

Information-Theoretic Channel for Multi-exposure Image Fusion

QIAOHONG HAO¹, QI ZHAO¹, MATEU SBERT^{2,*}, QINGHE FENG³,
COSMIN ANCUT⁴, MIQUEL FEIXAS², MARIUS VILA² AND JIAWAN ZHANG¹

¹*College of Intelligence and Computing, Tianjin University, Yaguan Road 135, Tianjin 300350, China*

²*Institute of Informatics and Applications, University of Girona, 17003 Girona, Spain*

³*College of Intelligent Engineering, Henan Institute of Technology, Pingyuan Road 699,
Xinxiang 453003, China*

⁴*ETcTI, Universitatea Politehnica Timisoara, Vasile Parvan 2, 300023, Timisoara, Romania*

*Corresponding author: mateu@ima.udg.edu, jwzhang@tju.edu.cn

Multi-exposure image fusion has emerged as an increasingly important and interesting research topic in information fusion. It aims at producing an image with high quality by fusing a set of differently exposed images. In this article, we present a pixel-level method for multi-exposure image fusion based on an information-theoretic approach. In our scheme, an information channel between two source images is used to compute the Rényi entropy associated with each pixel in one image with respect to the other image and hence to produce the weight maps for the source images. Since direct weight-averaging of the source images introduce displeasing artifacts, we employ Laplacian multi-scale fusion. Based on this pyramid scheme, images at every scale are fused by weight maps, and a final fused image is inversely reconstructed. Multi-exposure image fusion with the proposed method is easy to construct and implement and can deliver, in less than a second for a set of three input images of size 512×340 , competitive and compelling results versus state-of-art methods through visual comparison and objective evaluation.

Keywords: multi-exposure image fusion; information channel; conditional entropy; Rényi entropy; Laplacian pyramid

Received 3 November 2020; Revised 17 August 2021; Editorial Decision 6 September 2021

Handling editor: Dimitrios Tzovaras

1. INTRODUCTION

Image fusion is a process in image synthesis where various source images obtained from multiple sources of the same object/scene are combined to produce synthetic and visually improved images for human visual perception [1]. With the continuous improvement of the camera functions of mobile devices, such as mobile phones, photography is becoming more and more convenient and fast, and the content and performance of images are constantly enriched, making people invariably pursue technological perfection. Given the poor lighting condition and limited dynamic range of digital imaging devices, the range of colors that the acquired images can represent is limited. For a normal low dynamic range (LDR) digital camera, details in brighter and darker regions

can not be fully captured in one exposure, and rich details can not be displayed completely. Consequently, in order to obtain better quality images with appropriate contrast, vivid color and rich details, which characterize high dynamic range (HDR) images, a variety of imaging and enhancement techniques based on a single image have been proposed [2–6]. Nonetheless, these techniques can not restore saturated pixel in LDR images, and thus researchers are paying more and more attention to multi-image-based imaging techniques [7–17] to overcome this constraint that single image-based techniques have.

In a HDR image synthesis technique [18], the real exposure can be obtained by the exposure sequences, and then tone mapping should be performed on the HDR image to

adapt to the LDR of the display device [19–26]. Hardware solutions for solving HDR reconstruction have been proposed such as in [27], implementing a HDR imaging system with a highly programmable camera unit with high throughput A/D conversion, in [28], proposing a weighting function that produces statistically optimal estimates under the assumption of compound Gaussian noise, and in [29], presenting a unifying approach performing HDR assembly directly from raw sensor data. Also, several deep learning methods have emerged as in [30], to address the artifacts for dynamic scenes in HDR merge process, [31], for image reconstruction from a single exposure, and [32], to model the entire HDR video reconstruction process.

Another important technique is multi-exposure image fusion (MEF), which combines a sequence of LDR images to create an enhanced image to be displayed in LDR digital cameras. According to [33], 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.

Different kinds of fusion methods have been presented, such as weighted-average methods in space domain, principle component analysis, multi-scale fusion methods, which are introduced in detail in [34]. Generally, the fusion algorithms are split into three categories: pixel level, feature level, and decision level, according to the way of performing [35]. For the fusion algorithms, the key challenge is to select the information measure or quality measure index for the source image sequence that directly affects the detail extraction. However, there is much room for improvement in the preservation of the details of existing fusion algorithms.

Inspired by this, we propose a measure based on an information-theoretic channel to retain more details in the fused result. The information channel has been successfully used in medical imaging and in other areas of visual computing [36, 37]. Due to the relationship between the different exposure images, we do not measure images separately. We establish an information channel between two images, and use Rényi entropy to evaluate one image with the other one. Owing to the fact that we perform blending by weight map at pixel level, noise is introduced. To get around this problem, we perform fusion at multiple scale, and then reconstruct fused images at each scale into a final fused image. The proposed method can capture more details in the fusion results.

In this paper, we expand our preliminary work in [38] in several lines:

- We build information channels for every pair of images, instead of for only consecutive images ordered by exposure time in [38].

- We use Rényi entropy, an extension to Shannon entropy used in [38], to compute the information measure of a pixel in one image with respect to the other image.
- We show the computation complexity by dividing the fusion process into four steps including pre-treatment, channel building, weight map computing, and fusion by Laplacian pyramid frame.
- We weight the entropies obtained from the information channel of two images with the gaussian difference of the two images to allow to compensate for the increasing differences in exposure in the series of input images.
- We compare our Rényi results with six previous MEF methods with MEF-structural similarity index measure (MEF-SSIM), mutual information, entropy and entropy rate measures, against only comparison with entropy rate measure in [38].

The rest of this paper is organized as follows. Related work is given in Section 2. In Section 3.1, an overview of the proposed method is presented. Experimental results, parameter determination, comparison to the current state-of-the-art methods and computation complexity are introduced in Section 4 and finally the paper is concluded and prospected in Section 5.

2. RELATED WORK

In the multi-scale fusion framework, there are different ways to measure information. Burt and Kolczynski [39] first applied Laplacian pyramid in image fusion, where fusion at each level is guided by selecting maximal local energy. Instead of using the same fusion principle, Mertens *et al.* [9] adopted the quality measures of contrast, saturation, and well-exposure to guide the image fusion at each scale, obtaining more appealing results. Based on this method, Li *et al.* [40] developed a detail-enhanced method by employing a detail extraction component as a quadratic optimization problem so that more details can be preserved, and then further improved it in [41]. Ancuti *et al.* [42] developed a single scale fusion method, and used saliency as quality measure besides those utilized in [9].

To avoid losing detail, different filters to preserve details were applied in [11], [43], [44]. Li and Kang [11] produced weight maps by local contrast, brightness and color dissimilarity and then refined them by recursive filtering. Raman and Chaudhuri [43] used bilateral filters to preserve details. Then Li *et al.* [44] proposed a guided filtering-based weighted average method that highlights pixel saliency and spatial consistency with two-scale decomposition. Weighted guided image filtering was utilized by [17] to avoid amplifying the noise. Recently, Patel *et al.* [45] proposed a propagated image filter to generate weight maps of input images and used gradient domain postprocessing to improve the result. Ma and Wang

[46] introduced a new method to avoid disturbing noises, by decomposing an image patch into three components: signal strength, signal structure and mean intensity, measuring them separately based on patch strength and exposition measures, and then reconstructing the patches for the fused image.

Apart from the methods mentioned, Zhang *et al.* [47] combined source images as a weighted sum to form a fused image, and the weight was defined by gradient-directed quality assessment. Shen *et al.* [48] introduced a method for fused image based on perceptual quality measurement. Yu *et al.* [49] employed spatial distribution of intensity to fuse differently exposed images. Dong *et al.* [50] implemented a method based on registration to fuse images. Qin *et al.* [51] presented the fusion of images for dynamic scenes using optimization. Wang *et al.* [52], Shen *et al.* [53] and Wang *et al.* [54] proposed a segmentation method for image fusion. Probabilistic model-based techniques have also been used to construct a fused image [55, 56]. These probabilistic model-based techniques are more complex as compared to other methods. Merianos *et al.* [57] combined two image fusion methods with luminance channel and color channel to tackle multiple-exposure fusion. Ulucan *et al.* [33] developed a method based on weight map extraction via linear embeddings and watershed masking for static image exposure fusion.

