

No-reference quality metric for high dynamic range imaging system based on curvature analysis in tensor domain

GANGYI JIANG^{1,2}, JIAOWEN YU¹, MEI YU^{1,2*}, YANG SONG¹, ZONGJU PENG¹, FEN CHEN¹

¹Faculty of Information Science and Engineering, Ningbo University, Ningbo, China

²National Key Lab of Software New Technology, Nanjing University, Nanjing, China

High dynamic range imaging systems can offer a more complete representation of scene, aiming to capture all brightness information of a visible range of scene, even in extreme lighting conditions. This paper proposes a no-reference quality metric for high dynamic range image (HDRI), in which a combination of tensor decomposition and curvature analysis is used to construct an efficient feature set that is sensitive to degradation levels of patches in HDRIs. Tensor decomposition maintains the majority of color information of an HDRI, and the geometric structure information of the HDRI is then extracted by a curvature analysis. A quality-related label feature matrix is subsequently defined and obtained by using a feature set and a sparse dictionary with quality-related labels. Then, the multi-dimensional local feature set of the HDRI is determined from the quality-related label feature matrix. Finally, the local feature set and other global feature set are pooled to predict the quality of the HDRI. The prediction performance of the proposed metric is verified by three public test databases, and the experimental results indicate that both its Pearson linear correlation coefficient and Spearman rank-order correlation coefficient are better than those of other no-reference metrics. The proposed metric produces statistically better assessment results, implying a higher consistency with human visual perception.

Keywords: high dynamic range image, no-reference quality metric, tensor domain, curvature analysis, quality-related label feature matrix.

1. Introduction

High dynamic range (HDR) imaging affords true-to-life visual sensations and represents a step-change in viewing technology. The corresponding technologies are developed and applied to consumer electronics [1]. In high dynamic range image (HDRI) systems, an objective quality assessment of HDRIs is essential for various optimization processes. HDRIs fundamentally differ from traditional low dynamic range images (LDRIs) in that the pixels of the former are related (proportional) to the physical luminance of the scene [2]. HDRIs can thus provide richer colors and more realistic scene details in both under-exposed and over-exposed areas. HANHART *et al.* [3] tested 35 objective image quality metrics (IQMs) comprising 33 metrics for LDRIs and two metrics for HDRIs. They found that the existing LDR image quality metrics (LDR-IQMs) could not be directly used to assess the quality of an HDRI. Rather, the initial application of

a human visual perception (HVP) model before the usage of an LDR-IQM was required. Applicable HVP models include those of the physical luminance of the scene (Photometric) [4], perceptually uniform (PU) coding [5] and the logarithmic (Log) coded pixel values. However, the improved performance of these metrics after the usage of the HVP models varies considerably.

Some HDR image quality metrics (HDR-IQMs) have been presented, among which the HDR visual detection predictor (HDR-VDP) [6] is a representative example. MANTIUK *et al.* constructed a visual model [6] taking into consideration the aspects of high contrast vision, such as the scattering of light, the nonlinear response to light in the full luminance range, and local adaptation. To achieve intuitive prediction, MANTIUK *et al.* also developed a second version of HDR-VDP, denoted as HDR-VDP-2 [7], by adding a frequency-based pooling strategy. By solving the problem of the optimal pooling weights in the pooling strategy, NARWARIA *et al.* further developed HDR-VDP-2.2 [8], which affords more accurate quality prediction. AYDIN *et al.* proposed the dynamic range independent metric (DRIM) [9] for the quality assessment of HDRIs. However, the measurement results when using this metric are represented by three distortion maps, which are difficult to interpret owing to the lack of pooling of the different values. On the basis of a spatio-temporal analysis related to the fixation behavior of the human eye when viewing videos, NARWARIA *et al.* developed an HDR video quality metric (HDR-VQM) [10], which is also applicable to HDRIs. However, all of the above-mentioned metrics only consider the luminance information; and they are also full-reference HDR-IQMs, requiring the original HDRIs for the assessment of distorted HDRIs [6–10].

In fact, the original HDRI is usually unavailable for use as the reference image in many HDRI applications, and this limits the application of the full-reference HDR-IQMs. It is necessary for HDRI systems to design no-reference HDR-IQMs and to benchmark their performance against the subjective ground truth. Generally, no-reference LDR-IQMs can be classified into two categories [11]. The first category is based on hand-crafted image features such as the natural scene statistics (NSSs), information entropies and image filtering responses. An example is GM-LOG by XUE *et al.* [12], which extracts competitive histogram-based features from the joint statistics of the image gradient magnitudes and the Laplacian of Gaussian responses. The second category of LDR-IQMs comprises those based on training-related features such as the features obtained from various neural networks. WU *et al.* [13] proposed a data-driven no-reference IQM based on a quality-aware deep neural network, capable of automatic extraction of high-level features with parameter adjustment. Clearly, it has been also demonstrated that the no-reference LDR-IQMs are not suitable for the direct assessment of HDRIs.

In this study, we will propose a no-reference quality metric for an HDRI (NQM-HDRI). The proposed NQM-HDRI is mainly based on a curvature analysis in tensor domain. We have considered the fact that the higher peak brightness and richer colors of an HDRI are not accurately measured by using the features of most LDR-IQMs, resulting in a low-quality prediction accuracy. Hence, a combination of tensor decomposition and a curvature analysis in tensor domain is used to extract an efficient feature set that represents

the degradation levels of the patches in the HDRI. Tensor decomposition maintains most of the color information in the HDRI, and the geometric structure information of the HDRI can also be extracted with a curvature analysis. A quality-related label feature matrix (Q-LFM) is subsequently constructed using the feature sets of the patches and a sparse dictionary with quality-related labels. Then, the local feature set of the HDRI is extracted from the Q-LFM. Finally, the quality of the HDRI is predicted by combining local features with other global features. The main contributions of this paper are as follows:

- 1) It describes how information about the color and geometric structure of an HDRI can be used to obtain an efficient feature set through a curvature analysis in tensor domain, for the effective representation of the degradation level of patches of the HDRI.
- 2) It describes the use of the differing distortion levels of different regions of an HDRI to define and characterize the Q-LFM through different quality-related labels and obtains the local feature set of the HDRI from the Q-LFM.

