

Article

BMEFIQA: Blind Quality Assessment of Multi-Exposure Fused Images Based on Several Characteristics

Jianping Shi, Hong Li *, Caiming Zhong, Zhouyan He * and Yeling Ma

College of Science and Technology, Ningbo University, Ningbo 315300, China; shijianpingky@126.com (J.S.); zhongcaiming@nbu.edu.cn (C.Z.); mayeling@nbu.edu.cn (Y.M.)

* Correspondence: ky_lihong@nbu.edu.cn (H.L.); hezhouyan@nbu.edu.cn (Z.H.)

Abstract: A multi-exposure fused (MEF) image is generated by multiple images with different exposure levels, but the transformation process will inevitably introduce various distortions. Therefore, it is worth discussing how to evaluate the visual quality of MEF images. This paper proposes a new blind quality assessment method for MEF images by considering their characteristics, and it is dubbed as BMEFIQA. More specifically, multiple features that represent different image attributes are extracted to perceive the various distortions of MEF images. Among them, structural, naturalness, and colorfulness features are utilized to describe the phenomena of structure destruction, unnatural presentation, and color distortion, respectively. All the captured features constitute a final feature vector for quality regression via random forest. Experimental results on a publicly available database show the superiority of the proposed BMEFIQA method to several blind quality assessment methods.

Keywords: blind quality assessment; multi-exposure fused images; structure; naturalness; colorfulness



Citation: Shi, J.; Li, H.; Zhong, C.; He, Z.; Ma, Y. BMEFIQA: Blind Quality Assessment of Multi-Exposure Fused Images Based on Several Characteristics. *Entropy* **2022**, *24*, 285. <https://doi.org/10.3390/e24020285>

Academic Editor:
Gholamreza Anbarjafari

Received: 20 December 2021

Accepted: 8 February 2022

Published: 16 February 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

In view of the increasing development of image processing technologies, it has become more feasible to help humans perceive the real world in high-quality images. For example, multi-exposure fusion (MEF) and high dynamic range (HDR) imaging both belong to image enhancement technology, which can provide excellent detail information and an ideal, natural appearance [1]. HDR imaging technology requires additional processing, such as generation, image conversion to a low dynamic range image, and visualization on common displays. Unfortunately, all of these procedures can produce artifacts that affect the quality of HDR images [2]. Multi-exposure fusion technology is relatively simple, it does not need intermediate operations, and it has also been applied in some practical fields [3]. However, the unique process of integration from multiple different exposure images into one ultimate image will result in distortion due to the fusion weighting assignment. Obviously, the distortion will influence human perception and result in different human opinion scores. Hence, to measure the distortion degree objectively, it is necessary to develop a quality assessment method for MEF images.

In recent years, many researchers have developed some image quality assessment (IQA) methods [4–14]. Those methods can be categorized into full-reference (FR) methods, reduced-reference (RR) methods, and no-reference (NR) methods. The FR and RR methods both need original image information to make a comparison; however, it is opposite to the truth that there is never a pre-defined reference in the real application. Therefore, it is more urgent to design NR methods for IQA tasks.

Based on the considerations of the inherent characteristics of MEF images, a new NR-IQA method is proposed in this paper to predict the quality of MEF images more accurately, and it is named BMEFIQA.

The main contributions are detailed as follows:

- (1) Inspired by the various characteristics of MEF images, structural, naturalness, and colorfulness features are extracted from different standpoints to perceive their distortion.
- (2) On account of structure loss produced by abnormal exposure, the exposure map is weighted to the gradient similarity to detect structural distortion.
- (3) The experimental results demonstrate that the proposed method is competent for MEF images and superior to several NR-IQA methods.

The remainder of this paper is constructed as follows: In Section 2, related works are presented. In Section 3, the proposed BMEFIQA method is described in detail. The experimental results and analysis based on a public MEF image database are presented in Section 4. Finally, the conclusion is drawn in Section 5.

2. Related Works

Currently, there are many efficient NR-IQA methods for ordinary images. For instance, Moorthy et al. [5] proposed an NR-IQA method named the distortion identification-based image verity and integrity evaluation (DIIVINE), which is based on natural scene statistics (NSSs). In [6], Saad et al. utilized an NSS model of discrete cosine transform coefficients to design the IQA method, which is referred to as BLINDS-II. In [7], Mittal et al. extracted features from the empirical distribution of locally normalized luminance and the products of locally normalized luminance in the spatial domain; the method is named BRISQUE. Liu et al. [8] constructed the CurveletQA method by utilizing curvelet transform to extract a set of statistical features. Xue et al. [9] combined the gradient magnitude map with the Laplacian of Gaussian response to perceive the structural information of images and showed highly competitive performance, dubbed as GradLog. Fang et al. [10] employed the degree of deviation from NSS models to represent the unnatural character of contrast-distorted images, which is termed ContrastQA. Li et al. [11] proposed an NR-IQA method based on structural degradation, which is described by the gradient-weighted histogram of local binary pattern (LBP) calculation on the gradient map (GWH-GLBP). Liu et al. [12] developed the oriented gradient (OG) IQA method by studying the quality relevance of the relative gradient orientation. Gu et al. [13] combined local and global considerations to design an NR-IQA method called NIQMC. Oszust [14] captured the information carried by image derivatives of different orders by local features and used it for image quality prediction. Zhang et al. [15] integrated the features of natural image statistics and created a multivariate Gaussian model of image patches for quality assessment. Xu et al. [16] proposed an NR-IQA method based on high-order statistics aggregation, which needs a small codebook.

Except for the above IQA methods for ordinary images, some studies have also contributed to predicting the quality of tone-mapped images. For example, Gu et al. [17] devised an effective blind tone-mapped quality index (BTMQI) via the analysis of information, naturalness, and structure. Kundu et al. [18] derived the HDR image gradient-based evaluator (HIGRADE) based on standard measurements of the bandpass and newly conceived differential NSS. Although the above IQA methods for ordinary and tone-mapped images showed good performance, they are not appropriate for the prediction of the quality of MEF images, as their distortion type is distinguished from the above two types of images.

For MEF images, there also exist some IQA methods, which belong to FR methods. For instance, Zheng et al. [19] proposed a quantitative metric called the ratio of spatial frequency error; this is derived from the definition of spatial frequency, which reflects local intensity variation. Ma et al. [20] proposed an objective method based on the principle of structural similarity and a novel measure of patch structural consistency. Xing et al. [21] combined contrast information with structural similarity and saturation similarity to predict the quality of an MEF image. Fang et al. [22] utilized a pyramid subband contrast preservation scheme and an information theory-adaptive pooling strategy to establish a quality assessment method for MEF images of both static and dynamic scenes. Deng et al. [23] designed a method by extracting color, texture, and structural features. Martinez et al. [24] utilized the multi-scale scheme to compute structural similarities. Considering the limitation of such FR

methods, which need reference information, an NR method was proposed in our previous work [25]. However, there is room for improvement. Therefore, to predict the quality of an MEF image more accurately, its characteristics should be taken into consideration more comprehensively.