In addition, information theory has been used in image fusion in recent years. Goshtasby [8] used entropy to measure the information of each uniform piece of the image and then blended these input images by rational Gaussian functions, where the optimal block size and filter width are obtained by gradient-ascent algorithm, yet as [9] commented, halo artifacts around edges can not be well avoided. Herwig *et al.* [58] utilized local entropy with adaptive filter size to measure pixel-wise information. Bramon *et al.* [36] applied mutual information measures in the context of multi-modal medical 3D volumetric image fusion, using the more informative voxel intensity from the two source images as the voxel intensity of the fused image.

Our previous paper [38] investigated MEF based on mutual information measures I_1 and I_2 , introduced for fusion by Bramon *et al.* [36], and then compared them with the Shannon conditional entropy (CE) to measure contributions of pairwise pixels for MEF. The CE measure proved better than I_1 and I_2 . Upon this basis, we explore here the extension of Shannon entropy to Rényi entropy to improve the fused result, which allows for a trade-off between the details preservation and esthetic appeal and improve the weighting of the information maps by taking into account the gaussian differences between the source images.

3. INFORMATION CHANNEL FOR MEF

3.1. Overview of the proposed method

An overview of the processing steps of our approach is shown in Fig. 1. These steps are explained in detail in the follow-

ing sections. The information-theoretic channel (as shown in Fig. 2) is explained in Section 3.2. The information measures (conditional Shannon entropy and Rényi entropy) are introduced in detail in Section 3.3. Finally, we employ multi-scale fusion in Laplacian pyramid frame (Fig. 4, Section 3.4) to get pleasing fused image.

3.2. Information channel

Like in our prior work [38], two images are interpreted as random variables and an information channel is built between them to calculate pixel-wise weights. The two images, which have equal size and thus contain the same number of pixels, are represented by their normalized luminance histograms, which are denoted by X and Y , and the corresponding bins of the histogram are expressed as x and y (in this paper we compute luminances using the formula $0.11 \times B + 0.59 \times G + 0.3 \times R$, where R , G and B are the red, green and blue color components [59]). As shown in Fig. 2, the channel from X to Y , $X \rightarrow Y$ is given by a matrix M of conditional probabilities $p(Y|X)$, together with the marginal probabilities, $p(X)$ and $p(Y)$. If there are n bins in X , m bins in Y , the matrix is $n \times m$ dimensional, where $M(y, x)$ is $p(y|x)$ (see the diagram in Fig. 2b). Similarly, the channel from Y to X is expressed by a $m \times n$ dimensional matrix of $p(X|Y)$.

In Fig. 2b, $p(x)$ and $p(y)$ are the probabilities corresponding to bins x and y , and $p(y|x)$ is the conditional probability of y given x . $N(x)$ and $N(y)$ are the number of pixels whose pixel value falls within bin x and y , respectively, and $N(x, y)$ is the number of corresponding pixels on both images (that is, pixels at the same position in both images), which color in the first image falls in bin x and in the second image in bin y . The probabilities are computed as relative frequencies:

$$p(x) = \frac{N(x)}{\sum_{x \in X} N(x)}$$

$$p(y|x) = \frac{N(x, y)}{N(x)}. \quad (1)$$

3.3. Information measures

3.3. Conditional entropy

The mutual information between variables X and Y expresses their shared information. The ‘CE’ represents the uncertainty of one variable conditioned to the other variable. Deweese and Meister [60] proposed the concept of specific information of one element, and different decompositions of mutual information I_1, I_2 were proposed, adopted later by Bramon *et al.* [36]. In Zhao *et al.* [38] it was empirically found that the weights based on CE performed much better than weights based on mutual information measures I_1 and I_2 .

The reason why entropy can work better than mutual information might be found in that measures based on mutual information are directly related to the output distribution, as I_1

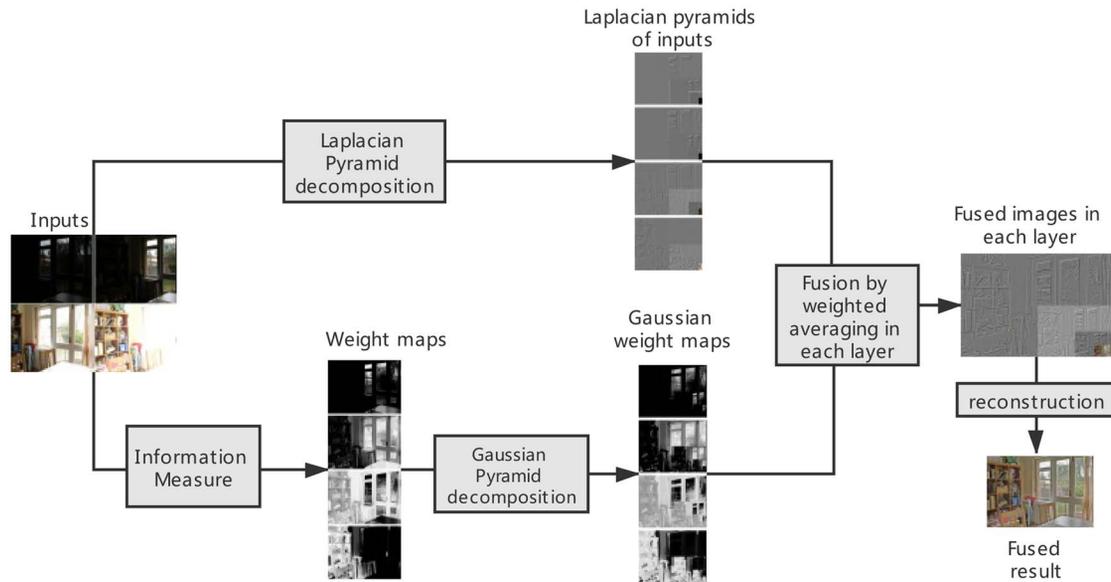


FIGURE 1. Overview of the different processing steps of our MEF approach.

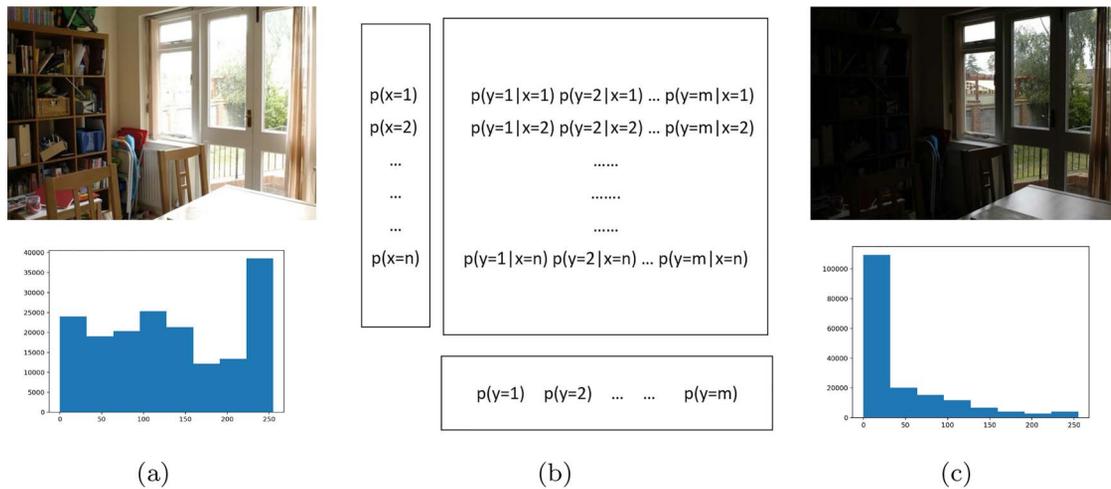


FIGURE 2. The information channel $X \rightarrow Y$ from X to Y .

can be expressed as the relative entropy or Kullback–Leibler distance of $p(Y|x)$ to $p(Y)$, or directly related to the input distribution $p(X)$, as I_2 can be expressed as $H(X) - H(Y|x)$. The input and output distributions for the channels between the different exposure images are the frequency histograms, which can be very different due to different exposure, see left and right image histogram in Fig. 2. This means that mutual information-based measures can be good to decide which pixel, in either of both images, holds more information, but not the relative size of this information. This is, to choose the more informative

pixel from the two images in a channel, as Bramon *et al.* [36] did, mutual information can be very good, but to use mutual information to weight average between both pixels would make the result distorted. In contrast, entropy, which stands alone as a measure, can be considered as the complementary of the distance to the uniform distribution, which is the same for all the images, and thus puts all entropies considered on the same ground.

