expand this research domain. The rest of this paper is organized as follows. Section 2 explains the proposed novel MEF approach. Experimental results and detailed visual and statistical comparisons are presented in Section 3. Finally, a brief conclusion with possible future directions is given in Section 4.

## 2. The proposed method of exposure fusion

As mentioned previously, MEF studies are diverse in extracting fusion weights, and the fine adjustment of these weight maps is challenging. The proposed approach depends on patch-based weight estimation (per pixel) via linear embeddings of images and watershed masking to obtain global fusion maps. To the best of our knowledge, the proposed MEF method is a novel framework, taking the advantage of linear embeddings of image pixel/patch spaces and watershed masks of images.

A simple flowchart of the proposed MEF algorithm is demonstrated in Fig. 1. The block diagram in this figure illustrates two branches for extracting linear embedding weights and watershed masks from the given stack of image exposures. These maps and masks are then combined to obtain global fusion masks in order to obtain a fused image. In the following, the developed method is detailed including the determination of three main exposures; linear embeddings of pixel/patch spaces of images for extracting weight maps; the well-known watershed masking together with the reasons why it is preferred rather than a binary masking; and finally, the blending procedure of multi-exposure images to obtain HDR-like fusion, which is later to be enhanced in a final post-processing step.

### 2.1. Determination of exposures

Although it can be extended to any number of exposures, the proposed MEF algorithm blends three different exposures to obtain a fused image. These exposures will be referred to as the *under exposed image* $\mathcal{U}$, *normal exposed image* $\mathcal{N}$ and *over exposed image* $\mathcal{O}$ in remaining parts of this study. These exposures needs to be carefully determined especially when the exposure stack has more than three images.

As a straightforward approach, histograms of different exposures can be utilized effectively in clustering the stack of input images into three classes. Therefore, histograms are directly employed

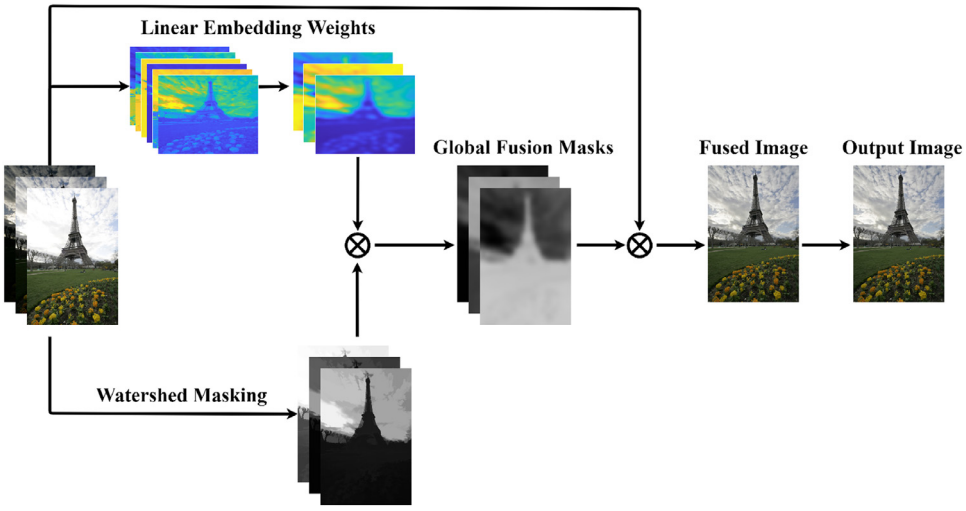



**Fig. 1.** A simple flowchart of the proposed MEF method.


**Fig. 2.** The flowchart of the determination of exposures.

as features in $k$-means clustering to determine the three exposure clusters of images. A sliding window based averaging technique is then applied in each cluster to obtain three different exposures (i.e., $\mathcal{U}$, $\mathcal{N}$, $\mathcal{O}$) of images to be used in the fusion process. A simple flowchart of the determination of exposures is illustrated in Fig. 2, and the corresponding algorithm is summarized in Algorithm 1.

**Algorithm 1** Determination of exposures.

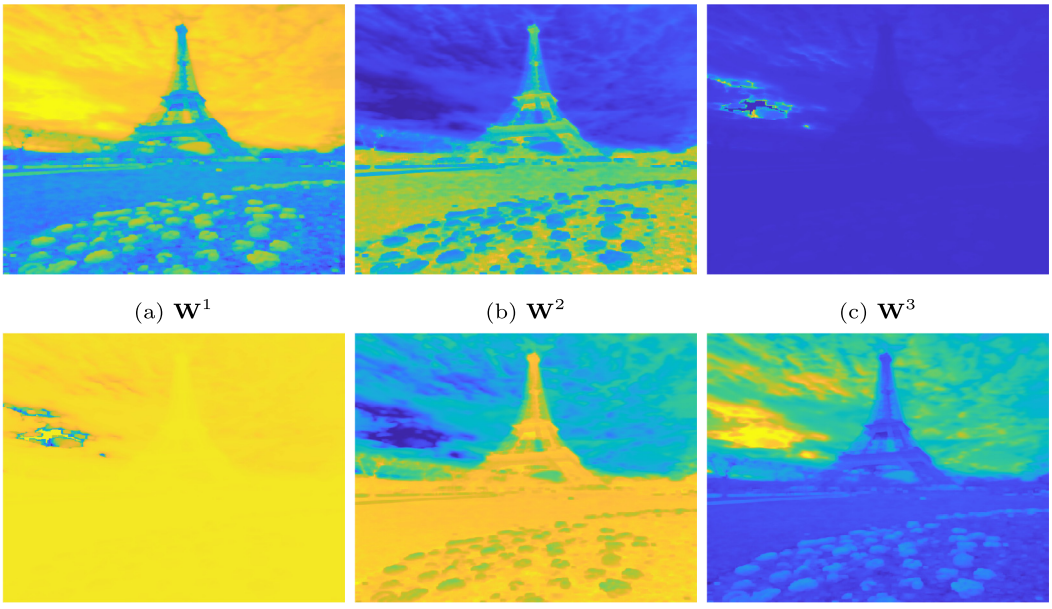
**Inputs:** Image stack $\{I_n\}$, $n = 1 \dots N$
**Outputs:** Determined exposures $\{\mathcal{U}, \mathcal{N}, \mathcal{O}\}$

**for all** Image $I_n \in \{I_n\}$ **do**
  $I_g = rgb2gray(I_n)$
  $\mathbf{v}_n = histogram(I_g)$
$label(i) = kMeans(\{\mathbf{v}_n\})$, $i = 1, 2, 3$
$\{\mathcal{U}, \mathcal{N}, \mathcal{O}\} = SlidingWindow(\forall I_n \in label(i))$

The sliding window technique is based on a simple averaging procedure to combine images in the same cluster (after $k$-means clustering). Instead of a pixel based averaging in these clusters, the averaging operation here operates on a very small neighborhood of pixel patches (e.g., $5 \times 5$ pixels) and overlaps between adjacent patches are implicitly allowed which are finally averaged uniformly in these regions. This forces all image patches to agree on the overlapped areas, hence satisfying local compatibility and smoothness while reducing the noise and artifacts in the combined image per exposure cluster.