3. The Proposed BMEFIQA Method

Since MEF images are fused by multiple images with different exposure levels [26], the weight assignment process will introduce some artifacts, such as detail loss, structural degradation, unnaturalness, and color distortion. Figure 1 gives a vivid presentation of MEF images generated by three different MEF algorithms, which show the different visual effects. Figure 1a, generated by Merten's algorithm [27], has the highest mean opinion score (MOS) and is the best quality. The detailed information is well preserved, except for the over-exposed part (i.e., outside the entrance of the cave). Moreover, it also has much more abundant color information than the others, and it seems more natural. At a first glance of Figure 1b, which is generated by Raman's algorithm [28], the blackened scene is the initial perception of humans. In fact, the brightness decline also leads to detail and color information loss. The MEF image generated by local energy weighting [29] shows proper brightness, but it loses the original naturalness, and it also introduces some artifacts around the wall and stones. Therefore, structural, naturalness, and color information are the main factors that influence the visual quality of MEF images.



Figure 1. An example of multi-exposure fusion images generated by three MEF algorithms. (a) MEF image generated by Merten's algorithm [27] (MOS = 7.6957); (b) MEF image generated by Raman's algorithm [28] (MOS = 4.2609); (c) MEF image generated by local energy weighting [29] (MOS = 4.3043).

To fill the gap of quality prediction for MEF images under the condition with a lack of reference information, a novel method named BMEFIQA is proposed in this paper, which considers three factors (i.e., structure, naturalness, and colorfulness). Figure 2 presents the pipeline of the proposed BMEFIQA method. Specifically, three types of quality-sensitive features corresponding to the above three factors are extracted to generate overall feature vectors. Then, random forest is utilized to train a quality regression model, thus aggregating all excavated features. More details are given in the following subsections.

3.1. Structural Features

Structural information usually carries the basic visual content of a scene, and the human visual system (HVS) has strong adaptability to extract the structure for visual perception [30]. For an MEF image, distortion introduction will always destroy its structural information, such as abnormal exposure (i.e., over-exposure or under-exposure). Therefore, the visual quality of an MEF image can be determined by measuring whether the structural information is damaged or not, especially for over-exposed and under-exposed regions.

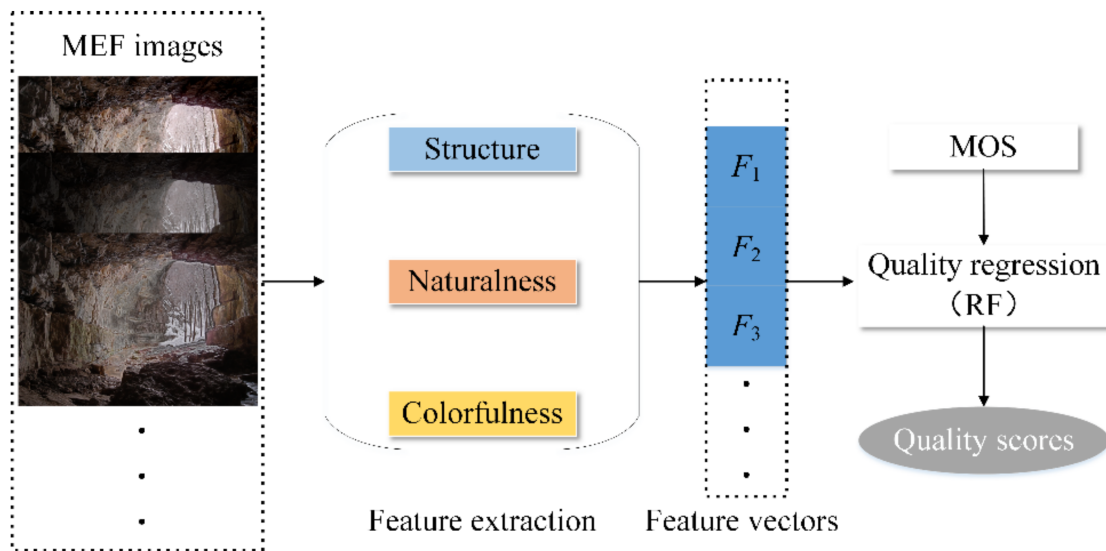


Figure 2. The framework of the proposed BMEFIQA method.

First, the exposure of the MEF image needs to be measured to distinguish the regions of over-exposure and under-exposure. Given MEF image I and converting it into a gray-scale one, its corresponding exposure map E_I can be calculated by measuring the distance between its normalized pixel intensity and the constant 0.5. Specifically, when the normalized pixel intensity is close to 0 or 1, the corresponding pixel is regarded as under-exposed or over-exposed. Therefore, E_I can be defined as

$$E_I = 1 - \exp \frac{(I_y - 0.5)^2}{2\tau^2} \quad (1)$$

where I_y is the normalized pixel intensity of the gray-scaled MEF image I , and τ is the standard variance of the Gaussian function, which is set to 0.2 according to previous experience [31].

Figure 3 shows three exposure maps of the MEF images in Figure 1. By comparing with Figure 1, it can be observed that the brighter regions are the under-exposed or over-exposed regions, and the fainter regions represent the normal exposed regions in Figure 3.

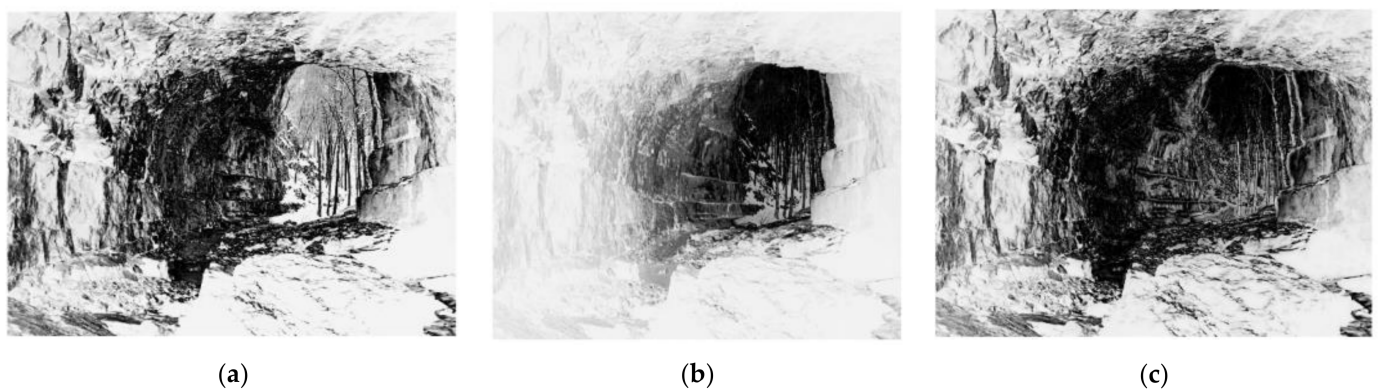


Figure 3. (a–c) The corresponding exposure maps of the MEF images in Figure 1.