The measure CE used in [38] comes from the CE of bin x , $H(Y|x)$, which gives the new distribution of Y , or remaining

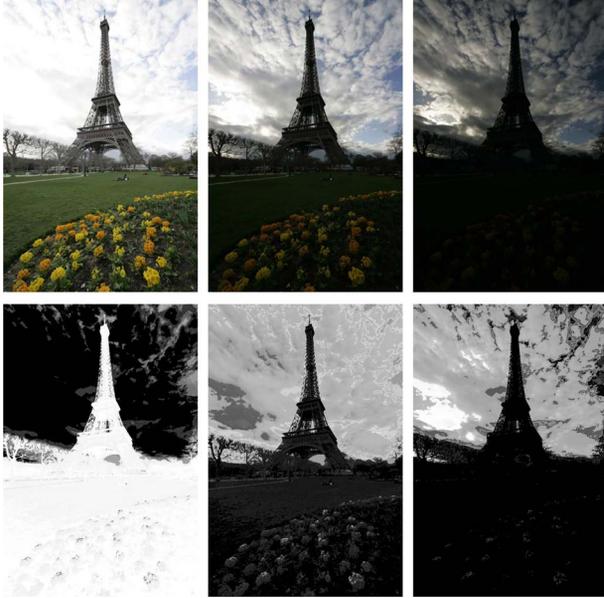


FIGURE 3. The first row of images are source images, and the second row of images are the corresponding weight maps.

uncertainty, once known $X = x$, and is given by Equation 2.

$$H(Y|x) = - \sum_{y \in Y} p(y|x) \log p(y|x). \quad (2)$$

Observe that Equation 2 is the Shannon entropy of the conditional distribution $p(Y|x)$. Next, the normalized weight map can be produced by Equation 3, where W_x and W_y represent weights of bins x and y , respectively. We assign to a pixel the weight corresponding to its color bin, this is, all the pixels in the same bin have the same weight. The pixel with higher CE has higher uncertainty, and thus has less correlation to the other image. Hence, weights are inversely proportional to CE:

$$\begin{aligned} W_x &= \frac{H(X|y)}{H(Y|x) + H(X|y)}, \\ W_y &= \frac{H(Y|x)}{H(Y|x) + H(X|y)}. \end{aligned} \quad (3)$$

For multiple source images (e.g. X , Y and Z), we consider the information of X , Y over Z (e.g. $H(X, Y|z)$), but its computation would be inefficient and inaccurate, as it is based on a very sparse three-dimensional matrix. Therefore, we compute weight maps in pairwise images and normalize them to the same scale.

3.3.1. Overexposed and underexposed regions

Observe that the color of pixels from overexposed and underexposed regions will be selected in the fusion process from the non-over or non-underexposed images. This can be seen

in the following way. For an overexposed image X , where the luminance of the overexposed region will have a maximum, constant value, say $l_{max} = l_n$, the corresponding region in the not overexposed image Y will have a scale of luminances, say to simplify l_{n-1}, l_{n-2} . This means that the entropy $H(X|l_n)$, as the conditional probabilities $p(X|l_{max})$ are distributed over l_{n-1}, l_{n-2} , is higher than $H(X|l_{n-1})$ and $H(X|l_{n-2})$, as very probably most of the pixels with l_{n-1}, l_{n-2} value in Y image are paired with l_n value in X image, and thus the pixel in the non-overexposed image Y will be given a higher weight. Observe that the same can be told of an underexposed image. An exception would be when all pixels with constant maximum luminance l_n in the overexposed image would be substituted by another constant luminance, say l_{n-1} , in the non-overexposed image, as in that case the weight of the pixel in both images would be the same. We have not observed this situation in our experiments, but if there would be the case, the effect would be damped as we build information channels for all pairs of source images.

Next, we present the Rényi extension of Shannon entropy.

3.3.1. Rényi entropy

Based on Shannon entropy, Alfred Rényi [61] proposed a generalized entropy, defined by

$$R_\alpha = \frac{1}{1 - \alpha} \log \sum_{x \in X} p(x)^\alpha, \quad (4)$$

where the parameter α is greater than or equal to 0. When α is equal to 1, Rényi entropy reverts to Shannon entropy.

The value of α modulates the relative importance of the probability distribution values $p(x)$. For $\alpha > 1$ the bigger $p(x)$ values, which correspond to bigger regions (remember that $p(x)$ is the relative number of pixels with luminance x) become more important and smaller regions become less important, and thus the distribution becomes less entropic and more detail aware. For $\alpha < 1$ the smaller $p(x)$ values and thus the smaller regions increase its importance, and the distribution becomes more entropic.

Following Eq. 4, the conditional Rényi entropy of Y over x is defined by

$$R_\alpha(Y|x) = \frac{1}{1 - \alpha} \log \sum_{y \in Y} p(y|x)^\alpha. \quad (5)$$

$R_\alpha(Y|x)$ indicates the information of luminance value x relative to histogram Y (or by abuse of language, image Y) and can be calculated based on $X \rightarrow Y$. Similarly, to get the information of luminance value y relative to histogram or image X , we need to compute $R_\alpha(X|y)$ based on $Y \rightarrow X$.

For two images X and Y , we build the information channels $X \rightarrow Y$ and $Y \rightarrow X$, based on which, the information of x on Y and of y on X will be measured by Rényi entropy. Similar to

Equation (3), the normalized weight of luminance value x is

$$W_x = \frac{R_\alpha(X|y)}{R_\alpha(Y|x) + R_\alpha(X|y)}, \quad (6)$$

and bin y is

$$W_y = \frac{R_\alpha(Y|x)}{R_\alpha(Y|x) + R_\alpha(X|y)}. \quad (7)$$

We call $R_\alpha(X|Y)$ and $R_\alpha(Y|X)$ the vectors of values for all y and x respectively. To compute the weights of two corresponding pixels (this is, in the same position) with quantized luminance values x, y , we look up in $R_\alpha(X|Y)$ using y and in $R_\alpha(Y|X)$ using x , and then apply Equations 6 and 7. For fusion of more than two images, we considered building an information channel between each pair of images, and adding up the information obtained from each channel to compute the information of each source image. However, the contribution of each channel should be different, as the greater the difference of the images in the channel, the less should be the association and contribution of the information. Thus, we give lesser weights to those channels in which images are more dissimilar in the following way.

3.3.2. Gaussian difference

Given a sequence of source images, S_1, S_2, \dots, S_n , for each source image S_k the channels with all the other images are constructed, i.e. $S_k \rightarrow S_t$, based on which we compute the corresponding $R_\alpha(S_k|S_t)$, where $t = 1 \dots n, t \neq k$. Then the weight of $R_\alpha(S_k|S_t)$ is defined by

$$w_{k,t}(i,j) = \frac{G_{k,t}(i,j)}{\sum_{l=1, l \neq k}^n G_{k,l}(i,j)}, \quad (8)$$

where (i,j) is the pixel with coordinates i and j and where $G_{k,t}$ is the Gaussian function of image difference:

$$G_{k,t} = \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{(S_k - S_t)^2}{2\sigma^2}}. \quad (9)$$

Note that as this difference is computed on all pixel luminance values, which are scaled to $[0,1]$, then $G_{k,t}$ has the same scale as the source image. Equation (8) illustrates the normalized weight of channel between the k th and t th images.