### 2.2. Weight maps via linear embeddings

In image processing, there is an observation which suggests that natural images are sampled from low-dimensional manifolds. Hence, densely-sampled images, or rather small texture patches, can be successfully reconstructed as a linear combination of their neighbors. This is generally referred to as neighbor embedding in image processing tasks [21–24], and is inspired by the manifold learning algorithms for dimensionality reduction [18,25,26]. In this study, the manifold sampling assumption is unified with exposure images in a given stack, which results in a new framework for weight map extraction in MEF. The general idea originates from the well-known dimensionality reduction technique called locally linear embedding (LLE) [18], based on the assumption that each exposure image is sampled from a manifold structure, and all these exposures should lie on or close to a locally linear patch of the underlying sampled manifold.



(a) $\mathbf{W}^1$ (b) $\mathbf{W}^2$ (c) $\mathbf{W}^3$

(d) $\mathbf{W}^4$ (e) $\mathbf{W}^5$ (f) $\mathbf{W}^6$

**Fig. 3.** An example of the extracted weight maps via linear embedding.

After determining three exposure images, namely, $\mathcal{U}$ (under exposed), $\mathcal{N}$ (normal exposed) and $\mathcal{O}$ (over exposed) via clustering and sliding window, a patch based scheme is designed for characterizing the intrinsic properties of manifold structures of spatially-local pixel/patch spaces of these exposures. While representing three exposure images by matrices **U, N** and **O** of sizes $r \times c$ pixels, $n \times n$ image patches are extracted with lexical ordering of image pixels, i.e., $i = 1 \dots rc$. Note that $n \times n$ patches $\mathbf{u}_i$, $\mathbf{n}_i$ and $\mathbf{o}_i$ extracted from **U, N** and **O** respectively, are centered around the pixel indexed by $i$ and these are all collocated patches originating from each different exposure. These image patches are later transformed into the stacked column vectors of size $n^2 \times 1$. The parameter $n$ is fixed to a sufficiently small neighborhood of size 5 pixels. Let us denote these patch triplets in a set as $\mathcal{P} = \{\mathbf{u}_i, \mathbf{n}_i, \mathbf{o}_i\}_{\forall i}$.

The main objective is to characterize point based structures by means of patch manifolds through spatially collocated exposure patches. Local geometric properties of each $n \times n$ neighborhood indexed by $i$ can be linearly characterized by solving three optimization problems given in Eq. (1) as follows,

$$\begin{aligned}
\{\mathbf{W}_i^1, \mathbf{W}_i^2\} &= \arg\min_{\{w_1, w_2\}} \|\mathbf{o}_i - [\mathbf{u}_i \;\; \mathbf{n}_i][w_1 \;\; w_2]^T\|_2^2 \quad \text{s.t.} \quad w_1 + w_2 = 1 \\
\{\mathbf{W}_i^3, \mathbf{W}_i^4\} &= \arg\min_{\{w_3, w_4\}} \|\mathbf{u}_i - [\mathbf{n}_i \;\; \mathbf{o}_i][w_3 \;\; w_4]^T\|_2^2 \quad \text{s.t.} \quad w_3 + w_4 = 1 \\
\{\mathbf{W}_i^5, \mathbf{W}_i^6\} &= \arg\min_{\{w_5, w_6\}} \|\mathbf{n}_i - [\mathbf{u}_i \;\; \mathbf{o}_i][w_5 \;\; w_6]^T\|_2^2 \quad \text{s.t.} \quad w_5 + w_6 = 1
\end{aligned} \tag{1}$$

where each individual patch in the set $\mathcal{P}$ is linearly embedded into the remaining two exposure patch subspaces leading to a set of weights $\{w_j\}_{j=1}^6$ and $\left\{\mathbf{W}_i^j\right\}_{j=1}^6$ denotes the set of linear embedding weight maps at each pixel location $i$, $\forall i$. Note that there exists a sum-to-one constraint in each optimization in order to enforce the approximation to lie in the subspace of the patch to be embedded and also to provide invariance to translations. The optimization problems in Eq. (1) can be easily solved by means of an inner product (Gram) matrix similar to [18].

An example of the extracted six weight maps from the *Tower* stack is given in Fig. 3. It can be clearly observed that each embedding weight map highlights specific parts of the exposures to be fused. In short, $\mathbf{W}^1$ and $\mathbf{W}^5$ originate from $\mathcal{U}$ for reconstructing $\mathcal{O}$ and $\mathcal{N}$, respectively. Similarly, $\mathbf{W}^2$ and $\mathbf{W}^3$ come from the image $\mathcal{N}$ for $\mathcal{O}$ and $\mathcal{U}$; and $\mathbf{W}^4$ and $\mathbf{W}^6$ are extracted from $\mathcal{O}$ for $\mathcal{U}$ and $\mathcal{N}$. These weight map pairs obtained from the same exposures are later combined absolutely in Eq. (2) to form fused linear embedding maps, namely $\mathbf{E}_1'$, $\mathbf{E}_2'$ and $\mathbf{E}_3'$ for $\mathcal{U}$, $\mathcal{N}$ and $\mathcal{O}$, respectively.

$$\begin{aligned}
\mathbf{E}_1' &= |\mathbf{W}^1| + |\mathbf{W}^5| \\
\mathbf{E}_2' &= |\mathbf{W}^2| + |\mathbf{W}^3| \\
\mathbf{E}_3' &= |\mathbf{W}^4| + |\mathbf{W}^6|
\end{aligned} \tag{2}$$

Since $\mathbf{E}_1'$, $\mathbf{E}_2'$ and $\mathbf{E}_3'$ are calculated from different exposures in the image stack, they are normalized to sum-to-one and then smoothed to provide local smoothness in the transition regions while avoiding possible noise and artifacts. The resulting embedding weights $\mathbf{E}_1$, $\mathbf{E}_2$ and $\mathbf{E}_3$ are obtained in Eq. (3) as follows,

$$\mathbf{E}_k = \left(\mathbf{E}_k' \otimes [\sum_{k=1}^{3} \mathbf{E}_k']^{-1}\right) * \mathbf{G} \tag{3}$$

where $\mathbf{G}$ is a Gaussian smoothing kernel, $\otimes$ and $*$ denote the element-wise multiplication and the convolution operators, respectively. Fig. 4 exemplifies the obtained final linear embedding weight maps for the *Tower* stack, and the corresponding algorithm is summarized in Algorithm 2.

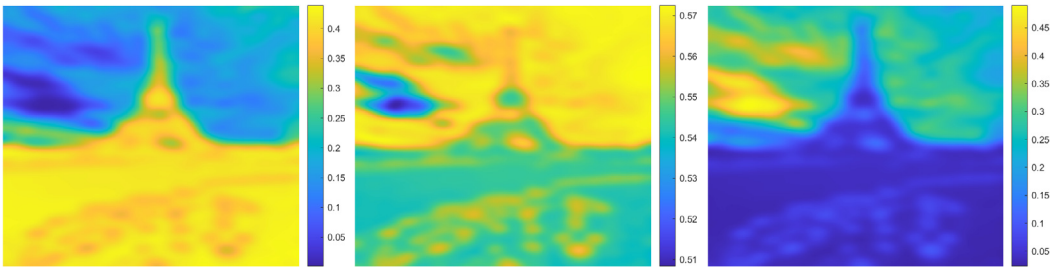
**Algorithm 2** Weight maps via linear embeddings.