As an effective structural information feature, the gradient is utilized to describe the structure loss phenomenon. To further measure the influence of abnormal exposure on the acquisition of gradient information, some fake MEF images are obtained by darkening and brightening the real MEF images, which are denoted as I_f . They can be produced as follows:

$$I_f = I_r \cdot C \quad (2)$$

where $C = c$, $c \in \{1/3.5, 1/5, 1/6.5, 3.5, 5, 6.5\}$, and I_r represents the gray-scaled MEF image I . After that, six fake MEF images can be generated.

Then, gradient maps G of such fake MEF images are calculated as

$$G = \sqrt{(I_f \otimes p_x)^2 + (I_f \otimes p_y)^2} \quad (3)$$

where p_x and p_y are the Prewitt filters, along with the horizontal and vertical directions, respectively. \otimes indicates the convolution operator. Additionally, the gradient map of the real MEF image I is also obtained by (3), which is denoted as G_I .

Then, the gradient similarities between I and the generated six fake MEF images are calculated. Let G_s be the similarities of G and G_I , which are calculated to quantify the influence of abnormal exposure. They are defined as

$$G_s = \frac{2G \cdot G_I + C_1}{G^2 + G_I^2 + C_1} \quad (4)$$

where C_1 is a constant to avoid zero denominators, which is set to 10^{-8} .

Since the distortions of the over-exposed and the under-exposed regions are easier for the human eye to perceive, combining the gradient similarities G_s with the exposure map E_I can make the distortion perception more accurate. Therefore, gradient combined with exposure weighting is defined as

$$G_e = \frac{\sum_k \sum_f G_s \cdot E_I}{\sum_k \sum_f E_I} \quad (5)$$

where k and f are the indexes of the horizontal and vertical pixels.

As the same number of fake MEF images, after combining six gradient similarity maps with E_I , the obtained G_e is a 6-dimensional feature that describes the structure loss of the MEF image.

From the statistical perspective, the NSS model is utilized in the gradient domain to represent the structural variation in the MEF images. Specifically, the gradient map G_I is processed by the local mean subtraction and divisive normalization to obtain the mean subtracted contrast normalized (MSCN) coefficients [7], which are expressed by

$$\hat{G}_I(i, j) = \frac{G_I(i, j) - \mu(i, j)}{\delta(i, j) + 1} \quad (6)$$

where $\hat{G}_I(i, j)$ are the MSCN coefficients of G_I at the position of (i, j) . $\mu(i, j)$ and $\delta(i, j)$ are the local mean and standard deviation of $G_I(i, j)$, respectively.

Different degrees of structural distortion will inevitably affect the distribution of the MSCN coefficients of G_I . As the distribution has a Gaussian-like appearance, a generalized Gaussian distribution (GGD) is utilized to match the MSCN coefficients, and the mathematical expression is given by

$$f(l, \alpha, \sigma^2) = \frac{\alpha}{2\beta\Gamma(1/\alpha)} \cdot e^{-(|l|/\beta)^\alpha} \quad (7)$$

where $\beta = \sigma\sqrt{\Gamma(1/\alpha)\Gamma(3/\alpha)}$, $\Gamma(\cdot)$ represents the gamma function. The parameters α and σ^2 represent the shape and variance of the Gaussian distribution, respectively. Therefore, the two parameters α and σ^2 are taken as the 2-dimensional global structural features.

In addition, the pairwise products of neighboring MSCN coefficients are also calculated to capture the relationship of the neighboring pixels. The MSCN coefficients are processed in four directions, and this process is expressed as

$$\begin{aligned}
H(i, j) &= \hat{G}_1(i, j) \hat{G}_1(i, j + 1) \\
V(i, j) &= \hat{G}_1(i, j) \hat{G}_1(i + 1, j) \\
D_1(i, j) &= \hat{G}_1(i, j) \hat{G}_1(i + 1, j + 1) \\
D_2(i, j) &= \hat{G}_1(i, j) \hat{G}_1(i + 1, j - 1)
\end{aligned} \tag{8}$$

where $H(i, j)$, $V(i, j)$, $D_1(i, j)$, and $D_2(i, j)$ are the results processed along with the horizontal, vertical, main diagonal, and sub-diagonal directions, respectively.

Then, an asymmetric generalized Gaussian distribution (AGGD) is utilized to fit each pairwise product, which is defined as

$$f(k; v, \omega_l^2, \omega_r^2) = \begin{cases} \frac{v}{(\beta_l + \beta_r) \Gamma(v-1)} \exp\left(-\left(\frac{-k}{\beta_l}\right)^v\right), & k < 0 \\ \frac{v}{(\beta_l + \beta_r) \Gamma(v-1)} \exp\left(-\left(\frac{-k}{\beta_r}\right)^v\right), & k \geq 0 \end{cases} \tag{9}$$

$$\varphi = (\beta_r - \beta_l) \frac{\Gamma(2/v)}{\Gamma(1/v)} \tag{10}$$

where $\beta_l = \omega_l \sqrt{\Gamma(v-1)/\Gamma(v-3)}$ and $\beta_r = \omega_r \sqrt{\Gamma(v-1)/\Gamma(v-3)}$.

The above parameters v , φ , ω_l^2 , and ω_r^2 constitute the compensation features that perceive the global structural distortion. As a result, the compensation features calculated in the four directions form a 16-dimensional feature vector.

However, the phenomena of structure loss also contain the loss of detailed information, which can be measured via entropy. Block entropy is calculated to perceive the local detail information variation, which is denoted as e_b . After calculating the local entropy in 8×8 blocks of each MEF image, the entropy calculation is used to measure the distribution of the obtained local entropy, which can be expressed as

$$e = - \sum_{n=1}^m p(e_b(n)) \log_2 p(e_b(n)) \tag{11}$$

where m is the block number of each MEF image, and $p(\cdot)$ is the probability density of the m -th block entropy value.

Furthermore, the mean value and standard deviation of the block entropy are also calculated to measure the overall detail loss, which are denoted as e_m and e_s , respectively. Finally, e , e_m , and e_s are integrated to form another set of 3-dimensional structural features.

3.2. Naturalness Features

Generally, a high-quality MEF image has a natural-like appearance. MEF algorithms may disrupt the natural statistical regularities in the spatial domain. From a global perspective, the naturalness can be quantified via the NSS-based model. An MEF image, I , is converted to a gray-scale image, I_r , and then the MSCN coefficients of I_r are calculated according to Equation (6). Moreover, the GGD model defined in Equation (7) is utilized to capture the statistical property, and the shape and variance parameters are taken as the first group of global naturalness features.