3.3.3. Information maps

We then weight average the information obtained from each channel to obtain $I'_k(i,j)$, the information of image S_k

$$I_k(i,j) = \left[\sum_{t=1, t \neq k}^n w_{k,t}(i,j) \cdot R_\alpha(S_k|S_t)(i,j) \right]^\beta, \quad (10)$$

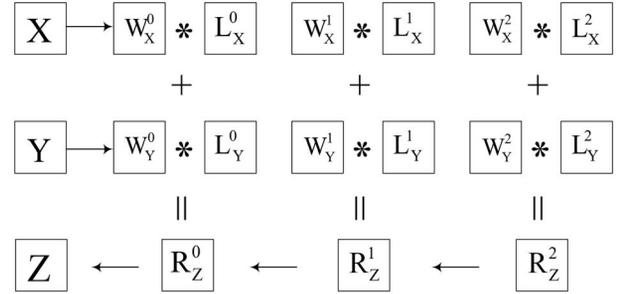


FIGURE 4. An example of fusion in the frame of Laplacian pyramid (3 layers). X, Y represent two source images, Z is the fused image. $L_X^k(L_Y^k)$ is the k th layer of source images, $W_X^k(W_Y^k)$ is the corresponding filtered weight map, and R_Z^k is the k th layer fused image, where $k \in \{0, 1, 2\}$.

where the $R_\alpha(S_k|S_t)(i,j)$ value is obtained by looking up the value in $R_\alpha(S_k|S_t)$ using the quantized luminance value of pixel (i,j) in image S_t , and β a power exponent to fine-tune the result. Normalizing the $I_k(i,j)$ values we finally obtain the information map $W_k(i,j)$ for image S_k

$$W_k(i,j) = \frac{I_k(i,j)}{\sum_{t=1, t \neq k}^n I_t(i,j)}. \quad (11)$$

Figure 3 shows a group of source images with decreasing exposure (top) and their respective weight maps (bottom) obtained by the proposed information measure. The brighter the weight maps, the higher the weights. Comparing the source images and weight maps, we can see that the first weight map has more information in the tower and grassland area, but less in the sky area, whereas the second and the third both show more details of sky, which is in line with our visual perception. In the third source image, we observe that there is a magnified area, which is the most informative among all images and indicated by the weight map clearly.

3.4. Laplacian fusion

The next step is to blend the source images according to their weights. However, the fused image produced by direct average weighting is usually unacceptable due to disturbing seams, which results from large difference and sharp transitions in the weight maps. Smoothing weight maps by Gaussian or bilateral filters is insufficient [9]. Here, we carry out ‘Laplacian pyramid’, which has sufficiently good performance in image fusion as remarked by [9, 42]. Source images are decomposed into multiple scale images, then blending by weight map is performed at each level, and finally all levels are reversely reconstructed into a fused image. To adapt the size of the image at each level, the weight maps obtained from the original source images are downsampled with Gaussian filter. The multiscale fusion diagram based on Laplacian pyramid is given in Fig. 4.

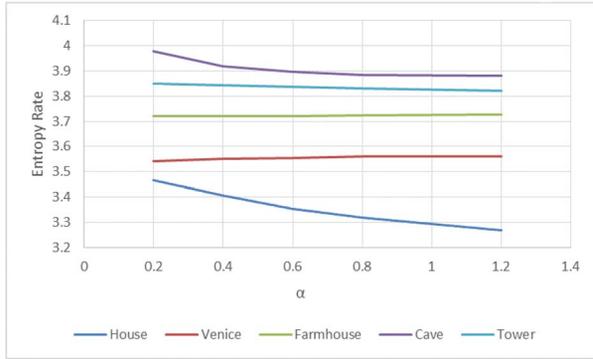


FIGURE 5. The entropy rates of fused images based on different values of α .

4. EXPERIMENTAL RESULTS

4.1. Parameter determination

In this section, we will discuss the determination of parameters in the proposed method used. Firstly, image probability distributions are given via the normalized histograms obtained from gray values. Note that we use less number of bins than the original value of 256. This is because the more the bins, the more sparse the channel matrix, and thus the more inaccurate the result. Also, the cost of the channel construction is quadratic in the number of bins. Zhao *et al.* [38] considered eight bins to be a good balance between the detail preservation and calculation performance. Although more details could be captured if more bins are used, there was no significant improvement in the results. Thus we follow Zhao *et al.* [38] and set 8 as number of bins. The information channels between two source images can be constructed once assigned bin number by Equation 1. Then based on the specific criteria, the weight maps of the two source images can be generated.

For the proposed Rényi entropy measure, we fix the values of the parameters after repeated experimentation. Using five pyramid layers gets attractive results. For the standard deviation σ of the Gaussian function to compute information channel weights, we set $\sigma = 0.5$. About power exponent β in Equation 10, which needs to be decided to fine-tune the result, we set its value as 2, which makes the information between different source images more distinguishable. Additionally, an important parameter is α to compute the Rényi entropy-based information measure. To analyze the appropriate values for α , we compute entropy rates of fused images based on different α , the curves are shown in Fig. 5.

In Fig. 5, we can see that there is more fluctuation in the curves of scenes House and Cave, Fig. 6, therefore we show these fused results based on different values of α in Fig. 7. From Fig. 7, we can observe that there are more details when α equals 0.2 for the two scenes. In the curves, these two fused results with 0.2 have higher entropy rate. We can also see in Table 1 how the value of $\alpha = 0.2$ gives the bigger MEF-SSIM [62]

TABLE 1. MEF-SSIM scores of fused results with different α corresponding to Fig. 7.

α	House	Cave
0.2	0.9029	0.9519
0.4	0.8846	0.9414
0.6	0.8853	0.9409
0.8	0.8852	0.9408
1.0	0.8895	0.9412
1.2	0.8830	0.9413

score for all the images considered. Thence, we selected α as 0.2 to generate the weight maps.

4.2. Comparison with other methods

4.2. Visual Effect

In the following, we compare the fused results of our previous method and our new method (Fig. 8e and g) from different scenes with other five methods, this is shown in Fig. 8. The source image sequences and the results by the comparison methods, except [46], [38] and [57] ones, are obtained from the MEF database created by Ma *et al.* [62, 63]. The source images in the MEF database are all aligned. For unaligned source images a preprocess to register them should be done first.

Ma *et al.* [46] and Merianos *et al.* [57] results are obtained with the software by the authors [64] and [65], respectively, the last one without its postprocessing option. The software by Merianos *et al.* [65] merges a black frame on the results, visible in the images, and allows only three input source images. Figure 9 gives the close-ups corresponding to the fused results in Fig. 8.

In the first row of Fig. 8, the House scene, we can observe that the left and right parts are inconsistent in color in the fused image of Mertens [9], as can be perceived in the color of the two chairs. The results by S.T.Li [11] and Li [44] have the same problems as Mertens ones. The close-up of the window is shown in the first row of Fig. 9. In it, we can see that the result of Ma [46] yields the region outside the window a little overexposed, whereas the results based on entropy (see Fig. 9e and g) show more details outside the window.

In the second row, the Venice scene, it can be seen in Li's result [44] that the color of sky is deep and inconsistent with the color of water reflection, something that does not happen in our result. Moreover, it is perceivable that our fused result has more details in the boat and water surface. On top of that, comparing the pole regions (lower right corner area) in the results by Mertens [9], Ma [46] and Merianos [57], our result is more detail-preserving and brighter. Also comparing the round roof region in the results by S.T.Li [11] and Li [44], our fused

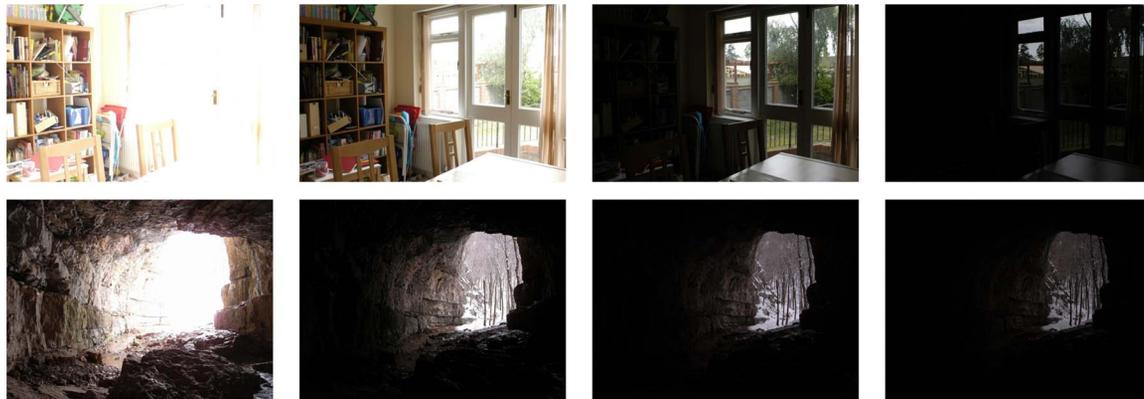


FIGURE 6. Source images for the house and cave scenes.