The remainder of this paper is organized as follows. Section 2 describes the details of the proposed NQM-HDRI. Section 3 uses three public test databases [4, 14, 15] to compare the proposed metric with other state-of-the-art metrics. Section 4 concludes the paper.

2. No-reference quality metric for high dynamic range images

A color HDRI has three color channels, namely R, G, and B, and their conversion into a luminance or gray domain results in the loss of color information. However, tensor

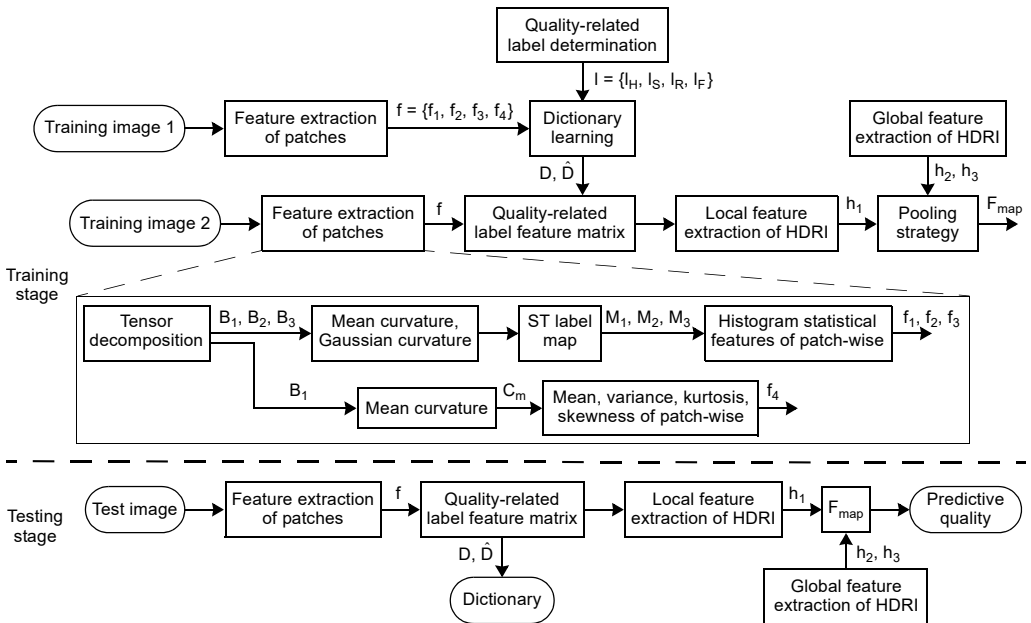


Fig. 1. Framework of the proposed NQM-HDRI.

decomposition can be used to maintain most of the information of the color HDRI [16]. Hence, signal processing approaches based on the tensor domain prevent the loss of color information, and can thus be used for HDRI feature extraction. Consequently, the first step in the implementation of the proposed IQM is the use of a curvature analysis in tensor domain to extract a feature set that is sensitive to the degradation levels of the patches of an HDRI. Afterwards, the NQM-HDRI is proposed as depicted in Fig. 1, which depicts the two implementation stages of the NQM-HDRI, namely the training and test stages. The training stage comprises four parts: *i*) extraction of the feature set f by a curvature analysis in tensor domain, *ii*) defining the Q-LFM using a sparse dictionary based on quality-related labels, *iii*) obtaining the multi-dimensional local feature set h_1 of the HDRI from the Q-LFM and subsequently obtaining the global feature sets h_2 and h_3 , and *iv*) pooling the local and global feature sets using a random forest. At the test stage, the proposed NQM-HDRI is performed. The process flow in Fig. 1 is described in detail in the following sub-sections.

2.1. Extraction of a feature set by a curvature analysis in tensor domain

Generally, the features of most LDR-IQMs are unsuitable for HDRIs because they are mostly extracted directly from the luminance or gray domain resulting in the loss of color information. Here, the tensor domain is used owing to its superior feature extraction from HDRIs. As noted earlier, the proposed NQM-HDRI uses a curvature analysis in tensor domain to extract the feature set that represents the distortion levels of the patches of an HDRI.

2.1.1. Tensor decomposition

Let $\mathcal{A} \in \mathbb{R}^{I_1 \times I_2 \times \dots \times I_N}$ denote an N -order tensor corresponding to an N -dimensional data array. Tensor decomposition is a high order extension of the singular value decomposition (SVD) of a matrix. There are some basic forms of tensor decomposition such as CANDECOMP/PARAFAC decomposition and Tucker decomposition. Here, Tucker decomposition [16] is used in this paper. By this approach, \mathcal{A} can be decomposed into an n -mode product of a core tensor and N matrices, as follows:

$$\mathcal{A} \approx [\mathcal{B}; U^{(1)}, U^{(2)}, \dots, U^{(N)}] = \mathcal{B} \times U^{(1)} \times \dots \times U^{(N)} \quad (1)$$

where \mathcal{B} is a core tensor with the same size as \mathcal{A} ; $U^{(n)} \in \mathbb{R}^{I_1 \times I_2}$ ($n = 1, 2, \dots, N$) is the factor matrix in each mode that satisfies all orthogonality conditions. \mathcal{B} can be expressed in terms of \mathcal{A} , as follows:

$$\mathcal{B} \approx \mathcal{A} \times (U^{(1)})^T \times \dots \times (U^{(N)})^T \quad (2)$$

In the feature extraction, tensor decomposition of an HDRI is first performed. The HDRI with the size of $W \times L \times 3$ contains three color channels, namely, R, G and B, defined by $\mathcal{A} = (A_R, A_G, A_B)$. After Tucker decomposition, \mathcal{B} can be obtained by

Eq. (2). A sub-tensor \mathcal{D} (i.e., $\mathcal{D} = (D_1, D_2, D_3)$) extracted from \mathcal{B} satisfies orderliness, i.e., $\|D_1\| \geq \|D_2\| \geq \dots \geq \|D_3\| \geq 0$. Let B_i ($i = 1, 2, 3$) denote matrix with the size of $W \times L$ in \mathcal{B} (i.e., $\mathcal{B} = (B_1, B_2, B_3)$), then the following relationship would be satisfied:

$$\|B_1\|^2 \geq \|B_2\|^2 \geq \|B_3\|^2 \geq 0 \tag{3}$$

where B_1, B_2 and B_3 are sub-bands I, II, and III, respectively. Figure 2 shows the sub-bands extracted from the reference and distorted HDRIs. Because of the difficulty of using an ordinary display for the HDRIs, the images in Fig. 2 are scaled by a certain factor. The corresponding energy responses are shown in Fig. 3, in which the different curves represent the energy responses of the column pixels of the different sub-bands. As can be observed from Fig. 3, sub-band I contains most of the energy of an HDRI. The blue curve, which corresponds to sub-band I, is obviously higher than the two other curves corresponding to the two other sub-bands.