From another perspective, naturalness is also affected by the overall brightness and contrast, which is seriously influenced by under-exposed and over-exposed conditions. It has been demonstrated that the mean and standard deviation values of the image intensity can represent the brightness and contrast of an image [32]; the entropy can also describe the distortion of brightness. Therefore, these moment features (mean and standard deviation) and the entropy feature are applied to build NSS models to capture the naturalness variation undergoing the distortions. Among them, Gaussian probability density functions are used to fit the mean and standard deviation. The specific definitions are as follows:

$$d_a = \frac{1}{\sqrt{2\pi}\zeta_a} \exp\left[-\frac{(a - \tau_a)^2}{2\zeta_a^2}\right] \tag{12}$$

$$d_s = \frac{1}{\sqrt{2\pi}\zeta_s} \exp\left[-\frac{(s - \tau_s)^2}{2\zeta_s^2}\right] \quad (13)$$

where a and s are the mean and standard deviation values, respectively. d_a and d_s are the possibilities of the MEF image being natural when a and s are given. The parameters in Equations (12) and (13) are set to $\zeta_a = 26.063$, $\tau_a = 118.559$, $\zeta_s = 12.858$, and $\tau_s = 57.274$, respectively.

In addition, the entropy feature is fitted by the extreme value probability density function, which is defined as

$$d_o = \frac{1}{\zeta_o} \exp\left[\frac{o - \tau_o}{\zeta_o} - \exp\left[\frac{o - \tau_o}{\zeta_o}\right]\right] \quad (14)$$

where o is the entropy value of I_r . d_o is the possibility of the MEF image being natural when o is given. ζ_o and τ_o are the parameters that are set to 0.258 and 7.54, respectively.

As a result, d_a , d_s , and d_o form the second group of global naturalness features. Combined with the first group of features, 5-dimensional features are used to describe the naturalness of the MEF image.

3.3. Colorfulness Features

Abundant color information is important for an outstanding MEF image, as it illustrates that the image has proper color saturation and realistic scene chroma. The weight assignment processes in the different MEF algorithms may emphasize different parts of the image content, as different humans do not always focus on the same things. As Figure 1 shows, different emphasis brings different distortion. Nevertheless, the color distortion measure is necessary and right. In previous work, it was proved that the color perception of human vision is mainly processed in the opponent color space [33]. Therefore, the opponent color space is obtained by the red–green–blue (RGB) color channels in the RGB color space. The transformation processes are expressed as $T_1 = R - G$, $T_2 = (R + G)/2 - B$; T_1 is the obtained red–green channel, and T_2 is the obtained yellow–blue channel.

A combination of the two transformed channels is utilized to represent the colorfulness of the MEF image. The specific definition is as follows:

$$C_t = \log(\gamma_{T_1}^2 / |a_{T_1}|^{0.2}) \cdot \log(\gamma_{T_2}^2 / |a_{T_2}|^{0.2}) \quad (15)$$

where $\gamma_{T_1}^2$ and $\gamma_{T_2}^2$ denote the variances in T_1 and T_2 , respectively. a_{T_1} and a_{T_2} denote the mean value of T_1 and T_2 , respectively.

In Figure 1, it can be seen that over-exposure and under-exposure will both affect the representation of excellent color with a decrease in contrast. Therefore, the contrast energies are calculated in the opponent color space to perceive the color distortion. Let C_e be the calculated contrast energy; it is obtained by

$$C_e = \frac{\varepsilon \cdot \rho(I_g)}{\rho(I_g) + \varepsilon \cdot \chi} - \vartheta_g \quad (16)$$

where $g \in \{T_1, T_2\}$; I_{T_1} and I_{T_2} denote the red–green and yellow–blue channels of I , respectively. $\rho(I_g) = \sqrt{(I_g \otimes p_x)^2 + (I_g \otimes p_y)^2}$; χ is the contrast gain to correct and normalize all filter responses, ε is the maximum value of $\rho(I_g)$, and ϑ_g is the noise threshold. χ and ϑ_g are fixed at the same settings as in [34].

Finally, the mean value of C_e is calculated as another feature except for C_t . In fact, according to the opponent color space, two channels can be used to perform the contrast energy calculation process, and 2-dimensional features are obtained.

Moreover, to obtain a more complete sense of color distortion, the NSS model is also executed in the two opponent color channels. Specifically, T_1 and T_2 are processed with the MSCN coefficient calculation defined in Equation (6) and the GGD model defined

in Equation (7), respectively. The obtained shape and variance parameters constitute 4-dimensional complementary color features.

3.4. Feature Aggregation and Quality Regression

So far, 39-dimensional features are extracted via structure, naturalness, and colorfulness analyses, denoted as S_F , N_F , and C_F , respectively. According to the obtained quality-sensitive features, the mapping relationship between the features and human subjective scores should be learned to achieve the purpose of quality prediction. As an effective regression manner, random forest (RF) is used to pool the high-dimensional features to indicate the quality of the MEF image. The specific expression can be represented as follows:

$$Q = \eta(F) \quad (17)$$

where Q is the final quality score of the MEF image; $\eta(\cdot)$ denotes the mapping function by RF; and F denotes the final all-feature vectors, $F = \{S_F, N_F, C_F\}$.

4. Experiment Results

4.1. Experimental Protocol

In the experiment, an open MEF subjective assessment database [35] is selected to evaluate the performance of the proposed BMEFIQA method. The database consists of 136 MEF images, which portray different scenarios. The corresponding perceptual quality score is subjectively tested by numerous observers, that is, human subjective scores, and the MOS values of each MEF image are obtained from them. These MEF images are generated by distinct types of MEF algorithms, including local energy weighted linear combination, global energy weighted linear combination, Li12 [36], Raman09 [28], ShutaoLi12 [37], Mertens07 [27], Gu12 [29], and ShutaoLi13 [38]. Specifically, 136 MEF images are derived from 17 source multi-exposure images. The detailed contents are reported in Table 1 and Figure 4.

Table 1. The details of the MEF image database [35].

| No. | Source Sequences | Size | Image Source |
|-----|------------------|----------------------------|--------------------|
| 1 | Balloons | $339 \times 512 \times 9$ | Erik Reinhard |
| 2 | Belgium house | $512 \times 384 \times 9$ | Dani Lischinski |
| 3 | Lamp1 | $512 \times 384 \times 15$ | Martin Cadik |
| 4 | Candle | $512 \times 364 \times 10$ | HDR Projects |
| 5 | Cave | $512 \times 384 \times 4$ | Bartlomiej Okonek |
| 6 | Chinese garden | $512 \times 340 \times 3$ | Bartlomiej Okonek |
| 7 | Farmhouse | $512 \times 341 \times 3$ | HDR Projects |
| 8 | House | $512 \times 340 \times 4$ | Tom Mertens |
| 9 | Kluki | $512 \times 341 \times 3$ | Bartlomiej Okonek |
| 10 | Lamp2 | $512 \times 342 \times 6$ | HDR Projects |
| 11 | Landscape | $512 \times 341 \times 3$ | HDRsoft |
| 12 | Lighthouse | $512 \times 340 \times 3$ | HDRsoft |
| 13 | Madison capitol | $512 \times 384 \times 30$ | Chaman Singh Verma |
| 14 | Memorial | $341 \times 512 \times 16$ | Paul Debevec |
| 15 | Office | $512 \times 340 \times 6$ | Matlab |
| 16 | Tower | $341 \times 512 \times 3$ | Jacques Joffre |
| 17 | Venice | $512 \times 341 \times 3$ | HDRsoft |