FIGURE 7. Fused results based on Rényi entropy with different α values for the House and Cave scenes.

images are better because of color consistency (it can also be observed from the close-up in the second row of Fig. 9).

For the third row of results, the Cave scene, our fused image contains more details inside the cave, and the visual effect using Rényi entropy (see Fig. 9g) is more natural.

For the Farmhouse scene, although the fused images are in general very similar, the fused image in Zhao [38] and Rényi contain more details in the window regions (the regions were cut out in Fig. 9).

For the Garden scene Rényi's results can render better the details (see close-up in Fig. 9), at the cost of a less contrasted sky.

For the Lighthouse scene, for Mertens [9], Ma [46] and Merianos [57] the foreground beach looks too darker, whereas for Reny, S.T.Li [11], Li [44] and Zhao [38] looks brighter and the rocks details are easier distinguished (see close-up in Fig. 9). On the other hand, Zhao's rendering of the house looks a bit washed out. Interestingly, this scene is where the methods differ most, apart from the beach observe for instance the different rendition of the streak of light in the sky in the left upper side of the images in Fig. 8.

Finally, in our Rényi result for the Tower scene, the details of the tower are clearly presented, along with the sunlight next

to it. Meanwhile, grass and flowers in tower scene also have a nice visual feeling, which can be seen from the last line in Fig. 9. All in all, from these results, it can be observed that our method can preserve more details and contrast in the over and underexposed regions.

We can also observe that in general the results produced by our new method based on Rényi entropy (Figs 8g and 9g) are brighter than the ones obtained with our previous method [38] based on Shannon CE (Figs 8e and 9e).

4.2. Objective Evaluation

In addition to subjective visual effect comparison, we also introduce objective evaluations, i.e. entropy of fused image, mutual information between source images and fused image, entropy rate of fused image, and MEF-SSIM. The entropy can roughly measure the amount of information in an image. The entropy rate of an image can measure the contrast, as proposed by Vila *et al.* [66]. Different from these two evaluations, mutual information reflects the consistency of the fused image with the source images.

First, from Figs 10-13 we observe that Rényi method notably improves, except for Lighthouse and Garden mutual informa-

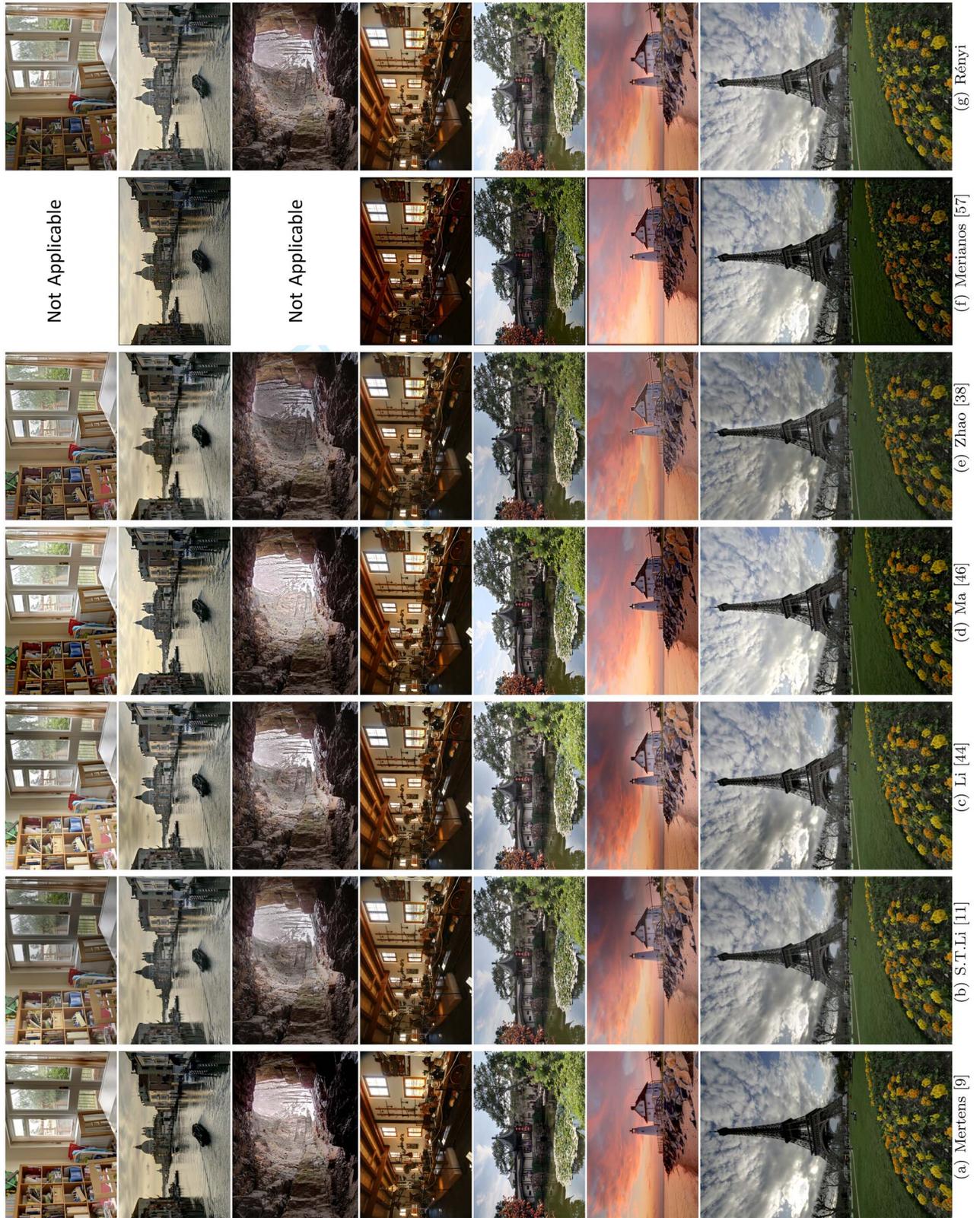


FIGURE 8. Fused results by different methods. House and Cave are not shown with Merianos’ method [57], as it only allows three input images.