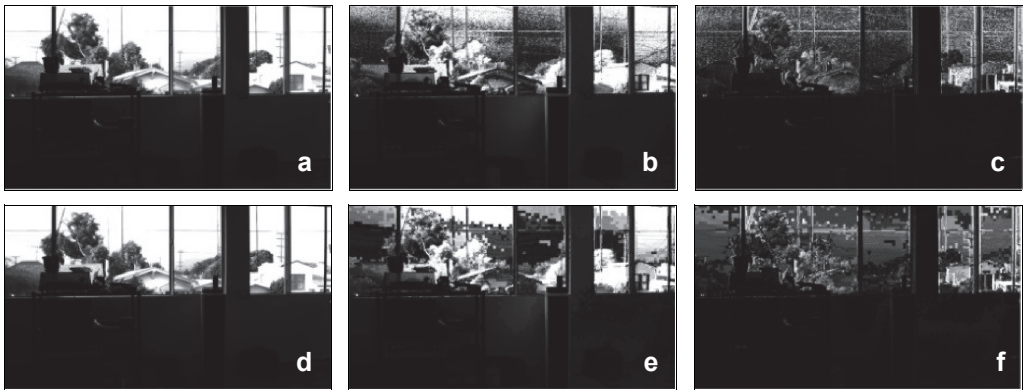


Fig. 2. Sub-bands I, II, and III of the reference (a–c) and distorted (d–f) HDRIs.

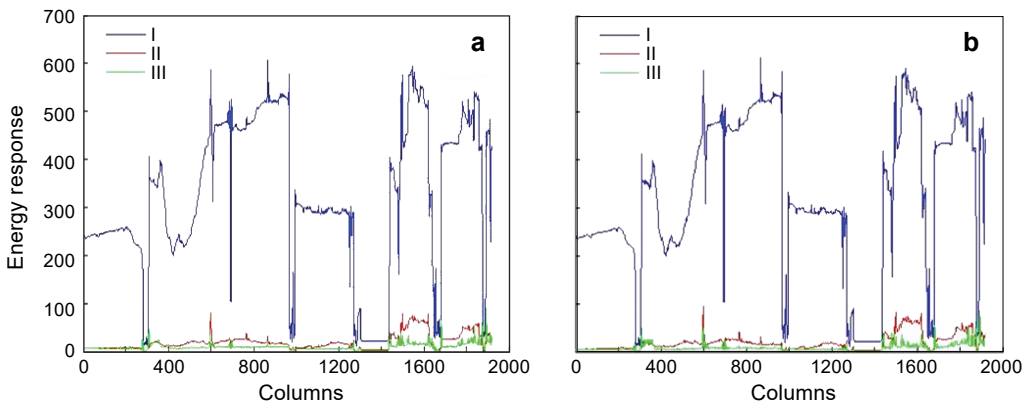


Fig. 3. Energy responses of the column pixels of the three sub-bands of the reference (a) and distorted (b) HDRIs.

2.1.2. Determination of the features of patches by a curvature analysis

An image can be considered to comprise surfaces with various convex and concave structures. The pixels of an image can be classified into different types on the basis of their geometric structure. Such a type of characteristic is called a surface type (ST) [17]. The three sub-bands of an HDRI, $B = \{B_i(x, y) | i = 1, 2, 3\}$, also have a characteristic, which can be defined by the mean curvature and Gaussian curvature. The mean curvature C_m and Gaussian curvature C_g [17] of each sub-band can be respectively computed by

$$C_m = \frac{(1 + B_x^2)B_{yy} + (1 + B_y^2)B_{xx} - 2B_x B_y B_{xy}}{2(1 + B_x^2 + B_y^2)^{3/2}} \quad (4)$$

$$C_g = \frac{B_{xx}B_{yy} - B_{xy}^2}{1 + B_x^2 + B_y^2} \quad (5)$$

where B_x and B_y are the first-order partial derivatives, and B_{xx} , B_{xy} and B_{yy} are the second-order partial derivatives. Figure 4 presents C_m and C_g analysis results for the reference and distorted HDRIs. It is obvious from Fig. 4 that the mean curvature of each sub-band provides more details about the structure than the Gaussian curvature.

Table 1 presents the nine fundamental ST labels for the different curvature analyses of an HDRI in tensor domain. As indicated in Table 1, each pixel of a sub-band can be labeled with one of the nine ST labels on the basis of results of the analysis described above. An ST label map can thus be obtained for each sub-band, each field of which is defined by an integer from 1 to 9. The ST label maps of the sub-bands of the reference and distorted HDRIs are shown in Fig. 5. The maps can be observed to be highly discriminative regarding the details of the structure. Thus, the feature set of each patch can be extracted from the ST label maps. The ST label maps for each image are identified as M_1 , M_2 , and M_3 , corresponding to the HDRI's three sub-bands. Each ST label map is divided into patches with a size of $k_1 \times k_1$, and a histogram of the ST labels is obtained for each patch. Consequently, each patch has a nine-dimensional feature set, with a total of 27 (3×9) dimensional features extracted from the corresponding patches of the ST label maps of the three sub-bands of an image, denoted as $\{f_1, f_2, f_3\}$.

It is evident from Fig. 5 that the ST label maps are sensitive to HDRI's distortions. However, the features extracted from the ST label maps are not sufficient to describe the

T a b l e 1. Definitions of the ST labels.