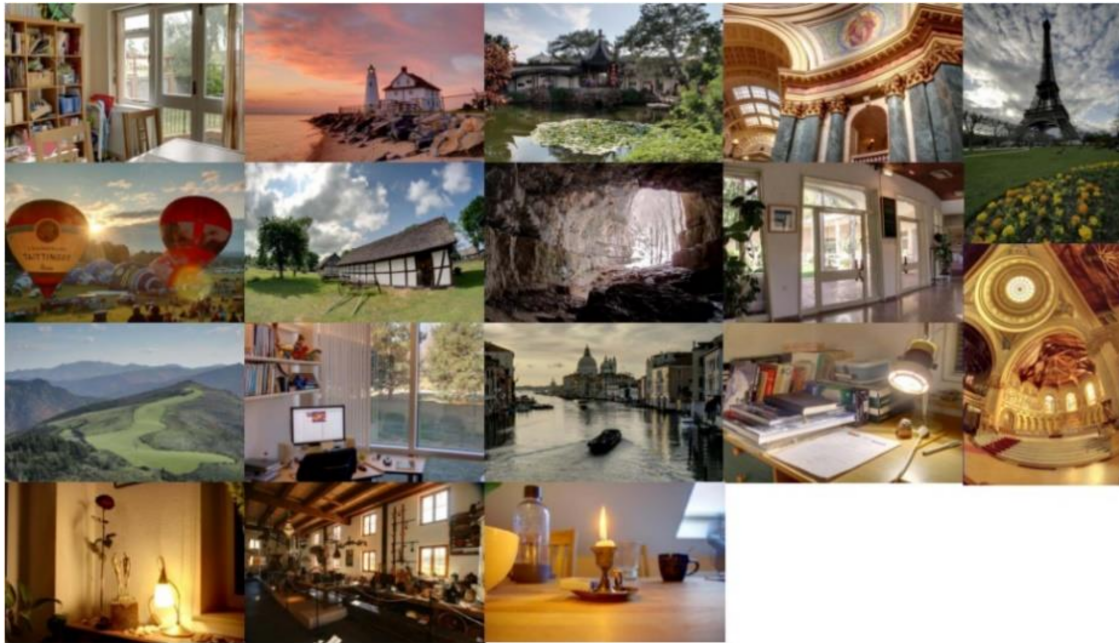


Figure 4. Source images in the benchmark MEF database [35].

Three standard performance evaluation criteria are utilized to evaluate the performance of the proposed method, and they include the Person linear correlation coefficient (PLCC), Spearman's rank-order correlation coefficient (SROCC), and root-mean-square error (RMSE). Specifically, PLCC, SROCC, and RMSE are used to evaluate the prediction accuracy, prediction monotonicity, and prediction error, respectively. Moreover, before the calculation of the PLCC value, a five-parameter logistic regression is used to perform nonlinear mapping between the predicted score and the subjective quality scores; the definition is given by

$$Q_f = \lambda_1 \left[\frac{1}{2} - \frac{1}{\exp(\lambda_2(Q - \lambda_3))} \right] + \lambda_4 Q + \lambda_5 \quad (18)$$

where Q_f is the fitted score, and λ_1 , λ_2 , λ_3 , λ_4 , and λ_5 are the fitting parameters, which can be obtained via the *nlinfit* function in MATLAB software. The initial values of these five fitting parameters are guided by the video quality expert group (VQEG) [39]. The final values are determined by nonlinear least-squares optimization between the subjective quality scores and Q_f . When calculating the PLCC value, the input (i.e., predicted score Q) should be replaced by Q_f .

An excellent quality assessment method should always have higher PLCC and SROCC values and a lower RMSE value. To obtain a robust criterion, we employ 17-fold cross-validation to test the proposed BMEFIQA method. Specifically, the MEF image database is divided into 17 subsets, 16 of which are used for RF model training, and the remaining are used for testing. Each subset contains MEF images belonging to the same scene. Finally, we report the average criteria values (i.e., PLCC, SROCC, and RMSE) across all fold trails when the train/test cycles are over for all the scenes.

4.2. Performance Comparison

To verify the superiority of the proposed method, two types of the NR-based method, which has demonstrated effectiveness for other images, are used to make the performance comparison. The first type contains ten methods, which are designed for ordinary images, dubbed as 2D NRIQA, namely, DIIVINE [5], BLIND-II [6], BRISQUE [7], CurveletQA [8], GradLog [9], ContrastQA [10], GLBP [11], OG [12], NIQMC [13], and SCORER [14]. The second type contains three methods designed for tone-mapped images, dubbed as TM-NRIQA, namely, BTMQI [17], HIGRADE-1 [18], and HIGRADE-2 [18]. All the comparison

methods are learning-based ones following the 17-fold cross-validation to obtain their performance for MEF images, which is consistent with the proposed method. To ensure that the results are not biased, the performances of the comparison methods are all obtained by running the original released codes from the authors.

The overall performance comparison results are presented in Table 2. To highlight the best performance, the PLCC, SROCC, and RMSE values are shown in bold. Obviously, the proposed BMEFIQA method reveals its better performance compared to the other comparison methods, which verifies its validity. We can draw some observations from Table 2. On the one hand, most 2D NRIQA methods exhibit poor performance, except for GradLog; the worst performance indicators for PLCC and SROCC are only 0.163 and 0.113, respectively. The reasons for this phenomenon can be roughly summarized as follows: First, the distortion of an MEF image presents differently from an ordinary 2D image, such as difficult semantic understanding and color deterioration when suffering from over-exposure and under-exposure. Second, the GradLog method combines gradient amplitude maps with the Laplace of Gaussian responses to sense structural information, which can describe the structural distortion of MEF images well. Hence, its performance outperforms other comparison 2D NRIQA methods. On the other hand, compared with 2D NRIQA methods, TM-NRIQA methods achieve better performance. This can be attributed to the fact that the distortion of tone-mapped images seems similar to that of MEF images. However, such results should also be improved to predict the quality of MEF images more accurately. The proposed BMEFIQA method considers the structure, naturalness, and colorfulness of an MEF image from three aspects. Its comprehensive consideration advocates that it obtains superior performance to the competing methods.