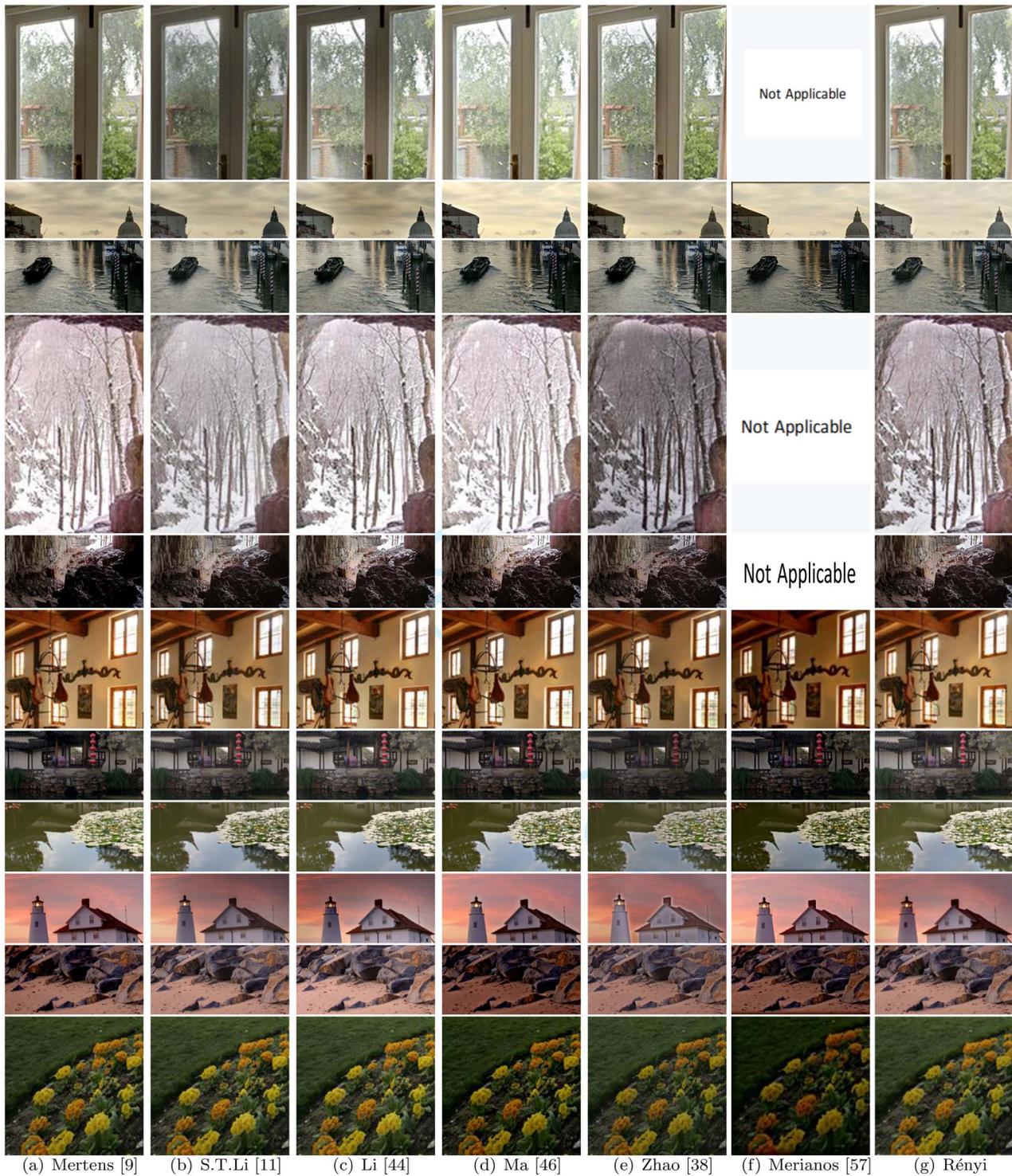


FIGURE 9. Close-up of fused result comparison with other methods. House and Cave are not shown with Merianos' method [57], as it only allows three input images.

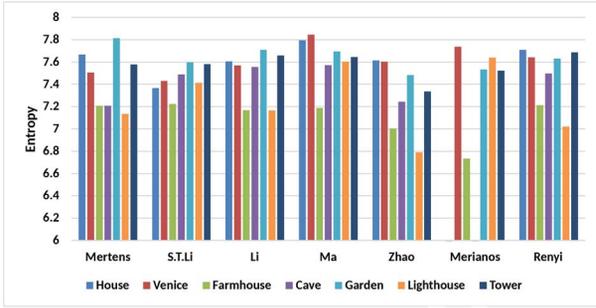


FIGURE 10. Entropy of fused results.

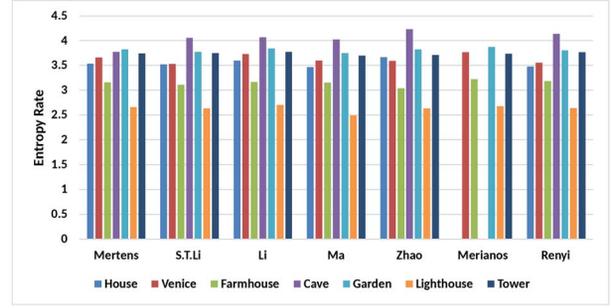


FIGURE 12. Entropy rate of fused results.

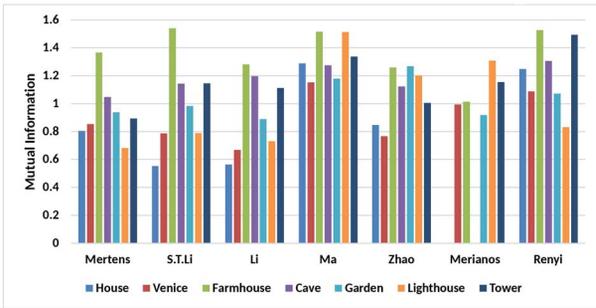


FIGURE 11. Mutual information of fused results.

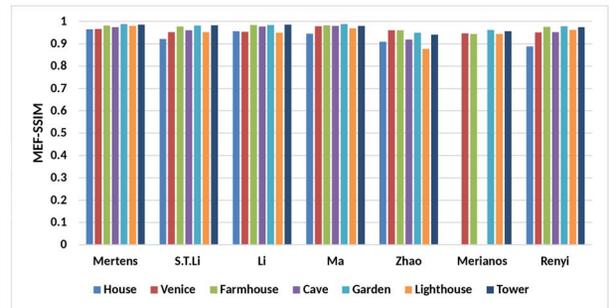


FIGURE 13. MEF-SSIM scores of fused results.

tion, on our previous CE method [38] in all three measures but one, entropy rate, where the values are very similar.

Now, from Fig. 10, we can observe that the resulting entropy of our approach is a second after Ma’s method. We can observe in Fig.11 that mutual information of Rényi’s results is very similar to Ma’s, except for Lighthouse, and both are better than the other methods.

The MEF-SSIM measure, introduced by Ma *et al.* [62], is also employed as an image quality assessment metric of fused results. The results of this measure are shown in Fig. 13, where it can be observed that all the methods score very similarly. Of the four measures considered, MEF-SSIM is the one with smaller differences between the compared methods. The best results are by Ma [46] and then Mertens [9] and Li which score almost identically, maybe because Li’s method is actually a detail-enhanced algorithm built upon Mertens’s method. Our method with Rényi entropy scores a bit better than S.T.Li [11] and Merianos [57] one. Although our Rényi method does not score within the three first ones from Fig. 13, the very small differences with the higher scoring methods make it competitive with respect to MEF-SSIM measure too, in addition to having a solid theoretic background. In summary, our results are competitive and promising when considering all these evaluation indexes.

4.3. Computation complexity

All the experiments are performed on a Windows system computer with Intel Core i7 CPU and 8GB memory. The fusion process is divided into four steps, including pre-treatment, channel building, weight map computing, and fusion by Laplacian pyramid frame. In the pre-treatment step, preprocessing the color image into gray scale image is carried out, as well as computing the relationship between each pair of images.

The consumed time in each step and the respective percentage of total time cost are shown in Table 2. Since we need to compute weights of each pixel, the cost of our method is proportional to the number of pixels in a source image. From Table 2, we can see that the biggest cost corresponds to channel building, as we build information channels between every two source images sequentially instead of computing it in parallel. On the other hand, computing the weight maps and fusing the results is only ~25 or 30% of the total time. It means that we could compute fused images for different values of α with small overhead, because the bulk of the computation, between 70 or 75%, has not to be repeated. The total time of the unoptimized method is ~1 second for three or four source image fusion.

It is difficult in general to compare the complexity with other methods, as usually it is not provided. Zhang [67] provides a benchmark for MEF methods with two source images,

TABLE 2. Time cost in each step, in seconds, and time cost in each step as percentage of total cost. N is the number of input source images and $w \times h$ is their resolution.

Scenes	$w \times h \times N$	Pre-treatment	Channel building	Weight maps	Lap-fusion	Total time
House	$512 \times 340 \times 4$	0.172 (17.01%)	0.585 (57.86%)	0.185 (18.30%)	0.069 (6.82%)	1.011
Venice	$512 \times 341 \times 3$	0.130 (21.10%)	0.293 (47.56%)	0.126 (20.45%)	0.067 (10.88%)	0.616
Cave	$512 \times 384 \times 4$	0.175 (15.10%)	0.668 (57.64%)	0.226 (19.50%)	0.090 (7.77%)	1.159
Farmhouse	$512 \times 341 \times 3$	0.124 (20.88%)	0.295 (49.66%)	0.107 (18.01%)	0.068 (11.45%)	0.594
Tower	$341 \times 512 \times 3$	0.136 (22.04%)	0.296 (47.97%)	0.115 (18.64%)	0.070 (11.35%)	0.617

underexposed and overexposed, and comments about the cost that regarding pixel-based, patch-based, edge-preserving-based and multi-scale-based methods, have running time generally in the same order of magnitude.

5. CONCLUSION AND FUTURE WORK

We have proposed a multi-exposure fusion approach based on building an information channel between each pair of images. Instead of considering each source image independently, we pondered the information that each image in a pair contains about the other one, measured using Rényi entropy. The fused results by the proposed method can preserve more details than other methods, and show the usefulness of considering the MEF problem from an information theoretic perspective. The information of each pixel is obtained by using the luminance distribution, being thus easy to implement. Our fused images deliver visually compelling and promising result versus state-of-art methods with subjective and objective comparison.