**Inputs:** Determined exposures $\{\mathbf{U}, \mathbf{N}, \mathbf{O}\} \leftarrow \{\mathcal{U}, \mathcal{N}, \mathcal{O}\}$
**Outputs:** Embedding maps $\{\mathbf{E}_k\}$, $k = 1, 2, 3$

Create patch triplet set $\mathcal{P} = \{\mathbf{u}_i, \mathbf{n}_i, \mathbf{o}_i\}_{\forall i}$
**for all** Triplet $(\mathbf{u}_i, \mathbf{n}_i, \mathbf{o}_i) \in \{\mathbf{u}_i, \mathbf{n}_i, \mathbf{o}_i\}$ **do**
    Solve Eq. (1) to obtain $\{\mathbf{W}_i^j\}$, $j = 1 \dots 6$
Blend derived weight maps via Eq. (2) to obtain $\{\mathbf{E}_k'\}$
Normalize and smooth $\{\mathbf{E}_k'\}$ via Eq. (3) to obtain $\{\mathbf{E}_k\}$

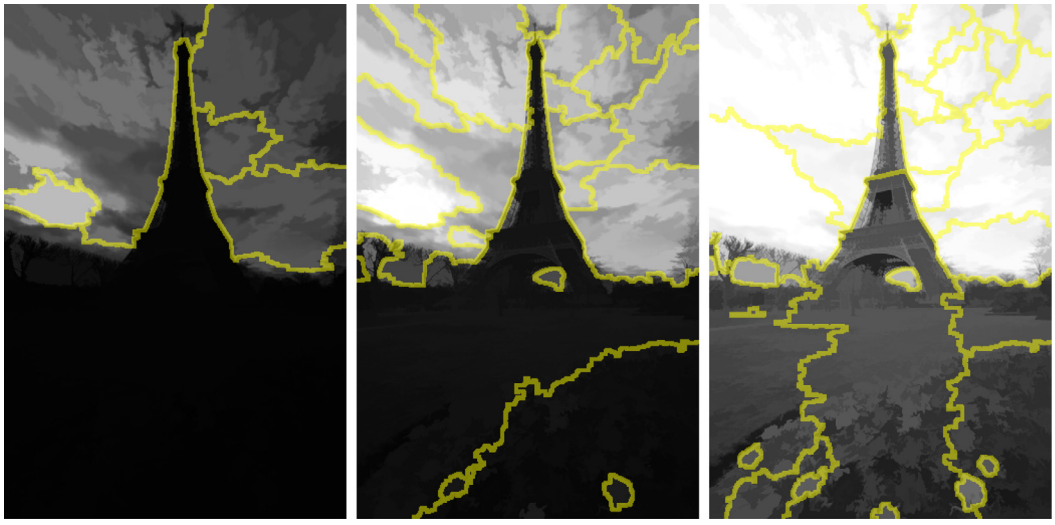
### 2.3. Watershed masking

A general strategy in MEF studies is to employ binary masks, e.g., hat function [7]. These masks often produce artifacts in the




(a) $\mathbf{E}_1$ (b) $\mathbf{E}_2$ (c) $\mathbf{E}_3$

**Fig. 4.** The final embedding maps to be used in the further stages of the algorithm.



(a) $\mathbf{M}_1$ (b) $\mathbf{M}_2$ (c) $\mathbf{M}_3$

**Fig. 5.** Watershed masks of the *Tower* stack.

(a) $\mathbf{G}_1$ (b) $\mathbf{G}_2$ (c) $\mathbf{G}_3$

**Fig. 6.** Global fusion masks of the *Tower* stack.

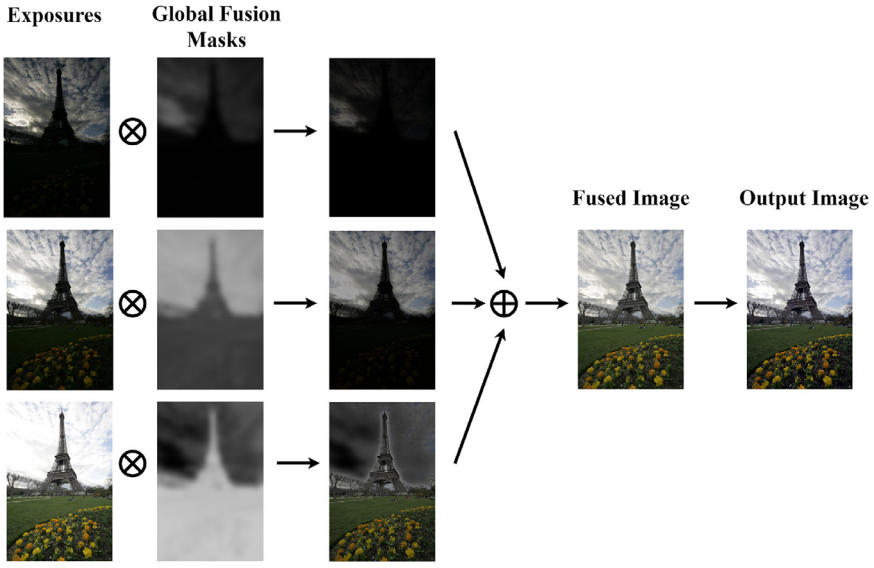



**Fig. 7.** The exposure fusion process.

**Table 1**
13 image stacks with different number of exposures used in the experiments.

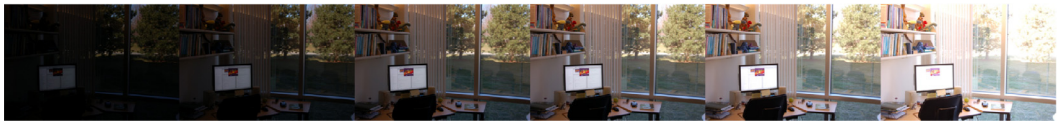
| **Name** | **Tower** | **Chinese Garden** | **Venice** | **Farmhouse** | **Landscape** |
|---|---|---|---|---|---|
| Size | 512 × 341 × 3 | 340 × 512 × 3 | 341 × 512 × 3 | 341 × 512 × 3 | 341 × 512 × 3 |
| **Name** | **Kluki** | **Cave** | **Office** | **Balloons** | **Belgium House** |
| Size | 341 × 512 × 3 | 384 × 512 × 4 | 340 × 512 × 6 | 339 × 512 × 9 | 384 × 512 × 9 |
| **Name** | **Lighthouse** | **Lamp** | **Madison Capital** | | |
| Size | 340 × 512 × 3 | 384 × 512 × 15 | 384 × 512 × 30 | | |

**Table 2**
MEF-SSIM scores for each stack given in Table 1. LE, GE, LEW and WSM fuse input exposures linearly with weight maps extracted by local energy, global energy, linear embeddings and watershed masking, respectively.