Table 2. The overall performance comparison results.

| Metrics | PLCC | SROCC | RMSE |
|------------|--------------|--------------|--------------|
| DIIVINE | 0.491 | 0.403 | 1.452 |
| BLINDS-II | 0.534 | 0.346 | 1.409 |
| BRISQUE | 0.414 | 0.380 | 1.517 |
| CurveletQA | 0.371 | 0.337 | 1.548 |
| GradLog | 0.631 | 0.567 | 1.293 |
| ContrastQA | 0.458 | 0.412 | 1.482 |
| GWH-GLBP | 0.163 | 0.113 | 1.645 |
| OG | 0.523 | 0.525 | 1.421 |
| NIQMC | 0.519 | 0.404 | 1.425 |
| SCORER | 0.481 | 0.494 | 1.461 |
| BTMQI | 0.452 | 0.343 | 1.487 |
| HIGRADE-1 | 0.561 | 0.566 | 1.380 |
| HIGRADE-2 | 0.585 | 0.583 | 1.352 |
| Proposed | 0.694 | 0.673 | 1.200 |

In order to compare the performance of the different methods more intuitively, Figure 5 shows the scatter plots for the objective predicted scores and MOS values of fourteen NR-IQA methods. They are conducted using the whole MEF database, and the scatter points are fitted by the logistic function. The closer these scatter points are to the fitting line, the better the performance. In Figure 5, it can be seen that the predicted scores of the proposed method show a higher correlation with the MOS values than the other NR-IQA methods.

4.3. Impacts of Different Features and Block Sizes

Based on the above performance comparison, the proposed BMEFIQA method is demonstrated to have a good-quality prediction capability. However, for the individual features, their role in the overall method remains ambiguous. To realize their purpose, the individual features S_F , N_F , and C_F and their different combinations are used to train the corresponding regression models to predict the quality of the MEF images separately. As shown in Table 3, the obtained regression models are written as Model- t , $t \in \{1, 2, \dots, 7\}$,

which indicates the number of individual features and their possible combinations. Their respective performance results are also given in Table 3, and the best performance results are shown in bold. The following can be observed: First, among the individual features, compared with the naturalness features N_F and color features C_F , the structural features S_F show a relatively strong performance. For the overall performance, S_F contributes the most. Second, among the different combinations, the combination of S_F and C_F produces the best performance. Although the performance of C_F itself is relatively weak, when it is combined with N_F , its performance is improved, indicating that it provides a good auxiliary effect. Third, when all S_F , N_F , and C_F features are incorporated together, the best performance can be obtained. Therefore, we believe that S_F , N_F , and C_F features are complementary to each other.

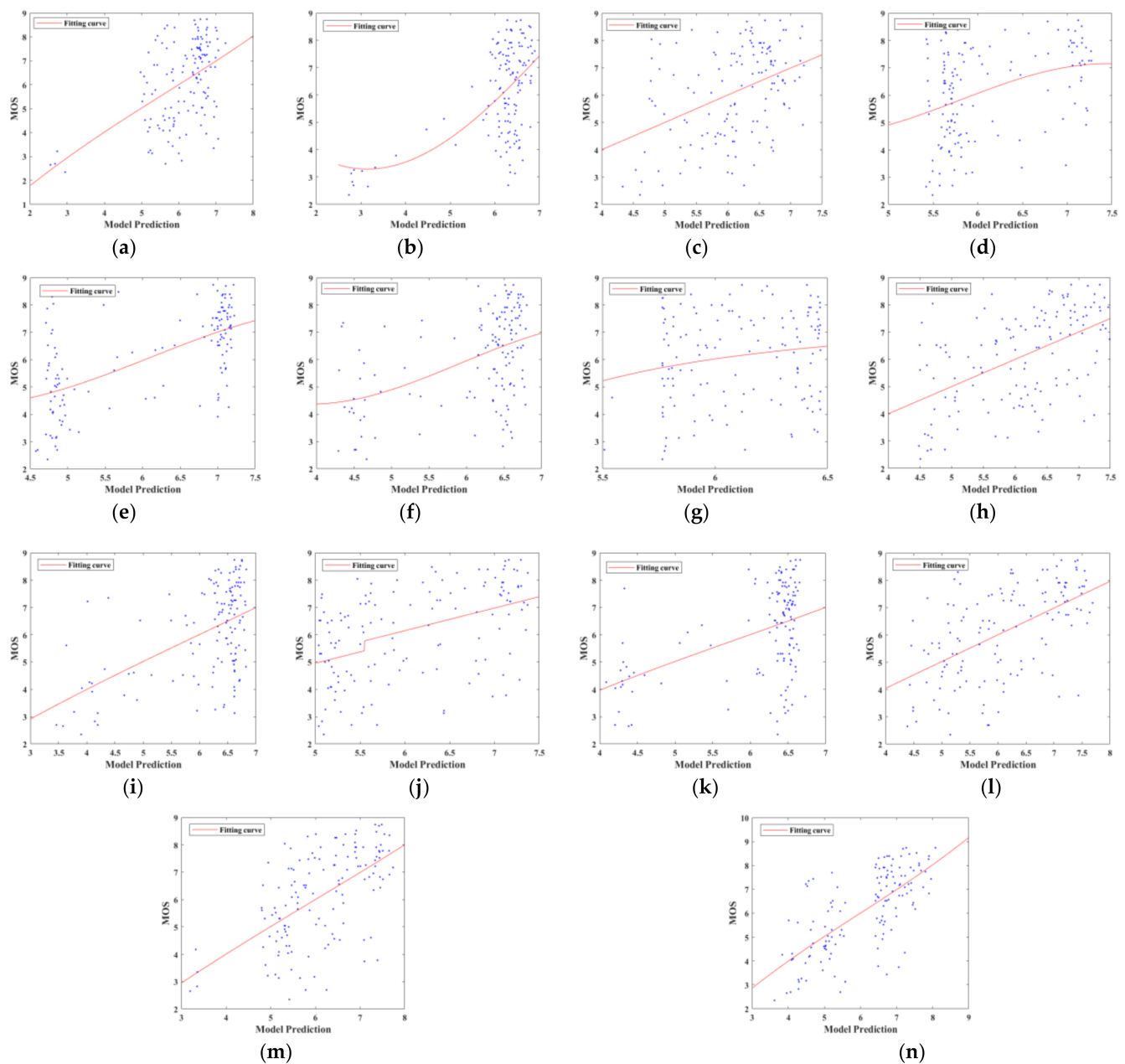


Figure 5. Scatter plots for the objective predicted scores and MOS values of sixteen NR-IQA methods. (a) DIIVINE [5]; (b) BLINDS_II [6]; (c) BRISQUE [7]; (d) CurveletQA [8]; (e) GradLog [9]; (f) ContrastQA [10]; (g) GWH-GLBP [11]; (h) OG [12]; (i) NIQMC [13]; (j) SCORER [14]; (k) BT-MQI [17]; (l) HIGRADE-1 [18]; (m) HIGRADE-2 [18]; (n) Proposed.