There are also some places that need to be improved. We have used only the luminance channel, so the color information is weakened, which explains that our fused results are a little dimmer. We will examine the noise that can be introduced in the uniform regions due to lacking information, and the effect of noise in input images on the weight maps and on the fused image.

Furthermore, we will also consider generalizing it to real-time applications by graphics card implementation, strengthening it through parameter self-adaption, and making it more interactive, allowing users to consider local region improvement. Also, we will study the use of the information channel approach to other image fusion applications, e.g. multi-focus and multi-spectral image fusion.

DATA AVAILABILITY

The data underlying this article will be shared on reasonable request to the corresponding author.

ACKNOWLEDGMENTS

MEF database is from Image and Vision Computing Laboratory (IVC, <https://ivc.uwaterloo.ca/>) at the University of

Waterloo (UW, <http://www.uwaterloo.ca>), Copyright © 2015 The University of Waterloo.

FUNDING

This work is supported by the National Natural Science Foundation of China under grant No.61702359, and by grant PID2019-106426RB-C31 and by grant PID2019-106426RB-C31 from the Agencia Estatal de Investigación (AEI) from Spanish Government. Part of this work was supported by a grant of the Romanian Ministry of Education and Research, CNCS – UEFISCDI, project number PN-III-P1-1.1-TE-2019-1111, within PNCDI III.

REFERENCES

- [1] Bhateja, V., Singhal, A. and Singh, A. (2019) Multi-exposure image fusion method using anisotropic diffusion. In *Third International Congress on Information and Communication Technology*, London, UK, 25-26 February, pp. 893–900. Springer, Berlin.
- [2] Wu, X., Liu, X., Hiramatsu, K. and Kashino, K. (2017) Contrast-accumulated histogram equalization for image enhancement. In *2017 IEEE International Conference on Image Processing (ICIP)*, Beijing, China, 17-20 September, pp. 3190–3194. IEEE, USA.
- [3] Kinoshita, Y., Yoshida, T., Shiota, S. and Kiya, H. (2017) Pseudo multi-exposure fusion using a single image. In *2017 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC)*, Kuala Lumpur, Malaysia, 12-15 December, pp. 263–269. IEEE, USA.
- [4] Su, H. and Jung, C. (2017) Low light image enhancement based on two-step noise suppression. In *2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, New Orleans, LA, USA, 5-9 March, pp. 1977–1981. IEEE, USA.
- [5] Chen, C., Chen, Q., Xu, J. and Koltun, V. (2018) Learning to see in the dark. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Salt Lake City, UT, USA, 18-22 June, pp. 3291–3300. IEEE Computer Society, USA.
- [6] Ren, X., Li, M., Cheng, W.-H. and Liu, J. (2018) Joint enhancement and denoising method via sequential decomposition. In *2018 IEEE International Symposium on Circuits and Systems (ISCAS)*, Florence, Italy, 27-30 May, pp. 1–5. IEEE, USA.

- [7] Hasinoff, S.W., Sharlet, D., Geiss, R., Adams, A., Barron, J.T., Kainz, F., Chen, J. and Levoy, M. (2016) Burst photography for high dynamic range and low-light imaging on mobile cameras. *ACM Trans. Graph.*, 35, 192.
- [8] Goshtasby, A.A. (2005) Fusion of multi-exposure images. *Image Vis. Comput.*, 23, 611–618.
- [9] Mertens, T., Kautz, J. and Van, R.F. (2009) Exposure fusion: A simple and practical alternative to high dynamic range photography. *Compu. Graphics Forum*, 28, 161–171.
- [10] Saleem, A., Beghdadi, A. and Boashash, B. (2012) Image fusion-based contrast enhancement. *EURASIP J. Image Video Proc.*, 2012, 10.
- [11] Li, S. and Kang, X. (2012) Fast multi-exposure image fusion with median filter and recursive filter. *IEEE Trans. Consum. Electron.*, 58, 626–632.
- [12] Wang, J., Xu, G. and Lou, H. (2015) Exposure fusion based on sparse coding in pyramid transform domain. In *Proceedings of the 7th International Conference on Internet Multimedia Computing and Service*, pp. 1–4. ACM, Zhangjiajie, Hunan, China 19–21 August.
- [13] Li, Z., Zheng, J., Zhu, Z. and Wu, S. (2014) Selectively detail-enhanced fusion of differently exposed images with moving objects. *IEEE Trans. Image Process.*, 23, 4372–4382.
- [14] Sakai, T., Kimura, D., Yoshida, T. and Iwahashi, M. (2015) Hybrid method for multi-exposure image fusion based on weighted mean and sparse representation. In *23rd European Signal Processing Conference (EUSIPCO)* Nice, France, August 31 - September 4, pp. 809–813. IEEE, USA.
- [15] Nejati, M., Karimi, M., Soroushmehr, S.R., Karimi, N., Samavi, S. and Najarian, K. (2017) Fast exposure fusion using exposedness function. In *2017 IEEE International Conference on Image Processing (ICIP)*, Beijing, China, 17–20 September, pp. 2234–2238. IEEE, USA.
- [16] Ram Prabhakar, K., Sai Srikar, V. and Venkatesh Babu, R. (2017) Deepfuse: A deep unsupervised approach for exposure fusion with extreme exposure image pairs. In *Proceedings of the IEEE International Conference on Computer Vision*, Venice, Italy, 22–29 October, pp. 4714–4722. IEEE Computer Society, USA.
- [17] Li, Z., Zheng, J., Zhu, Z., Yao, W. and Wu, S. (2015) Weighted guided image filtering. *IEEE Trans. Image Process.*, 24, 120–129.
- [18] Debevec, P.E. and Malik, J. (1997) Recovering high dynamic range radiance maps from photographs. In *Proceedings of the 24th Annual Conference on Computer Graphics and Interactive Techniques, SIGGRAPH 1997*, Los Angeles, CA, USA, 3–8 August, pp. 369–378. ACM.
- [19] Reinhard, E., Stark, M., Shirley, P. and Ferwerda, J. (2002) Photographic tone reproduction for digital images. *ACM Trans. Graph.*, 21, 267–276.
- [20] Oh, T.-H., Lee, J.-Y., Tai, Y.-W. and Kweon, I.S. (2014) Robust high dynamic range imaging by rank minimization. *IEEE Trans. Pattern Anal. Mach. Intell.*, 37, 1219–1232.
- [21] Kinoshita, Y., Shiota, S., Iwahashi, M. and Kiya, H. (2016) An remapping operation without tone mapping parameters for hdr images. *IEICE Trans. Fundam. Electron. Commun. Comput. Sci.*, 99, 1955–1961.
- [22] Kinoshita, Y., Shiota, S. and Kiya, H. (2017) Fast inverse tone mapping based on reinhard’s global operator with estimated parameters. *IEICE Trans. Fundam. Electron. Commun. Comput. Sci.*, 100, 2248–2255.
- [23] Huo, Y. and Zhang, X. (2016) Single image-based hdr imaging with crf estimation. In *2016 International Conference On Communication Problem-Solving (ICCP)*, Taipei, Taiwan, 7–9 September, pp. 1–3. IEEE, USA.
- [24] Murofushi, T., Iwahashi, M. and Kiya, H. (2013) An integer tone mapping operation for hdr images expressed in floating point data. In *2013 IEEE International Conference on Acoustics, Speech and Signal Processing*, Vancouver, BC, Canada, 26–31 May, pp. 2479–2483. IEEE, USA.
- [25] Murofushi, T., Dobashi, T., Iwahashi, M. and Kiya, H. (2014) An integer tone mapping operation for hdr images in openxr with denormalized numbers. In *2014 IEEE International Conference on Image Processing (ICIP)*, Paris, France, 27–30 October, pp. 4497–4501. IEEE, USA.
- [26] Dobashi, T., Murofushi, T., Iwahashi, M. and Kiya, H. (2014) A fixed-point global tone mapping operation for hdr images in the rgbe format. *IEICE Trans. Fundam. Electron. Commun. Comput. Sci.*, 97, 2147–2153.
- [27] Unger, J. and Gustavson, S. (2007) High dynamic range video for photometric measurement of illumination. *Electron. Imag.*, 6501, 1–10.
- [28] Granados, M., Ajdin, B., Wand, M., Theobalt, C. and Lensch, H.P.A. (2010) Optimal hdr reconstruction with linear digital cameras. In *2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, San Francisco, CA, USA, 13–18 June, pp. 215–222. IEEE Computer Society, USA.
- [29] Kronander, J., Gustavson, S., Bonnet, G., Ynnerman, A. and Unger, J. (2014) A unified framework for multi-sensor hdr video reconstruction. *Signal Proc. Image Commun.*, 29, 203–215.
- [30] Kalantari, N.K. and Ramamoorthi, R. (2017) Deep high dynamic range imaging of dynamic scenes. *ACM Trans. Graph.*, 36, 144.
- [31] Eilertsen, G., Kronander, J., Denes, G., Mantiuk, R.K. and Unger, J. (2017) Hdr image reconstruction from a single exposure using deep cnns. *ACM Trans. Graph.*, 36, 178.
- [32] Kalantari, N.K. and Ramamoorthi, R. (2019) Deep hdr video from sequences with alternating exposures. *Compu. Graphics Forum*, 38, 193–205.
- [33] Ulucan, O., Karakaya, D. and Turkan, M. (2021) Multi-exposure image fusion based on linear embeddings and watershed masking. *Signal Process.*, 178, 107791.
- [34] Goshtasby, A.A. and Nikolov, S. (2007) Image fusion: Advances in the state of the art. *Info. Fusion*, 8, 114–118.
- [35] Li, S., Kang, X., Fang, L., Hu, J. and Yin, H. (2017) Pixel-level image fusion: A survey of the state of the art. *Info. Fusion*, 33, 100–112.
- [36] Bramon, R., Boada, I., Bardera, A., Rodriguez, J., Feixas, M., Puig, J. and Sbert, M. (2012) Multimodal data fusion based on mutual information. *IEEE Trans. Vis. Comput. Graph.*, 18, 1574–1587.
- [37] Feixas, M., Sbert, M., and Chen M. (eds) (2020) The role of the information channel in visual computing. *Adv. Info-Metrics*, Oxford University Press (OUP), 464–490. ISBN: 9780190636685.
- [38] Zhao, Q., Sbert, M., Feixas, M. and Xu, Q. (2018) Multi-exposure image fusion based on information-theoretic channel. In *2018 IEEE International Conference on Image Process-*