| | Algorithms | | | | | | | | | | | | Ours | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | LE | GE | Mertens | Raman | Gu | Li12 | S.Li | Li13 | Ma | Paul | H.Li | Liu | LEW | WSM | Proposed |
| *Tower* | 0.898 | 0.912 | **0.986** | 0.895 | 0.931 | 0.950 | 0.984 | **0.986** | **0.986** | 0.977 | 0.981 | 0.983 | 0.959 | 0.965 | 0.981 |
| *Chinese Garden* | 0.917 | 0.928 | 0.989 | 0.911 | 0.927 | 0.951 | 0.982 | 0.984 | 0.985 | 0.982 | 0.977 | 0.988 | 0.975 | 0.991 | **0.990** |
| *Venice* | 0.845 | 0.913 | 0.966 | 0.892 | 0.889 | 0.937 | 0.951 | 0.954 | 0.940 | 0.954 | 0.947 | 0.973 | 0.967 | 0.952 | **0.978** |
| *Farmhouse* | 0.941 | 0.916 | 0.981 | 0.877 | 0.932 | 0.958 | 0.977 | **0.985** | 0.984 | 0.971 | 0.974 | 0.978 | 0.929 | 0.973 | 0.978 |
| *Landscape* | 0.901 | 0.961 | 0.976 | 0.953 | 0.941 | 0.948 | 0.972 | 0.942 | 0.993 | 0.972 | 0.954 | **0.994** | 0.993 | 0.996 | 0.986 |
| *Lighthouse* | 0.793 | 0.944 | **0.980** | 0.938 | 0.934 | 0.968 | 0.953 | 0.950 | 0.970 | 0.965 | 0.962 | 0.985 | 0.978 | 0.973 | 0.975 |
| *Kluki* | 0.851 | 0.907 | **0.980** | 0.902 | 0.922 | 0.948 | 0.965 | 0.968 | 0.970 | 0.952 | 0.957 | 0.973 | 0.955 | 0.943 | 0.963 |
| *Cave* | 0.861 | 0.837 | 0.974 | 0.693 | 0.933 | 0.923 | 0.961 | **0.978** | 0.948 | 0.964 | 0.959 | 0.947 | 0.930 | 0.929 | 0.968 |
| *Office* | 0.831 | 0.955 | 0.984 | 0.907 | 0.899 | 0.954 | 0.972 | 0,967 | 0.988 | 0.973 | 0.970 | 0.979 | 0.985 | 0.987 | **0.991** |
| *Balloons* | 0.770 | 0.862 | **0.969** | 0.768 | 0.913 | 0.941 | 0.944 | 0.948 | 0.965 | 0.893 | 0.949 | 0.917 | 0.941 | 0.944 | 0.961 |
| *Belgium House* | 0.732 | 0.874 | 0.971 | 0.809 | 0.896 | 0.954 | 0.947 | 0.964 | **0, 973** | 0.899 | 0.945 | 0.946 | 0.945 | 0.953 | 0.965 |
| *Lamp* | 0.577 | 0.836 | **0.969** | 0.729 | 0.875 | 0.945 | 0.964 | 0.928 | 0.956 | 0.851 | 0.916 | 0.910 | 0.882 | 0.929 | 0.930 |
| *Madison Capitol* | 0.779 | 0.886 | 0.977 | 0.763 | 0.864 | 0.949 | 0.918 | 0.968 | **0.983** | 0.932 | 0.936 | 0.944 | 0.969 | 0.970 | 0.981 |
| **Average** | 0.822 | 0.902 | **0.977** | 0.849 | 0.912 | 0.948 | 0.961 | 0.967 | 0.973 | 0.945 | 0.956 | 0.963 | 0.953 | 0.963 | 0.973 |

regions where sharp texture and color changes occur in the scene. Regular smoothing filters can be applied to avoid these artifacts but this process may cause other undesirable artifacts, e.g., halo effects. Alternatively, edge-aware smoothing such as cross-bilateral filters can be employed; however, it is not trivial to control the spatial and intensity values via sufficient parameters in these type of filters. Instead, in this study, watershed segmentation [19] is adopted for mask extraction in order to acquire natural texture and color transitions while avoiding unwanted effects and artifacts in the fused image. It is worth mentioning here that, to the best available knowledge, this is the first time that the watershed segmentation has been employed in the MEF problem.

The watershed mask extraction process begins with the conversion of each exposure image $\mathcal{U}$, $\mathcal{N}$ and $\mathcal{O}$ into grayscale. Reconstruction-based opening and closing morphological operations [27] are then applied on these grayscale images. Similar to [28], an opening-by-reconstruction (erosion followed by a morphological reconstruction) is followed by a dilation and reconstruction in order to clean up images, i.e., to remove small blemishes without disturbing the overall structures. The aim is to avoid oversegmentation results in the watershed transform. A disk-shaped structuring element with radius 11 is employed to perform all these operations. In the rest of this paper, the outputs acquired from the above morphological operations, which are verified via watershed segmentation, will be referred to as watershed masks $\mathbf{M}_1$, $\mathbf{M}_2$ and $\mathbf{M}_3$ for $\mathcal{U}$, $\mathcal{N}$ and $\mathcal{O}$, respectively. The extracted masks of the *Tower* stack are shown in Fig. 5, and the corresponding algorithm is detailed in Algorithm 3.



(a) Six input exposures in the stack.

(b) Mertens (0.984) (c) S.Li (0.972) (d) Proposed (0.991)

**Fig. 8.** Visual comparison of the proposed method with Mertens and S.Li for *Office*.

(a) Three input exposures in the stack.

(b) Mertens (0.989) (c) Ma (0.985) (d) Proposed (0.990)

**Fig. 9.** Visual comparison of the proposed method with Mertens and Ma for *Chinese Garden*.

### 2.4. Exposure fusion

The global fusion masks are obtained in Eq. (4) via a linear combination of the extracted information contained in the watershed masks and the embedding weight maps as follows,

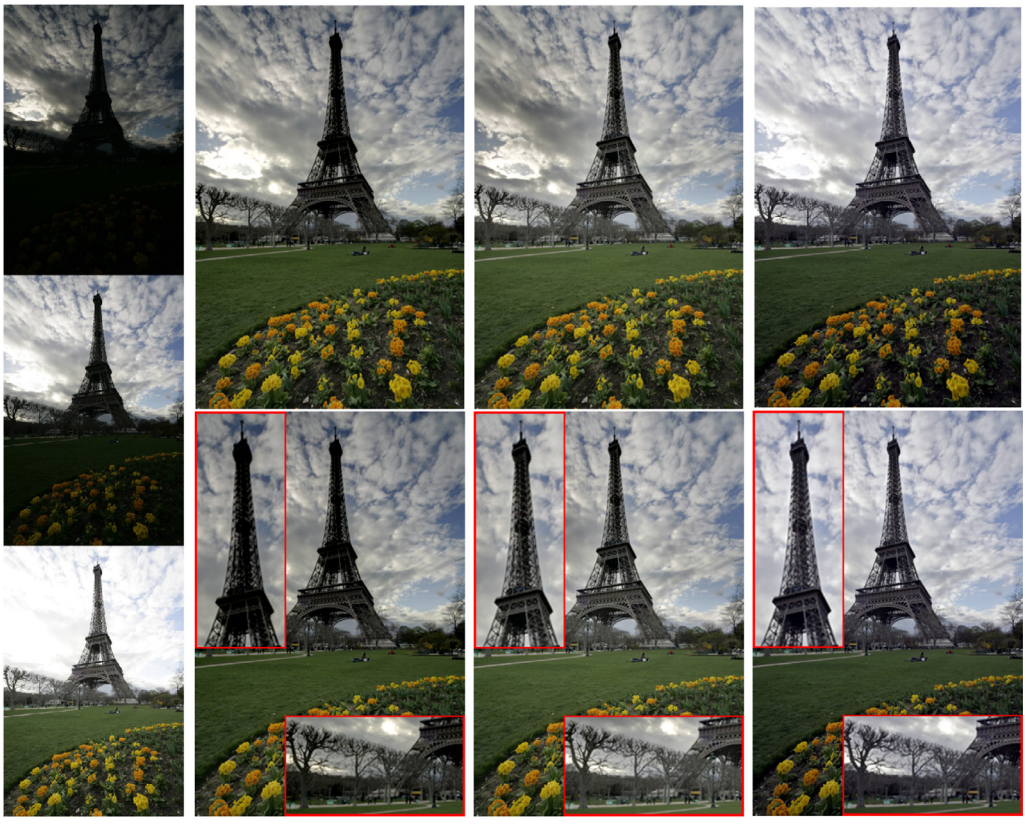
$$\begin{aligned} \mathbf{G_1} &= \mathbf{M_1} \otimes \mathbf{E_3} \\ \mathbf{G}_2 &= \mathbf{M}_2 \otimes \mathbf{E}_2 \\ \mathbf{G}_3 &= \mathbf{M}_3 \otimes \mathbf{E}_1 \end{aligned} \tag{4}$$