Curvature analysis result	$C_m > 0$	$C_m = 0$	$C_m < 0$
$C_g > 0$	ST = 1	ST = 2	ST = 3
$C_g = 0$	ST = 4	ST = 5	ST = 6
$C_g < 0$	ST = 7	ST = 8	ST = 9

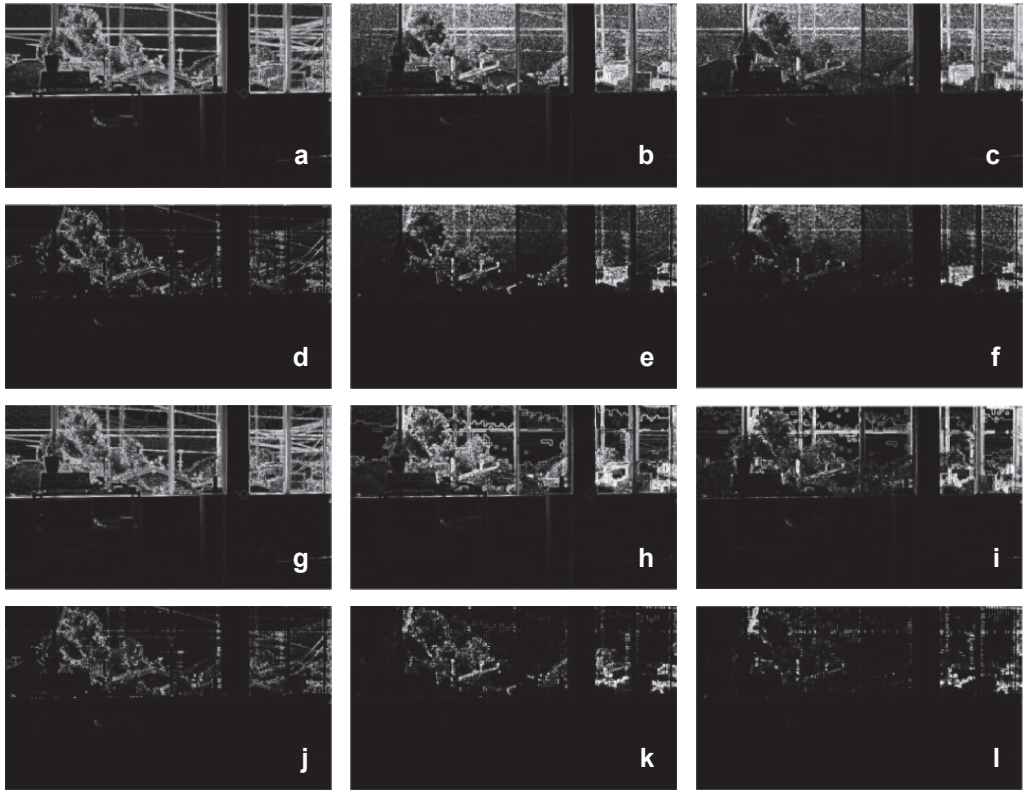


Fig. 4. Curvature analysis results in tensor domain of the mean curvature of the three sub-bands (I, II and III) of the reference HDRI (a–c), the Gaussian curvature of the three sub-bands of the reference HDRI (d–f), the mean curvature of the three sub-bands of the distorted HDRI (g–i), and the Gaussian curvature of the three sub-bands of the distorted HDRI (j–l).

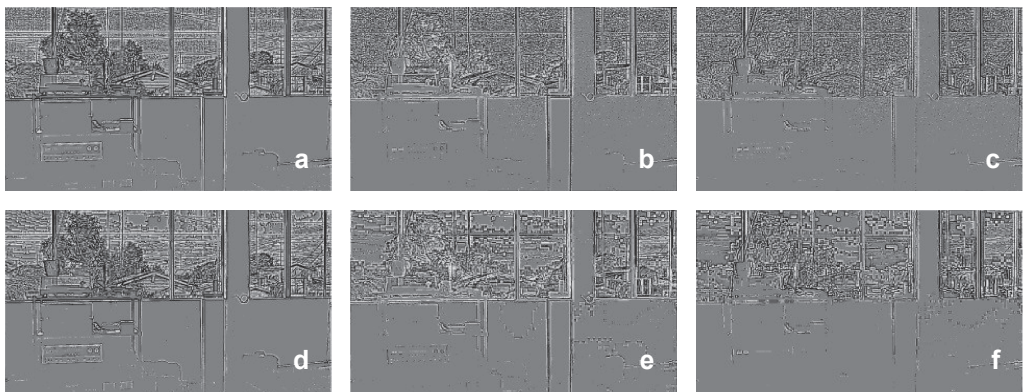


Fig. 5. ST label maps of the three sub-bands of the reference (a–c) and distorted (d–f) HDRIs.

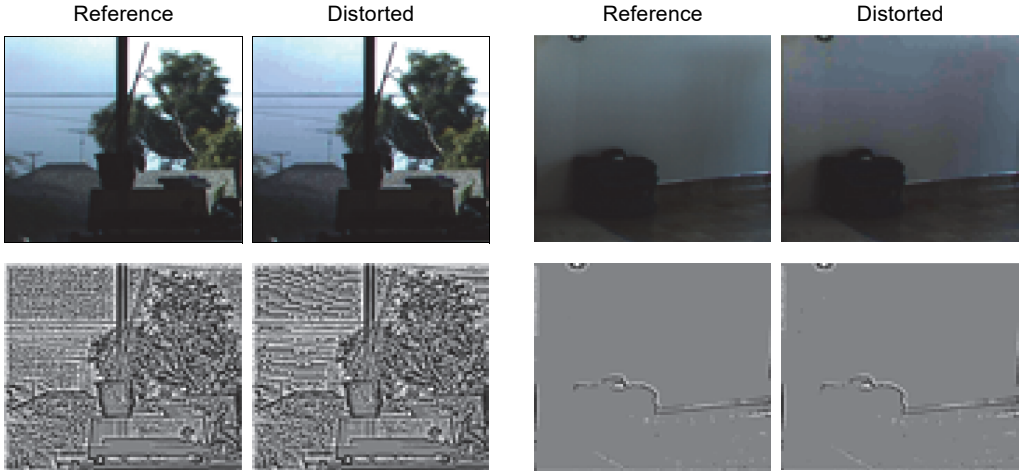


Fig. 6. Sample details extracted from the ST label maps of sub-band I.

geometric structure of the HDRI. To illustrate the necessity of complementary features, some structural details extracted from the ST label maps of sub-band I of the HDRI in Fig. 5 are shown in Fig. 6. From the first two images in Fig. 6, there are obvious differences between the ST label maps of the reference and distorted images. However, the latter two images in Fig. 6 show that the ST label maps cannot clearly reveal the apparent distortions in an image. Figure 6 shows that the same ST label may be assigned to pixels with different structural characteristics. It is therefore necessary to obtain other complementary features.