Table 3. Performance comparison of different feature combinations.

| Models | S _F | N _F | C _F | PLCC | SROCC | RMSE |
|---------|----------------|----------------|----------------|--------------|--------------|--------------|
| Model-1 | ✓ | × | × | 0.646 | 0.506 | 1.272 |
| Model-2 | × | ✓ | × | 0.557 | 0.457 | 1.385 |
| Model-3 | × | × | ✓ | 0.389 | 0.416 | 1.535 |
| Model-4 | ✓ | ✓ | × | 0.651 | 0.520 | 1.263 |
| Model-5 | ✓ | × | ✓ | 0.658 | 0.640 | 1.255 |
| Model-6 | × | ✓ | ✓ | 0.604 | 0.599 | 1.328 |
| Model-7 | ✓ | ✓ | ✓ | 0.694 | 0.673 | 1.200 |

In the block entropy calculation process, as different block sizes may affect the quality prediction performance, the block size should be determined to obtain the optimal performance. In the experiment, performances are tested under different block sizes, as shown in Table 4. Specifically, the block size varies from 8 to 64, and the corresponding PLCC, SROCC, and RMSE values are given. Additionally, the run time under different block sizes is also developed. From the results, we can observe that the PLCC and SROCC values and the run time decrease with an increase in block size. Under comprehensive consideration, the block size is set to 8.

Table 4. Performance comparison with different block sizes and the corresponding run time.

| Block Size | PLCC | SROCC | RMSE | Time (s) |
|------------|-------|-------|-------|----------|
| 8 × 8 | 0.694 | 0.673 | 1.200 | 1.3383 |
| 16 × 16 | 0.677 | 0.655 | 1.227 | 0.4744 |
| 32 × 32 | 0.683 | 0.658 | 1.218 | 0.2857 |
| 64 × 64 | 0.688 | 0.659 | 1.210 | 0.2347 |

4.4. Run Time

In a practical application, in addition to performance, efficiency is also important for the methods. Therefore, the mean implementation times for all MEF images in the MEF database of the proposed method and the other competing methods are reported in Table 5. All the original codes are implemented on a Windows 10 1.80 GHz Intel Core i7-8565U CPU with 8GB RAM and NVIDIA MX 150, using MATLAB 2017b. In Table 5, it can be seen that the proposed BMEFIQA method has a moderate execution time.

Table 5. The results of execution time by different methods.

| Methods | DIIVINE | BLINDS_II | BRISQUE | CurveletQA |
|----------|-----------|------------|----------|------------|
| Time (s) | 6.6024 | 14.7361 | 0.0414 | 2.5141 |
| Methods | GradLog | ContrastQA | GWH-GLBP | OG |
| Time (s) | 0.0343 | 0.0259 | 0.0576 | 0.0314 |
| Methods | NIQMC | SCORER | BTMQI | HIGRADE-1 |
| Time (s) | 1.9241 | 0.5878 | 0.0758 | 0.2602 |
| Methods | HIGRADE-2 | Proposed | | |
| Time (s) | 1.9040 | 1.3383 | | |

4.5. Discussion

In this study, a novel NR-IQA method that can handle the quality prediction task for MEF images is proposed. Specifically, the proposed BMEFIQA method considers the structure, naturalness, and colorfulness of MEF images, and it extracts the corresponding features of these three aspects. Through the performance comparison in the previous section, the proposed method shows its potential in predicting the quality of MEF im-

ages. It can be used in the application terminal to monitor quality with no reference information provided.

Although the proposed BMEFIQA method considers some image attributes, it is limited for kaleidoscopic images in real cases. They may be influenced by some special situations, which cannot be foreseen. As a result, regarding the performance of the proposed method, there is still room for improvement. For instance, the human attention mechanism will affect the perception of MEF image quality, and humans will have different artistic biases toward MEF images in different scenes. Deep learning is well known for its ability to automatically learn the features of some images. In future work, deep learning-based methods or more handcrafted features with the attention mechanism can be incorporated to improve the performance of the BMEFIQA method. Moreover, the MEF image database should be expanded for data training and testing to increase the robustness and practicability of the method.

5. Conclusions

This paper proposes a blind quality assessment method for multi-exposure fusion (MEF) images based on structure, naturalness, and colorfulness analyses of MEF images, named BMEFIQA. For the structure analysis, exposure map calculation is implemented to weight the gradient similarities, statistical modeling is used in the gradient domain, and entropy calculation is performed to characterize the loss of structural information. For the naturalness analysis, statistical modeling in the spatial domain combines the statistics of moment features and entropy, which are built to characterize scene naturalness. For colorfulness analysis, opponent color space is used to calculate contrast energy and build the statistical model. Finally, random forest is used to fuse the extracted features into prediction quality. Experiments on the public MEF image database demonstrate the superiority of the proposed BMEFIQA method.

Author Contributions: Conceptualization, H.L. and C.Z.; methodology, C.Z. and Z.H.; validation, J.S., H.L. and C.Z.; formal analysis, Z.H. and Y.M.; data curation, J.S. and Y.M.; writing—original draft preparation, J.S.; writing—review and editing, Z.H. and C.Z.; funding acquisition, H.L. and C.Z. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by National Natural Science Foundation of China, grant number 62172242; the Natural Science Foundation of Ningbo, grant number 202003N4155; and the Foundation of Zhejiang Province Education Department, grant number Y202146540.

Institutional Review Board Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Xu, H.; Ma, J.; Zhang, X.P. MEF-GAN: Multi-exposure image fusion via generative adversarial networks. *IEEE Trans. Image Process.* **2020**, *29*, 7203–7216. [\[CrossRef\]](#)
2. Luo, T.; Jiang, G.; Yu, M.; Xu, H.; Gao, W. Robust high dynamic range color image watermarking method based on feature map extraction. *Signal Process.* **2019**, *155*, 83–95. [\[CrossRef\]](#)
3. Qi, Y.; Zhou, S.; Zhang, Z.; Luo, S.; Lin, X.; Wang, L.; Qiang, B. Deep unsupervised learning based on color un-referenced loss functions for multi-exposure image fusion. *Inf. Fusion* **2021**, *66*, 18–39. [\[CrossRef\]](#)
4. Wang, Z.; Bovik, A.C. Reduced and no reference visual quality assessment. *IEEE Signal Process. Mag.* **2011**, *29*, 29–40. [\[CrossRef\]](#)
5. Moorthy, A.K.; Bovik, A.C. Blind image quality assessment: From natural scene statistics to perceptual quality. *IEEE Trans. Image Process.* **2011**, *20*, 3350–3364. [\[CrossRef\]](#) [\[PubMed\]](#)
6. Saad, M.A.; Bovik, A.C.; Charrier, C. Blind image quality assessment: A natural scene statistics approach in the DCT domain. *IEEE Trans. Image Process.* **2012**, *21*, 3339–3352. [\[CrossRef\]](#) [\[PubMed\]](#)
7. Mittal, A.; Moorthy, A.K.; Bovik, A.C. No-reference image quality assessment in the spatial domain. *IEEE Trans. Image Process.* **2012**, *21*, 4695–4708. [\[CrossRef\]](#)
8. Liu, L.; Dong, H.; Huang, H.; Bovik, A.C. No-reference image quality assessment in curvelet domain. *Signal Process. Image Commun.* **2014**, *29*, 494–505. [\[CrossRef\]](#)