where $\mathbf{G}_1$, $\mathbf{G}_2$ and $\mathbf{G}_3$ represent global fusion masks for $\mathcal{U}$, $\mathcal{N}$ and $\mathcal{O}$, respectively. Note here that global fusion masks are obtained in a way that the extracted linear embedding information contained in over- and under-exposures are exchanged in between via the corresponding watershed masks. Therefore, well-exposed regions in both under- and over-exposures are highlighted, while the normal exposure image contributes to both. These global masks are demonstrated in Fig. 6 for the *Tower* stack.

(a) Three input exposures in the stack.

(b) Mertens (0.966) (c) Li13 (0.954) (d) Proposed (0.978)

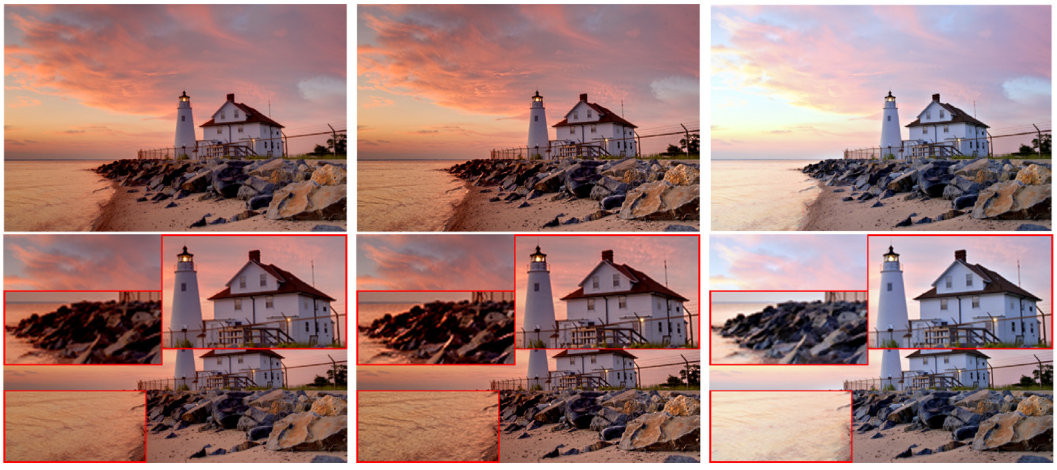
**Fig. 10.** Visual comparison of the proposed method with Mertens and Li13 for *Venice*.



(a) Exposures. (b) Mertens (0.986) (c) Li13 (0.986) (d) Proposed (0.981)

**Fig. 11.** Visual comparison of the proposed method with Mertens and Li13 for *Tower*.

(a) Three input exposures in the stack.



(b) Mertens (0.980) (c) Li12 (0.968) (d) Proposed (0.975)

**Fig. 12.** Visual comparison of the proposed method with Mertens and Li12 for *Lighthouse*.

(a) Four input exposures in the stack.

(b) Mertens (0.974) (c) Li13 (0.978) (d) Proposed (0.968)

**Fig. 13.** Visual comparison of the proposed method with Mertens and Li13 for *Cave*.

(a) Nine input exposures in the stack.

(b) Fused image without post-processing.

(c) Output image after post-processing.

(d) LE

(e) GE

(f) Gu

(g) S.Li

(h) Li12

(i) Li13

(j) Mertens

(k) Raman

(l) Ma

**Fig. 14.** Visual comparison of different methods for *Balloons*.

**Algorithm 3** Watershed masking.

**Inputs:** Input image $I \in \{\mathcal{U}, \mathcal{N}, \mathcal{O}\}$ and structuring element **SE**
**Outputs:** Watershed mask **M**

$I_g = rgb2gray(I)$
$I_e = imErode(I_g, \mathbf{SE})$
$I_{obr} = imReconstruct(I_e, I_g)$
$I_{obrd} = imDilate(I_{obr}, \mathbf{SE})$
$I_{obrcbr} = imReconstruct(\sim I_{obrd}, \sim I_{obr})$
$\mathbf{M} = \sim I_{obrcbr}$

The fused image **F** can then be recovered through a weighted blending of input images with the corresponding global fusion masks as given in Eq. (5). The exposure fusion process is illustrated in Fig. 7, and the corresponding algorithm is summarized in Algorithm 4.

**Algorithm 4** Exposure fusion.

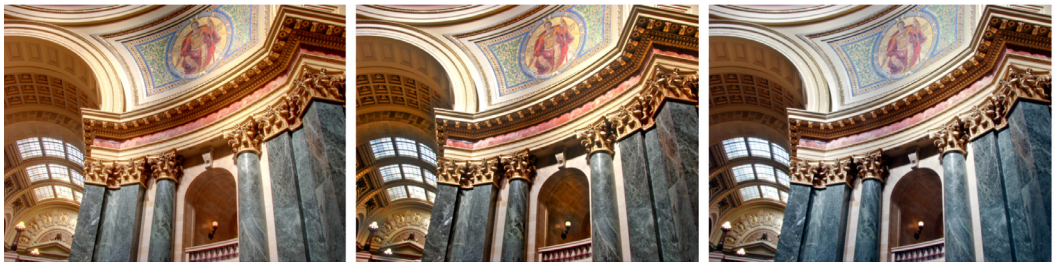
**Inputs:** $\{\mathbf{U}, \mathbf{N}, \mathbf{O}\}$, $\{\mathbf{E}_k\}$, $\{\mathbf{M}_k\}$, $k = 1, 2, 3$
**Outputs:** Fused image **F**

Derive global masks $\{\mathbf{G}_k\}$ via Eq. (4)
Apply Eq. (5) to obtain **F**

$$\mathbf{F} = \mathbf{U} \otimes \mathbf{G}_1 + \mathbf{N} \otimes \mathbf{G}_2 + \mathbf{O} \otimes \mathbf{G}_3. \tag{5}$$

After obtaining the fused image **F** via Eq. (5), a simple contrast enhancement based post-processing is employed in order to correct local low-light regions or unsatisfactory color intensities.