On the basis of the above analysis, it is obvious that sub-band I contains most of the energy of an HDRI, and that C_m of each sub-band contains more structural details than C_g . C_m of sub-band I can thus be considered to be more important than the other sub-band curvature analysis results. C_m of sub-band I is therefore divided into patches for the extraction of the feature set, including the mean, variance, kurtosis and skewness, which are denoted by f_4 . Overall, the curvature analysis in tensor domain is used to extract 31-dimensional features to constitute the feature set \mathbf{f} , where $\mathbf{f} = \{f_1, f_2, f_3, f_4\}$.

2.2. Construction of Q-LFM and local HDRI feature set

The next step after the feature extraction of patches is to obtain the Q-LFM by using a sparse representation based on quality-related labels. The acquisition of the quality-related labels is thus important to the proposed metric. They are obtained from four state-of-the-art full-reference IQMs, namely, HDR-VDP-2.2 [8], SSIM [18], RFSIM [19], and FSIM [20]. The luminance information of the HDRIs is employed for this purpose by dividing it into non-overlapping patches with a size of $k_1 \times k_1$. Each patch is then assessed by the above-mentioned four IQMs and the corresponding quality-related labels, respectively denoted by l_H , l_S , l_R and l_F , are obtained. A four-dimensional quality-related label vector denoted by $\mathbf{l} = \{l_H, l_S, l_R, l_F\}$ is thus constructed.

In the training process of a sparse dictionary, a large number of patches are randomly selected. For each patch i , 31-dimensional feature vectors \mathbf{f}_i and four-dimensional quality-related label vectors \mathbf{l}_i are obtained by the procedure described above. Sparse dictionary learning can be formulated by

$$\{D, \hat{D}, \mathbf{l}\} = \arg \min_{\{D, \hat{D}, \alpha\}} \sum_i \left\{ \|\mathbf{f}_i - D\alpha_i\|_F^2 + \|\mathbf{l}_i - \hat{D}\alpha_i\|_F^2 + \lambda \|\alpha_i\|_1 \right\} \quad (6)$$

where λ is a constant, and α_i denotes the sparse coefficients shared by \mathbf{f}_i over the dictionary D and \mathbf{l}_i over the dictionary \hat{D} . The quality-related labels are reconstructed using the sparse dictionary obtained by the K-SVD and OMP algorithms [21], and the corresponding formulas are

$$\alpha_i = \arg \min_{\alpha_i} \left\{ \|\mathbf{f}_i - D\alpha_i\|_F^2 + \lambda \|\alpha_i\|_1 \right\} \quad (7)$$

where $\hat{\mathbf{l}}_i = \hat{D}\alpha_i$ denotes the reconstructed quality-related label vectors.

A Q-LFM with a size of $m \times n \times 4$ is constructed by integrating the quality-related label vector of each patch of the HDRI. Here, $m = \lfloor W/k_1 \rfloor$, $n = \lfloor L/k_1 \rfloor$, and $\lfloor \cdot \rfloor$ denotes rounding down. The reconstructed quality-related label vectors are used to represent the distortion levels in the different regions of an HDRI. The Q-LFM is then divided into non-overlapping patches with a size of $k_2 \times k_2 \times 4$. Finally, the multi-dimensional local feature set of the HDRI, denoted by \mathbf{h}_i , is obtained as the mean values for the $k_2 \times k_2 \times 4$ patches.

2.3. Construction of a global HDRI features set

The above-mentioned HDRI feature set contains all of the local features from patches with a size of $k_2 \times k_2 \times 4$; thus, the global features of the HDRI are also modeled in the spatial domain, *i.e.*, the mean subtracted contrast normalized (MSCN) pixels and the σ -field of the HDRI [22]. The luminance pixels of the HDRI are preprocessed by mean subtraction and divisive normalization. Let $W \times L$ be the dimension of the luminance information I of the HDRI and $I(i, j)$ be the pixel value at the position (i, j) , $i \in \{1, 2, \dots, W\}$, $j \in \{1, 2, \dots, L\}$, then MSCN pixels are generated by

$$\hat{I}(i, j) = \frac{I(i, j) - \mu(i, j)}{\sigma(i, j) + 1} \quad (8)$$

where the local mean $\mu(i, j)$ and standard deviation $\sigma(i, j)$ are defined as

$$\mu(i, j) = \sum_{r=-R}^{r=R} \sum_{s=-S}^{s=S} w_{r,s} I(i+r, j+s) \quad (9)$$

$$\sigma(i, j) = \sqrt{\sum_{r=-R}^{r=R} \sum_{s=-S}^{s=S} w_{r,s} \left[I(i+r, j+s) - \mu(i, j) \right]^2} \quad (10)$$

where $w = \{w_{r,s} | r = -R, \dots, R, s = -S, \dots, S\}$ is a symmetric local convolution window centered at the (i, j) -th pixel. R and S determine the size of the local patch considered in the calculation of the mean and standard deviation. The MSCN pixels are modeled using a generalized Gaussian distribution (GGD) [22] and used in NQM-HDRI, denoted as h_2 . Although the MSCN coefficients are definitely more homogenous for the original images, the signs of the horizontally adjacent coefficients exhibit a regular structure, which is disturbed in the presence of distortion. This structure is modeled using the neighboring MSCN coefficients along the horizontal direction, denoted as $H(i, j)$

$$H(i, j) = \hat{I}(i, j)\hat{I}(i, j + 1) \quad (11)$$

where $\hat{I}(i, j)$ and $\hat{I}(i, j + 1)$ can be obtained by Eqs. (8)–(10). The very general asymmetric generalized Gaussian distribution (AGGD) model [22] is adopted to model the empirical distributions of pairwise products and used in NQM-HDRI, denoted h_3 .