- ing(ICIP)*, Athens, Greece, 7-10 October, pp. 1872–1876. IEEE, USA.
- [39] Burt, P.J. and Kolczynski, R.J. (1993) Enhanced image capture through fusion. In *1993 International Conference on Computer Vision*, Berlin, Germany, 11-14 May, pp. 173–182. IEEE Computer Society, USA.
- [40] Li, Z.G., Zheng, J.H. and Rahardja, S. (2012) Detail-enhanced exposure fusion. *IEEE Trans. Image Process.*, 21, 4672–4676.
- [41] Li, Z., Wei, Z., Wen, C. and Zheng, J. (2017) Detail-enhanced multi-scale exposure fusion. *IEEE Trans. Image Process.*, 26, 1243–1252.
- [42] Ancuti, C.O., Ancuti, C., Vleeschouwer, C.D. and Bovik, A.C. (2017) Single-scale fusion: An effective approach to merging images. *IEEE Trans. Image Process.*, 26, 65–78.
- [43] Raman, S. and Chaudhuri, S. (2009) Bilateral filter based compositing for variable exposure photography. In *Eurographics*, pp. 1–4.
- [44] Li, S., Kang, X. and Hu, J. (2013) Image fusion with guided filtering. *IEEE Trans. Image Process.*, 22, 2864–2875.
- [45] Patel, D., Sonane, B. and Raman, S. (2017) Multi-exposure image fusion using propagated image filtering. In *Proceedings of International Conference on Computer Vision and Image Processing* Roorkee, India, 9-12 September, pp. 431–441. Springer, Berlin.
- [46] Ma, K. and Wang, Z. (2015) Multi-exposure image fusion: A patch-wise approach. In *2015 IEEE International Conference on Image Processing (ICIP)*, Quebec City, QC, Canada, 27-30 September, pp. 1717–1721. IEEE, USA.
- [47] Zhang, W. and Cham, W.-K. (2011) Gradient-directed multi-exposure composition. *IEEE Trans. Image Process.*, 21, 2318–2323.
- [48] Shen, R., Cheng, I. and Basu, A. (2012) Qoe-based multi-exposure fusion in hierarchical multivariate gaussian CRF. *IEEE Trans. Image Process.*, 22, 2469–2478.
- [49] Yu, M., Song, E., Jin, R., Liu, H., Xu, X. and Ma, G. (2015) A novel method for fusion of differently exposed images based on spatial distribution of intensity for ubiquitous multimedia. *Multimed. Tools Appl.*, 74, 2745–2761.
- [50] Dong, X., Shen, J. and Shao, L. (2016) Hierarchical superpixel-to-pixel dense matching. *IEEE Trans. Circuits Syst. Video Technol.*, 27, 2518–2526.
- [51] Qin, X., Shen, J., Mao, X., Li, X. and Jia, Y. (2014) Robust match fusion using optimization. *IEEE Trans. Cybernet.*, 45, 1549–1560.
- [52] Wang, W., Shen, J., Li, X. and Porikli, F. (2015) Robust video object cosegmentation. *IEEE Trans. Image Process.*, 24, 3137–3148.
- [53] Shen, J., Hao, X., Liang, Z., Liu, Y., Wang, W. and Shao, L. (2016) Real-time superpixel segmentation by dbscan clustering algorithm. *IEEE Trans. Image Process.*, 25, 5933–5942.
- [54] Wang, W., Shen, J., Yang, R. and Porikli, F. (2017) Saliency-aware video object segmentation. *IEEE Trans. Pattern Anal. Mach. Intell.*, 40, 20–33.
- [55] Shen, R., Cheng, I., Shi, J. and Basu, A. (2011) Generalized random walks for fusion of multi-exposure images. *IEEE Trans. Image Process.*, 20, 3634–3646.
- [56] Song, M., Tao, D., Chen, C., Bu, J., Luo, J. and Zhang, C. (2011) Probabilistic exposure fusion. *IEEE Trans. Image Process.*, 21, 341–357.
- [57] Merianos, I. and Mitianoudis, N. (2019) Multiple-exposure image fusion for HDR image Synthesis Using Learned Analysis Transformations. *J. Imaging*, 5, 32.
- [58] Herwig, J. and Pauli, J. (2010) An information-theoretic approach to multi-exposure fusion via statistical filtering using local entropy. In *Proceedings of the Seventh IASTED International Conference on Signal Processing, Pattern Recognition and Applications*, pp. 50–57. ACTA Press.
- [59] Gonzalez, R.C. and Woods, R.E. (2007) *Digital Image Processing (3rd Edition)*. Prentice-Hall, Inc, New Jersey, US.
- [60] Dewese, M.R. and Meister, M. (1999) How to measure the information gained from one symbol. *Netw. Comput. Neural Syst.*, 10, 325–340.
- [61] Rényi, A. (1960) On measures of information and entropy. In *Proceedings of the fourth Berkeley Symposium on Mathematics Statistics and Probability*, pp. 547–561. Society for Industrial and Applied Mathematics, California, Berkeley June 20-July 30.
- [62] Ma, K., Zeng, K. and Wang, Z. (2015) Perceptual quality assessment for multi-exposure image fusion. *IEEE Trans. Image Process.*, 24, 3345–3356.
- [63] Zeng, K., Ma, K., Hassen, R. and Wang, Z. (2015) Perceptual Evaluation of Multi-exposure Image Fusion Algorithms. *The 6th International Workshop on Quality of Multimedia Experience (QoMEX)*, 2014.
- [64] Kede Ma Webpage, <https://kedema.org/Publications.html#2015> (accessed August 6, 2021).
- [65] Nikolaos Mitianoudis Webpage, <http://utopia.duth.gr/nmitiano/index.html> (August 3, 2021).
- [66] Vila, M., Bardera, A., Feixas, M., Bekaert, P. and Sbert, M. (2014) Analysis of image informativeness measures. In *2014 IEEE International Conference on Image Processing (ICIP)*, Paris, France, 27-30 October, pp. 1086–1090. IEEE, USA.
- [67] Zhang, X. (2021) Benchmarking and comparing multi-exposure image fusion algorithms. *Inf. Fusion*, 74, 111–131.