(a) The determined three exposures of the stack.



(b) Mertens (0.977) (c) Ma (0.983) (d) Proposed (0.981)

**Fig. 15.** Visual comparison of different methods for *Madison Capitol*.

(a) Three input exposures in the stack.

(b) Mertens (0.976) (c) Ma (0.993) (d) Proposed (0.986)

**Fig. 16.** Visual comparison of different methods for *Landscape*.

To achieve this, the top 1% and the bottom 1% of all pixel values of the image are saturated to stretch the contrast of **F**. The final output obtained through this post-processing presents both statistically superior results and visually more plausible and natural-looking images.

## 3. Experimental setup and results

### 3.1. Quality assessment metric for MEF

The performance of the proposed MEF approach is compared statistically with the other algorithms using the perceptual quality assessment method, which is a multi-scale structural similarity framework (MEF-SSIM) [20]. MEF-SSIM basically measures patch structural consistency for MEF and provides statistical analysis results in the range [0, 1], in which outcomes closer to 1 indicate better perceptual quality.

In order to assess the quality of the fused image, MEF-SSIM forms a multi-input (i.e., distinct exposures) structural comparison element based on the structural-similarity metric (SSIM) [29]. While neglecting the patch luminance components because of under/over exposedness, the structural comparison element ($S$) depends only on contrast and structure components of input images in Eq. (6) as follows,

$$S(\{x_n\}, y) = \frac{2\sigma_{\hat{x}y} + C}{\sigma_{\hat{x}}^2 + \sigma_y^2 + C} \tag{6}$$

where $\{x_n\}$ represents collocated set of patches in all $N$ images in the input stack and $y$ is the corresponding patch in the fused image. $\hat{x} = \hat{c} \cdot \hat{s}$ denotes the desired output (fused) patch as a function of the desired contrast $\hat{c}$, i.e., the highest contrast of $\{x_n\}$, and the desired structure $\hat{s}$, i.e., a weighted average of the input structure vectors. $\sigma_{\hat{x}}^2$ and $\sigma_y^2$ demonstrate local variances of $\hat{x}$ and $y$ respec-

(a) Three input exposures in the stack.

(b) Li13 (0.985) (c) Ma (0.984) (d) Proposed (0.978)

**Fig. 17.** Visual comparison of different methods for *Farmhouse*.

(a) The determined three exposures of the stack.

(b) Mertens (0.971) (c) Ma (0.973) (d) Proposed (0.965)

**Fig. 18.** Visual comparison of different methods for *Belgium House*.

tively, $\sigma_{\hat{x}y}$ is the local covariance between $\hat{x}$ and $y$. $C$ is a small constant handling the low contrast saturation effects [29].

The MEF-SSIM comparison is applied on local patches across the entire image, resulting in a spatial quality map which gives an indication of the structural quality. These local values are then averaged to acquire the overall MEF-SSIM score of the fused image. The luminance consistency in the fused image is further considered with a multi-scale extension by a set of scale-level quality scores.

### 3.2. Experimental results and discussion

In the first set of experiments, the proposed MEF algorithm is compared against 12 approaches in literature, including Mertens [2], Gu [3], Li12 [4], S.Li [5], Li13 [6], H.Li [7], Paul [8], Ma [9], Raman [13] and Liu [15] over 13 different image stacks with different number of exposures [20,30]. The details of these image stacks are summarized in Table 1.

Table 2 presents the detailed MEF-SSIM scores, as well as the average accuracy rates (in the final row), acquired through different algorithms. In this table, LE and GE stand for two simple methods which linearly combine input exposures using local energy and global energy as weight maps, respectively. Furthermore, LEW and WSM denote linear embedding weights and watershed masking applied individually to the MEF problem as weight maps, respectively. As a final note here, the default settings for each algorithm, including the proposed method, are adopted for comparison without any optimization. The results provided for Ma [9] are obtained by executing the code reached from the official web-page of the author in [31]. In addition, the results reported for Paul [8], H.Li [7] and Liu [15] are obtained by executing the code reached from web-pages [32–34], respectively. Since Fusion-

(a) Three input exposures in the stack.

(b) Mertens (0.980) (c) Ma (0.970) (d) Proposed (0.963)

**Fig. 19.** Visual comparison of different methods for *Kluki*.

(a) The determined three exposures of the stack.

(b) Mertens (0.969) (c) S.Li (0.964) (d) Proposed (0.930)

**Fig. 20.** Visual comparison of different methods for *Lamp*.

Net [15] only works with three input exposures, the proposed algorithm in Algorithm 1 is applied to accord with this limit.

The proposed algorithm produces highly competitive results among all MEF-SSIM scores and it is able to outperform some of the state-of-the-art approaches on the average with the given image sequences. Moreover, linear embedding weights (LEW) and watershed masking (WSM) are employed individually in the fusion process in order to demonstrate the impact of the extracted weights in the proposed algorithm. Although these individual weights sometimes outperform their combination, the combined algorithm is clearly more effective on average. It is also worth noting that the proposed method produces visually more plausible results for several image sequences in cases when statistical results do not provide the best results reported in Table 2.

The fused images obtained for *Office*, *Chinese Garden* and *Venice* result in 0.991, 0.990, 0.978 MEF-SSIM scores respectively, which provide superior results compared with the other methods. As observed in Fig. 8 for *Office*, the proposed MEF algorithm produces better visual details in the shelf region, especially for the toys, when compared with Mertens and S.Li. Moreover, the specific features of the Mathworks Environment and the *Peppers* image can be seen more clearly on the computer screen. However, it is important to note that there exists a slight saturation problem (e.g., the pathway behind the tree) in the proposed result. The algorithm tends to lose information in cases with excessively over-exposed images in the stack. Nevertheless, this method produces the best score statistically, with visually plausible output image for this stack.

As seen in Fig. 9, specific features in *Chinese Garden* are better preserved in the output of the proposed algorithm. While the colors for sky appear artificial in Mertens, both Ma and the proposed method uncover a natural sky scene. Furthermore, it is visible that Ma loses some information in the region of the moving man (en-

(a) Three input exposures in the stack.

(b) Mertens (c) Ma (d) Proposed

**Fig. 21.** Visual comparison of different methods for *Flowers*.

(a) Three input exposures in the stack.

(b) Mertens (c) Ma (d) Proposed

**Fig. 22.** Visual comparison of different methods for *SeaRock*.

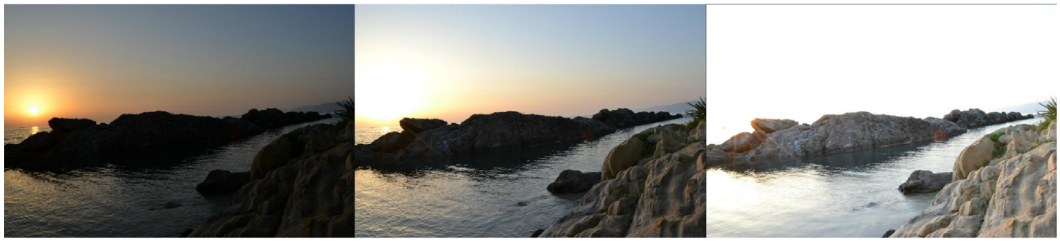
larged in the last row of Fig. 9) whereas Mertens and the proposed method result in output images which preserve these details subject to lower brightness. Additionally, the reflections on the water are more informative and visually plausible in the proposed output. However, the details on the rooftop and ivies seem to have a lower contrast when compared to Ma.