2.4. Pooling strategy

In this paper, the combination of local and global feature sets is pooled by a random forest (RF). An RF is a versatile machine-learning algorithm that can be used for regression and classification tasks [23]. Many decision trees are generated by the algorithm, which classifies or selects on the basis of certain properties of the RF. The output of an RF is the highest number of votes in the category options or the mean of all of the votes in the case of a regression problem. Some researchers have demonstrated the high prediction accuracy of an RF, which is also not prone to over-fitting. In the present study, the inputs are the subjective assessment values and the feature sets h_i . The mapping function F_{map} is obtained by an RF, which can determine the mapping relationship between the features and the subjective assessment values.

In the test stages, the test HDRI is first divided into non-overlapping patches of size $k_1 \times k_1$, and the feature set f of each patch is extracted by the feature extraction procedure. Second, the sparse dictionary is used to reconstruct the quality-related label vectors $\hat{\mathbf{I}}$ of each patch, all of which can be used to construct an $m \times n \times 4$ Q-LFM. Third, the Q-LFM is divided into patches of size $k_2 \times k_2 \times 4$, and the local feature set h_1 of the test HDRI is obtained by the mean processing of each patch from the Q-LFM. The global feature sets h_2 and h_3 are also obtained. Finally, the predictive quality of the test HDRI is pooled by the mapping function F_{map} .

3. Experimental analysis

To demonstrate the performance of the proposed NQM-HDRI, images obtained from three databases are used to perform some test experiments, denoted as database1 [14], database2 [15], and database3 [4]. Database1 comprises 150 HDRIs of 10 original HDRIs with different scenes, including buildings, landscapes and characters, and computer graphics. Figure 7 shows the ten original HDRIs, which have an ultra-high resolution, are in the .hdr format, and have a size of $W \times L \times 3$, where $W = 1080$ and



Fig. 7. Ten original HDRIs of different scenes in database1 [14].

$L = 1920$. Database2 comprises 264 HDRIs of 24 original HDRIs of different scenes. The 24 original HDRIs have an ultra-high resolution, are in the .pfm format and have a size of $1080 \times 944 \times 3$. Database3 comprises 65 HDRIs of 15 original HDRIs of different scenes, which have the same resolution, format, and size as database1.

With regard to their performance criteria, three measurements [24] are applied to the IQMs, the Spearman rank-order correlation coefficient (SROCC), Pearson linear correlation coefficient (PLCC), and root mean square error (RMSE). The SROCC is generally used to measure the prediction monotonicity of an IQM and only considers the ranks of the data points, ignoring the relative distance between them. The PLCC is a measure of the linear dependence between two quantities after nonlinear regression, whereas the RMSE is used to measure the prediction accuracy. Two aspects are applied to the present experimental results comprising feature selection and comparison with other IQMs.

3.1. Comparative analysis of the feature selection of patches

The data presented in this sub-section are illustrated with database1 as examples. To generate a valid data set for the sparse training, 100 000 patches are randomly selected from 30 HDRIs in training images-1. The HDRIs are of different scenes and distortion levels. The size of the patches is 8×8 , *i.e.*, $k_1 = 8$. Using the sparse dictionary, the quality-related labels of the patches from an HDRI are reconstructed to obtain a Q-LFM with a size of $m \times n \times 4$, where $m = 135$ and $n = 240$. The multi-dimensional feature sets h_1 of the HDRIs are determined by applying the mean process to the Q-LFM patches, where they have a size of $32 \times 32 \times 4$, *i.e.*, $k_2 = 32$. Finally, the local feature set h_1 of the HDRI is extracted. Training images-2 is used to train the pooling strategy. In the pooling strategy training process, we randomly choose 80% of the residual HDRIs for training and the other 20% of the residual HDRIs for testing. To balance the effect of random data selection, the performance evaluation is performed 1 000 times, and the median performance is computed.

In this sub-section, three schemes are designed and used to evaluate the significance of feature selection. In each scheme, the methods for obtaining a sparse dictionary, the Q-LFM and mapping function F_{map} are the same, except the feature set of dictionary training. Scheme 1 involves a multi-dimensional feature set (9×3) based on the ST label maps in the RGB domain. Scheme 2 involves a feature set (9×3) transformed from

T a b l e 2. Results of a performance evaluation for feature selection.

Feature selection schemes	SROCC	PLCC	RMSE
Scheme 1	0.6098	0.6623	0.778
Scheme 2	0.7898	0.8085	0.613
Scheme 3	0.8839	0.9001	0.454

the RGB domain into tensor domain on the basis of scheme 1, denoted by $\{f_1, f_2, f_3\}$. scheme 3 involves the feature set $\mathbf{f} = \{f_1, f_2, f_3, f_4\}$ obtained by the proposed curvature analysis in tensor domain, based on scheme 2. It is worth noting that the local feature set h_1 is obtained on the basis of dictionary trained by $\mathbf{f} = \{f_1, f_2, f_3, f_4\}$. The experimental results are presented in Table 2.

Scheme 1 is observed to have some effects on the quality assessment of the HDRIs. However, its performance is similar to those of most existing LDR-IQMs, and the accuracy requires improvement. Scheme 2 produces SROCC and PLCC indexes of 0.7898 and 0.8085, respectively, which makes it superior to most existing LDR-IQMs. Compared with scheme 1, these indexes represent improvements of about 29% and 22%, respectively. This indicates that HDRI feature extraction in tensor domain is more effective than that in the RGB domain. However, scheme 3 further increases the SROCC and PLCC indexes to 0.8839 and 0.9001, showing that the local statistical feature f_4 effectively complements those extracted by scheme 2. The feature set f obtained by using curvature analysis in tensor domain is thus confirmed to be capable of identifying distortions in HDRIs.

3.2. Comparison with other IQMs

The proposed metric is compared with other metrics such as BRISQUE [22], C-DIIVINE [25], BLIINDS2 [26], HDR-VQM [10], and so on, which comprise full-reference and no-reference IQMs. The results for the different metrics are presented in Tables 3 and 4. In order to have comparable results, the RMSE index in Table 4 is calculated after all MOS values are scaled to the range of [0, 100].