9. Xue, W.; Mou, X.; Zhang, L.; Bovik, A.C.; Feng, X. Blind image quality assessment using joint statistics of gradient magnitude and Laplacian features. *IEEE Trans. Image Process.* **2014**, *23*, 4850–4862. [CrossRef]
10. Fang, Y.; Ma, K.; Wang, Z.; Lin, W.; Fang, Z.; Zhai, G. No-reference quality assessment of contrast-distorted images based on natural scene statistics. *IEEE Signal Process. Lett.* **2014**, *22*, 838–842. [CrossRef]
11. Li, Q.; Lin, W.; Fang, Y. No-reference quality assessment for multiply-distorted images in gradient domain. *IEEE Signal Process. Lett.* **2016**, *23*, 541–545. [CrossRef]
12. Liu, L.; Hua, Y.; Zhao, Q.; Huang, H.; Bovik, A.C. Blind image quality assessment by relative gradient statistics and adaboosting neural network. *Signal Process. Image Commun.* **2016**, *40*, 1–15. [CrossRef]
13. Gu, K.; Lin, W.; Zhai, G.; Yang, X.; Zhang, W.; Chen, C.W. No-reference quality metric of contrast-distorted images based on information maximization. *IEEE Trans. Cybern.* **2017**, *47*, 4559–4565. [CrossRef]
14. Oszust, M. Local feature descriptor and derivative filters for blind image quality assessment. *IEEE Signal Process. Lett.* **2019**, *26*, 322–326. [CrossRef]
15. Zhang, L.; Zhang, L.; Bovik, A.C. A Feature-Enriched Completely Blind Image Quality Evaluator. *IEEE Trans. Image Process.* **2015**, *24*, 2579–2591. [CrossRef]
16. Xu, J.; Ye, P.; Li, Q.; Du, H.; Liu, Y.; David, D. Blind Image Quality Assessment Based on High Order Statistics Aggregation. *IEEE Trans. Image Process.* **2016**, *25*, 4444–4457. [CrossRef]
17. Gu, K.; Wang, S.; Zhai, G.; Ma, S.; Yang, X.; Lin, W.; Zhang, W.; Gao, W. Blind quality assessment of tone-mapped images via analysis of information, naturalness, and structure. *IEEE Trans. Multimed.* **2016**, *18*, 432–443. [CrossRef]
18. Kundu, D.; Ghadiyaram, D.; Bovik, A.C.; Evans, B.L. No-reference quality assessment of tone-mapped HDR pictures. *IEEE Trans. Image Process.* **2017**, *26*, 2957–2971. [CrossRef]
19. Zheng, Y.; Essock, E.A.; Hansen, B.C.; Haun, A.M. A new metric based on extended spatial frequency and its application to DWT based fusion algorithms. *Inf. Fusion.* **2007**, *8*, 177–192. [CrossRef]
20. Ma, K.; Zeng, K.; Wang, Z. Perceptual quality assessment for multi-exposure image fusion. *IEEE Trans. Image Process.* **2015**, *24*, 3345–3356. [CrossRef]
21. Xing, L.; Zeng, H.; Chen, J.; Zhu, J.; Cai, C.; Ma, K. Multi-exposure image fusion quality assessment using contrast information. In Proceedings of the International Symposium on Intelligent Signal Processing and Communication Systems, Xiamen, China, 6–9 November 2017.
22. Fang, Y.; Zeng, Y.; Zhu, H.; Zhai, G. Image quality assessment of image fusion for both static and dynamic scenes. In Proceedings of the IEEE International Conference on Multimedia and Expo (ICME), Shanghai, China, 8–12 July 2019.
23. Deng, C.W.; Li, Z.; Wang, S.G.; Liu, X.; Dai, J.H. Saturation-based quality assessment for colorful multi-exposure image fusion. *Int. J. Adv. Robot. Syst.* **2017**, *14*, 1–15. [CrossRef]
24. Martinez, J.; Pistonesi, S.; Maciel, M.C.; Flesia, A.G. Multiscale fidelity measure for image fusion quality assessment. *Inf. Fusion* **2019**, *50*, 197–211. [CrossRef]
25. He, Z.; Song, Y.; Zhong, C.; Li, L. Curvature and Entropy Statistics-Based Blind Multi-Exposure Fusion Image Quality Assessment. *Symmetry* **2021**, *13*, 1446. [CrossRef]
26. Peng, Y.; Feng, B.; Yan, Y.; Gao, X. Research on multi-exposure image fusion algorithm based on detail enhancement. In Proceedings of the International Conference on Mechanical Engineering, Guangzhou, China, 14 October 2021.
27. Mertens, T.; Kautz, J.; Van Reeth, F. Exposure fusion: A simple and practical alternative to high dynamic range photography. *Comput. Graph. Forum.* **2009**, *28*, 161–171. [CrossRef]
28. Raman, S.; Chaudhuri, S. Bilateral filter based compositing for variable exposure photography. In Proceedings of the Eurographics (Short Papers), Munich, Germany, 30 March–3 April 2009; pp. 1–3.
29. Gu, B.; Li, W.; Wong, J.; Zhu, M.; Wang, M. Gradient field multi-exposure images fusion for high dynamic range image visualization. *J. Vis. Commun. Image Represent.* **2012**, *23*, 604–610. [CrossRef]
30. Wang, Z.; Bovik, A.C.; Sheikh, H.R.; Simoncelli, E.P. Image quality assessment: From error visibility to structural similarity. *IEEE Trans. Image Process.* **2004**, *13*, 600–612. [CrossRef]
31. Lee, D.H.; Yoon, Y.J.; Kang, S.; Ko, S.J. Correction of the overexposed region in digital color image. *IEEE Trans. Consum. Electron.* **2014**, *60*, 173–178. [CrossRef]
32. Gonzalez, R.C.; Woods, R.E. Image enhancement in the spatial domain. *Digit. Image Process.* **2002**, *2*, 75–147.
33. Ennis, R.J.; Zaidi, Q. Geometrical structure of perceptual color space: Mental representations and adaptation invariance. *J. Vis.* **2019**, *19*, 1. [CrossRef]
34. Choi, L.K.; You, J.; Bovik, A.C. Referenceless prediction of perceptual fog density and perceptual image defogging. *IEEE Trans. Image Process.* **2015**, *24*, 3888–3901. [CrossRef]
35. Multi-exposure Fusion Image Database. Available online: <http://ivc.uwaterloo.ca/database/MEF/MEFDatabase.php> (accessed on 11 July 2015).
36. Li, Z.; Zheng, J.; Rahardja, S. Detail-enhanced exposure fusion. *IEEE Trans. Image Process.* **2012**, *21*, 4672–4676.
37. Li, S.; Kang, X. Fast multi-exposure image fusion with median filter and recursive filter. *IEEE Trans. Consum. Electron.* **2012**, *58*, 626–632. [CrossRef]

-
38. Li, S.; Kang, X.; Hu, J. Image fusion with guided filtering. *IEEE Trans. Image Process.* **2013**, *22*, 2864–2875.
 39. Antkowiak, J.; Baina, T.J. *Final Report from the Video Quality Experts Group on the Validation of Objective Models of Video Quality Assessment*; ITU-T Standards Contributions COM: Geneva, Switzerland, 2000.