In the fused *Venice* image obtained via the proposed method, the details are greatly preserved as shown in Fig. 10. Even though the algorithm produces a brighter image than the remaining methods, the fine details are generally much more effectively recovered. In particular, the sky is more natural and clouds, more vivid. The specific features on the boat can be easily distinguished and the details are more visible on the scene, while some details are lost (e.g., of the boat) in Mertens, which has the second best MEF-SSIM score.

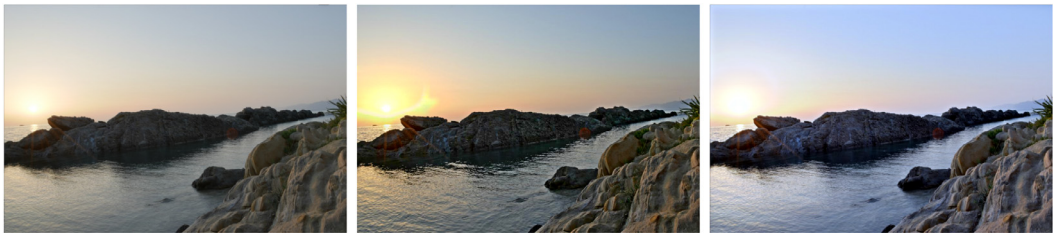
The visual comparison of the fused *Tower* stack is given in Fig. 11. Although MEF-SSIM scores of Mertens, Ma and Li13 are slightly better than the proposed method, lack-of-contrast regions and detail-loss in several areas are present in the outputs of these algorithms. Since the Eiffel Tower is low contrast, highlights are almost non-existent in the dark regions in Mertens and Li13. Also, there are lost details mainly on the upper side of the tower in these methods. The proposed algorithm, in contrast, shows better visual quality, with brighter details of the tower and more natural-looking clouds.

For *Lighthouse* in Fig. 12, the specific features and fine details in low-light regions, i.e., the rocks and house, are more visible in the proposed method compared to Mertens. On the other hand, Mertens and Li12 preserve more details on the sea and sky whereas some saturation problems exist in these regions in the proposed output. As aforementioned, this undesired outcome occurs due to the extremely over exposed image in the input sequence. The *Kluki* stack has a similar problem; however, clouds still appear visually plausible in the proposed algorithm.

Fig. 13 further compares the fused *Cave* stack with four different exposures. Li13 presents the best MEF-SSIM score for this image and Mertens has a similar statistical score to the proposed algorithm. The left-side of the cave entrance is artificially dark in Li13,



(a) Three input exposures in the stack.



(b) Mertens (c) Ma (d) Proposed

**Fig. 23.** Visual comparison of different methods for *SecretBeach*.

(a) Three input exposures in the stack.

(b) Mertens (c) Ma (d) Proposed

**Fig. 24.** Visual comparison of different methods for *OldHouse*.

but more natural in the proposed result. In addition, in Mertens, the fused image has less contrast, especially at the top of the cave and at the rock on the right.

A visual comparison of the fused outputs for the *Balloons* stack containing nine exposures and also the effect of contrast enhancement-based post-processing can be observed in Fig. 14. Additional visual comparisons are also provided in Figs. 15–20 for *Madison Capitol, Landscape, Farmhouse, Belgium House, Kluki* and *Lamp* image stacks, respectively.

All above experiments were carried out on an AMD Ryzen(TM) 5 3600x CPU @ 3.80 GHz 6-core 16GB RAM machine using Matlab R2020a and 13 test sequences detailed in Table 1 are fused in 1.81 s on the average, ranging between 1.5 s to 2.29 s.

In the second set of experiments, the proposed MEF algorithm is tested with higher resolution images in order to increase the variety of the analysis. The dataset in this setup includes five different static image stacks, namely *Flowers, SeaRock, SecretBeach, OldHouse, Rovinia*, of sizes 720 $\times$ 1080 pixels with three exposures each [35]. While the best MEF-SSIM scores in Table 3 are reached via the proposed approach (0.969 on average), the average execution times of the best three algorithms are 0.50 s, 7.85 s and 6.37 s for Mertens, Ma and the proposed method, respectively. The MEF-SSIM scores for Mertens, Raman and S.Li are aligned with [35] and the results provided for Ma are obtained through [31].

**Table 3**
MEF-SSIM scores for each stack adopted from Merianos and Mitianoudis [35].

| | Algorithms | | | | |
|---|---|---|---|---|---|
| | Mertens | Raman | S.Li | Ma | Proposed |
| *Flowers* | 0.964 | 0.906 | 0.921 | 0.987 | **0.989** |
| *SeaRock* | 0.932 | 0.896 | 0.913 | 0.933 | **0.958** |
| *SecretBeach* | 0.951 | 0.927 | 0.888 | 0.899 | **0.963** |
| *OldHouse* | 0.974 | 0.959 | 0.907 | 0.987 | **0.991** |
| *Rovinia* | 0.934 | 0.881 | 0.913 | 0.935 | **0.942** |
| **Average** | 0.951 | 0.914 | 0.908 | 0.948 | **0.969** |

(a) Three input exposures in the stack.

(b) Mertens (c) Ma (d) Proposed

**Fig. 25.** Visual comparison of different methods for *Rovinia*.

The visual comparisons are also provided in Figs. 21–25 for *Flowers, SeaRock, SecretBeach, OldHouse, Rovinia*, respectively. It can be observed from these illustrations that the details and colors are better preserved in the proposed technique, which results in more natural-looking outputs. Furthermore, the obtained images are more vivid and visually appealing. This mainly results from the successful usage of linear embeddings and watershed masking to fuse an image stack, which has significant potential.

## 4. Conclusion

HDR-like image reconstruction through MEF is a common study field in image processing and computer vision, and weight map extraction typically presents the novel part of different algorithms. This paper contains a proposal for a new MEF approach based on linear embeddings of pixel/patch spaces of images and watershed masking. In the developed method, linear embedding weights are extracted from differently exposed images and the corresponding watershed masks are used to adjust these maps according to the informative parts of input images for the final fusion step. After a fused image is acquired, the low-light areas and unsatisfactory color intensities are corrected via a simple local brightness enhancement algorithm. As a result, statistically successful and natural-looking output images are obtained. To the best of our knowledge, the proposed framework in this study is novel in that it exploits linear embeddings of image spaces and watershed masks of images.

The main drawback of the proposed method is that the visual quality and statistical scores tend to decrease slightly when the input stack contains excessively over-exposed or excessively dark under-exposed images. This problem can simply be solved by manually discarding extreme exposures; alternatively automated techniques can be envisaged by means of, for example, outlier detection techniques. Since this process was not within the scope of this study, such extension is designated as future work. Furthermore, the proposed algorithm can be adapted for dynamic image stacks in the future. As a final remark, the main computational complexity lies on the three optimization problems given in Eq. (1) for linear embeddings. This run-time complexity can be greatly reduced by parallel implementations on GPU processors.

## Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## CRediT authorship contribution statement

**Oguzhan Ulucan:** Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Writing - original draft, Visualization. **Diclehan Karakaya:** Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Writing - original draft, Visualization. **Mehmet Turkan:** Conceptualization, Methodology, Formal analysis, Writing - review & editing, Supervision.

## Acknowledgement

The authors are thankful to Dr. Hui Li and Prof. Dr. Lei Zhang for sharing their code, and to Burak Ciceksoy for his help in creating illustrations of images.

## References

[1] K. Ma, Z. Wang, Multi-exposure image fusion: a patch-wise approach, in: IEEE Int. Conf. Image Process., 2015, pp. 1717–1721.
[2] T. Mertens, J. Kautz, F. Van Reeth, Exposure fusion: a simple and practical alternative to high dynamic range photography, Comp. Graph. Forum 28 (2009) 161–171.
[3] B. Gu, W. Li, J. Wong, M. Zhu, M. Wang, Gradient field multi-exposure images fusion for high dynamic range image visualization, J. Vis. Commun. Image Represent. 23 (4) (2012) 604–610.
[4] Z.G. Li, J.H. Zheng, S. Rahardja, Detail-enhanced exposure fusion, IEEE Trans. Image Process. 21 (11) (2012) 4672–4676.
[5] S. Li, X. Kang, Fast multi-exposure image fusion with median filter and recursive filter, IEEE Trans. Consum. Electron. 58 (2) (2012) 626–632.
[6] S. Li, X. Kang, J. Hu, Image fusion with guided filtering, IEEE Trans. Image Process. 22 (7) (2013) 2864–2875.
[7] H. Li, L. Zhang, Multi-exposure fusion with CNN features, in: IEEE Int. Conf. Image Process., 2018, pp. 1723–1727.
[8] S. Paul, I.S. Sevcenco, P. Agathoklis, Multi-exposure and multi-focus image fusion in gradient domain, J. Circuit Syst. Comp. 25 (10) (2016) 1650123.
[9] K. Ma, H. Li, H. Yong, Z. Wang, D. Meng, L. Zhang, Robust multi-exposure image fusion: a structural patch decomposition approach, IEEE Trans. Image Process. 26 (5) (2017) 2519–2532.
[10] S.-h. Lee, J.S. Park, N.I. Cho, A multi-exposure image fusion based on the adaptive weights reflecting the relative pixel intensity and global gradient, in: IEEE Int. Conf. Image Process., 2018, pp. 1737–1741.

[11] N. Hayat, M. Imran, Ghost-free multi exposure image fusion technique using dense sift descriptor and guided filter, J. Vis. Commun. Image Represent. 62 (2019) 295–308.
[12] C. Florea, F. Albu, C. Vertan, A. Drimbarean, Logarithmic tools for in-camera image processing, in: Irish Signal Syst. Conf., 2008, pp. 394–399.
[13] S. Raman, S. Chaudhuri, Bilateral filter based compositing for variable exposure photography, in: Eurograph., 2009, pp. 1–4.
[14] Y. Que, H.J. Lee, Densely connected convolutional networks for multi-exposure fusion, in: Int. Conf. Comput. Sci. Comput. Intell., 2018, pp. 417–420.
[15] Q. Liu, H. Leung, Variable augmented neural network for decolorization and multi-exposure fusion, Inf. Fusion 46 (2019) 114–127.
[16] Y. Hu, R. Zhen, H. Sheikh, CNN-based deghosting in high dynamic range imaging, in: IEEE Int. Conf. Image Process., 2019, pp. 4360–4364.
[17] R. Li, S. Liu, G. Liu, T. Sun, J. Guo, Multi-exposure photomontage with hand-held cameras, Compt. Vis. Image Underst. 193 (2020) 102929.
[18] S.T. Roweis, L.K. Saul, Nonlinear dimensionality reduction by locally linear embedding, Science 290 (5500) (2000) 2323–2326.
[19] S. Beucher, C. Lantuejoul, Use of watersheds in contour detection, in: Proc. Int. W. Image Process., Rennes, 1979. pp. –.
[20] K. Ma, K. Zeng, Z. Wang, Perceptual quality assessment for multi-exposure image fusion, IEEE Trans. Image Process. 24 (11) (2015) 3345–3356.
[21] M. Turkan, D. Thoreau, P. Guillotel, Iterated neighbor-embeddings for image super-resolution, in: IEEE Int. Conf. Image Process., 2014, pp. 3887–3891.
[22] M. Turkan, D. Thoreau, P. Guillotel, Optimized neighbor embeddings for single-image super-resolution, in: IEEE Int. Conf. Image Process., 2013, pp. 645–649.
[23] M. Turkan, D. Thoreau, P. Guillotel, Self-content super-resolution for ultra-HD up-sampling, in: Proc. European Conf. Vis. Media Prod., 2012, pp. 49–58.
[24] H. Chang, D.-Y. Yeung, Y. Xiong, Super-resolution through neighbor embedding, in: IEEE Conf. Comp. Vis. Patt. Recog., vol. 1, 2004. pp. I–I.
[25] D.L. Donoho, C. Grimes, Hessian eigenmaps: locally linear embedding techniques for high-dimensional data, Proc. Natl. Acad. Sci. 100 (10) (2003) 5591–5596.
[26] J.B. Tenenbaum, V. De Silva, J.C. Langford, A global geometric framework for nonlinear dimensionality reduction, Science 290 (5500) (2000) 2319–2323.
[27] L. Vincent, Morphological grayscale reconstruction in image analysis: applications and efficient algorithms, IEEE Trans. Image Process. 2 (2) (1993) 176–201.
[28] Marker-controlled watershed segmentation, 2020, https://www.mathworks.com/help/images/marker-controlled-watershed-segmentation.html#responsive_offcanvas accessed June 17.
[29] Z. Wang, A.C. Bovik, H.R. Sheikh, E.P. Simoncelli, Image quality assessment: from error visibility to structural similarity, IEEE Trans. Image Process. 13 (4) (2004) 600–612.
[30] K. Zeng, K. Ma, R. Hassen, Z. Wang, Perceptual evaluation of multi-exposure image fusion algorithms, in: IEEE Int. W. Quality Multimedia Exp., 2014, pp. 7–12.
[31] K. Ma, Robust multi-exposure image fusion: a structural patch decomposition approach, 2020. https://ece.uwaterloo.ca/~k29ma/codes/SPD-MEF.rar accessed April 03.
[32] S. Paul, Multi-exposure and multi-focus image fusion in gradient domain, 2020. https://github.com/sujoyp/gradient-domain-imagefusion accessed June 10.
[33] H. Li, Multi-exposure fusion with CNN features, 2020https://github.com/xiaohuiben/MEF-CNN-feature accessed June 10.
[34] H. Leung, Variable augmented neural network for decolorization and multi-exposure fusion, 2020, https://github.com/yqx7150/DecolorNet_FusionNet_code accessed June 10.
[35] I. Merianos, N. Mitianoudis, Multiple-exposure image fusion for HDR image synthesis using learned analysis transformations, J. Imaging 5 (3) (2019) 32.