LDR-IQMs are thus usually applied after the use of an HVP model such as the Photometric [4], PU [5] and Log models. For database1, this improves the accuracy of LDR-IQMs, as can be observed from Table 3, although the improvement varies among the different combinations. The MSE and PSNR can be observed to be the most correlated in the Log domain, and the least correlated in the Photometric domain. In addition, among the objective metric results for the PU and Log domains, the SSIM has the highest correlation coefficients. In the case of VIF, the improvements are the highest for the PU and photometric domains. Table 3 also lists the results for HDR-VDP-2.2 and HDR-VQM, which can be observed to be relatively accurate. The above-mentioned IQMs are all full-reference metrics. For the no-reference metrics, BLIINDS2 can be observed to be better correlated in the PU and Log domains. BRISQUE is cor-

Table 3. Comparison of different IQMs for database1 and database2.

Reference mode	HVP model	IQMs	Database1			Database2			
			SROCC	PLCC	RMSE	SROCC	PLCC	RMSE	
FR	Photometric	MSE	0.3881	0.4051	23.526	0.7227	0.7080	22.163	
		PSNR	0.4018	0.4409	23.096	0.7183	0.7132	22.000	
		SSIM	0.4953	0.5016	22.261	0.8792	0.8655	15.719	
		VIF	0.7482	0.7603	16.715	0.8723	0.8666	15.659	
	PU	MSE	0.4791	0.4824	22.540	0.8617	0.8559	16.232	
		PSNR	0.4791	0.5297	21.826	0.8617	0.8606	15.984	
		SSIM	0.8553	0.8661	12.861	0.9503	0.9532	9.489	
		VIF	0.7634	0.7845	15.956	0.9306	0.9349	11.142	
	Log	MSE	0.5943	0.6114	20.362	0.8892	0.8856	14.574	
		PSNR	0.5943	0.6456	19.651	0.8892	0.8870	14.494	
		SSIM	0.8935	0.8965	11.400	0.9268	0.9235	12.038	
		VIF	0.4884	0.5079	22.167	0.8385	0.8354	17.249	
	NR	HDR-VDP-2.2	HDR-VDP-2.2	0.8911	0.8989	11.276	0.9503	0.9531	9.496
			HDR-VQM	0.8874	0.8949	11.481	0.9572	0.9612	8.657
		Log	C-DIIVINE	0.3631	0.4572	19.072	0.6228	0.7127	16.788
			BRISQUE	0.5001	0.5847	16.876	0.5446	0.6103	18.298
BLIINDS2			0.8495	0.8628	10.522	0.8703	0.9149	9.292	
PU		C-DIIVINE	0.4086	0.4452	19.202	0.8183	0.8330	13.242	
		BRISQUE	0.8042	0.8183	12.114	0.8620	0.8960	10.336	
		BLIINDS2	0.8392	0.8542	10.920	0.8694	0.9125	9.476	
		NQM-HDRI	0.9031	0.9118	8.876	0.9139	0.9393	8.038	

related well in the PU domain and poorly correlated in the Log domain. C-DIIVINE has the poorest correlation in the Log and PU domains. However, despite the improvements afforded by the combined use of the HVP model with the LDR-IQMs, their performance results are still lower than those of the proposed NQM-HDRI. The PLCC and SROCC indexes of the proposed NQM-HDRI are 0.9031 and 0.9118, respectively, which are higher than those of the other IQMs. The RMSE of the proposed NQM-HDRI is also as low as 8.876, which represents a higher HDRI quality prediction accuracy compared with the other metrics.

For database2, a good PLCC and SROCC which are all over 0.8 are obtained by most full-reference IQMs after the use of PU or Log, especially for the SSIM and VIF. Although the performance in the Photometric domain is relatively worse than that in PU and Log domains, the SSIM and VIF also achieve good results. HDR-VDP-2.2 and HDR-VQM, as true HDR-IQMs, are the best performing metrics when compared to all other IQMs. In contrast to full-reference IQMs, no-reference IQMs show the worst performance with the PLCC and SROCC. However, no-reference IQMs can be applied more widely, because the reference HDRI is not essential. No-reference IQMs are relatively

T a b l e 4. Comparison of different IQMs for database3.

Reference mode	HVP models	IQMs	Database3		
			SROCC	PLCC	RMSE
FR	Photometric	MSE	0.5711	0.5095	25.684
		PSNR	0.5737	0.5594	24.742
		SSIM	0.6770	0.6708	22.138
		VIF	0.4864	0.6144	23.551
	PU	MSE	0.8065	0.8024	17.814
		PSNR	0.8065	0.8009	17.874
		SSIM	0.9121	0.9201	11.688
		VIF	0.9083	0.9181	11.828
	Log	MSE	0.8719	0.8820	14.067
		PSNR	0.8710	0.8819	14.071
		SSIM	0.8179	0.8255	16.847
		VIF	0.6653	0.7065	21.126
		HDR-VDP-2.2	0.9298	0.9408	10.120
		HDR-VQM	0.9193	0.9332	10.725
NR	Log	C-DIIVINE	0.3944	0.3992	26.2355
		BRISQUE	0.5833	0.4065	26.3532
		BLIINDS2	0.5250	0.6963	20.7921
	PU	C-DIIVINE	0.2654	0.4418	25.6699
		BRISQUE	0.3667	0.3780	28.1244
		BLIINDS2	0.5833	0.7135	19.1574
		NQM-HDRI	0.7833	0.8477	14.127

sufficient for realizing a satisfactory prediction accuracy considering the PU domain, which are all over 0.8. Although the proposed NQM-HDRI is insufficient compared with some full-reference IQMs, it is also superior to most no-reference IQMs, whose PLCC and SROCC are 0.9139 and 0.9393, respectively.

Owing to fewer HDRIs in database3, it is not suitable for sparse dictionary training. Thus, the sparse dictionary trained from database1 is used here and the experimental results are presented in Table 4. Similar to database1 and database2, most full-reference IQMs have achieved good results in the PU and Log domains, with all PLCC and SROCC indexes being over 0.8. Moreover, the performance of HDR-VDP-2.2 and HDR-VQM are all over 0.9. However, the results of the no-reference IQMs show that they do not sufficient realize a satisfactory prediction accuracy considering the HVP models in database3. Compared with other no-reference IQMs, the performance of the NQM-HDRI evidently improves, whose PLCC and SROCC are 0.7833 and 0.8477, respectively. To summarize, the proposed NQM-HDRI is indeed more accurate for database1, database2 and database3 than the other no-reference IQMs, which indicates it provides more accurate quality prediction for HDRIs. What needs to be explained in the experimental analysis is that part of the data in Tables 3 and 4 are derived from [4].

4. Conclusion

Although the higher peak brightness and richer colors offered by high dynamic range images (HDRIs) enhance the detection of artifacts, they also make it more difficult to apply a quality metric. Some researchers have actually shown that low dynamic range image quality metrics are unsuitable for HDRIs. This paper thus proposes a no-reference quality metric for an HDRI (NQM-HDRI) based on a curvature analysis in tensor domain. On the basis of the color information and geometric structure information of the HDRI, a curvature analysis is used to extract an effective feature set that is sensitive to the distortion levels in the HDRI. Specifically, signal processing in tensor domain enables the maintenance of most of the color information, and a curvature analysis is used to extract the geometric structure information. A quality-related label feature matrix (Q-LFM) is then constructed by dictionary learning, and the local feature set of the HDRI extracted from the Q-LFM is obtained. By combining local and global feature sets, the predictive quality of the HDRI is finally determined by a pooling strategy. Three public databases are used to demonstrate the high predictive performance of the proposed NQM-HDRI. The metric nevertheless still has some limitations. The extracted feature set is only sensitive to compression distortions and the accuracy of the metric for evaluating other distortions requires improvement. There are also some limitations regarding the quality-related labels considering that the degrees of distortion of specific patches may vary among different distorted versions of the HDRI. Immediate further study would include a refinement of the proposed NQM-HDRI to address the above-mentioned limitations, and further extending it to HDR video quality assessment.

Acknowledgements – This work was supported in part by the Natural Science Foundation of China under Grant 61671258 and 61620106012, in part by the Zhejiang Provincial Natural Science Foundation of China under Grant LY15F010005, and in part by the K. C. Wong Magna Fund at Ningbo University.

References

- [1] CHALMERS A., DEBATTISTA K., *HDR video past, present and future: a perspective*, Signal Processing: Image Communication **54**, 2017, pp. 49–55, DOI: [10.1016/j.image.2017.02.003](https://doi.org/10.1016/j.image.2017.02.003).
- [2] MANTIUK R.K., *Practicalities of predicting quality of high dynamic range images and video*, [In] *2016 IEEE International Conference on Image Processing (ICIP)*, 2016, pp. 904–908, DOI: [10.1109/ICIP.2016.7532488](https://doi.org/10.1109/ICIP.2016.7532488).
- [3] HANHART P., BERNARDO M.V, PEREIRA M., PINHEIRO A.M.G., EBRAHIMI T., *Benchmarking of objective quality metrics for HDR image quality assessment*, EURASIP Journal on Image and Video Processing **2015**(1), 2015, article ID 39, DOI: [10.1186/s13640-015-0091-4](https://doi.org/10.1186/s13640-015-0091-4).
- [4] ZERMAN E., VALENZISE G., DUFAUX F., *An extensive performance evaluation of full-reference HDR image quality metrics*, Quality and User Experience **2**(1), 2017, article ID 5, DOI: [10.1007/s41233-017-0007-4](https://doi.org/10.1007/s41233-017-0007-4).
- [5] AYDIN T.O., MANTIUK R., SEIDEL H.-P., *Extending quality metrics to full luminance range images*, Proceedings of SPIE **6806**, 2008, article ID 68060B, DOI: [10.1117/12.765095](https://doi.org/10.1117/12.765095).
- [6] MANTIUK R., DALY S.J., MYSZKOWSKI K., SEIDEL H.-P., *Predicting visible differences in high dynamic range images: model and its calibration*, Proceedings of SPIE **5666**, 2005, pp. 204–214, DOI: [10.1117/12.586757](https://doi.org/10.1117/12.586757).

- [7] MANTIUK R., KIL JOONG KIM, REMPEL A.G., HEIDRICH W., *HDR-VDP-2: a calibrated visual metric for visibility and quality predictions in all luminance conditions*, ACM Transactions on Graphics **30**(4), 2011, article ID 40, DOI: [10.1145/2010324.1964935](https://doi.org/10.1145/2010324.1964935).
- [8] NARWARIA M., MANTIUK R., DA SILVA M., LE CALLET P., *HDR-VDP-2.2: a calibrated method for objective quality prediction of high-dynamic range and standard images*, Journal of Electronic Imaging **24**(1), 2015, article ID 010501, DOI: [10.1117/1.JEI.24.1.010501](https://doi.org/10.1117/1.JEI.24.1.010501).
- [9] AYDIN T.O., MANTIUK R., MYSZKOWSKI K., SEIDEL H.-P., *Dynamic range independent image quality assessment*, ACM Transactions on Graphics **27**(3), 2008, article ID 69, DOI: [10.1145/1360612.1360668](https://doi.org/10.1145/1360612.1360668).
- [10] NARWARIA M., DA SILVA M.P., LE CALLET P., *HDR-VQM: An objective quality measure for high dynamic range video*, Signal Processing: Image Communication **35**, 2015, pp. 46–60, DOI: [10.1016/j.jvcir.2018.10.016](https://doi.org/10.1016/j.jvcir.2018.10.016).
- [11] JINGTAO XU, PENG YE, QIAOHONG LI, HAIQING DU, YONG LIU, DOERMANN D., *Blind image quality assessment based on high order statistics aggregation*, IEEE Transactions on Image Processing **25**(9), 2016, pp. 4444–4457, DOI: [10.1109/TIP.2016.2585880](https://doi.org/10.1109/TIP.2016.2585880).
- [12] WUFENG XUE, XUANQIN MOU, LEI ZHANG, BOVIK A.C., XIANGCHU FENG, *Blind image quality assessment using joint statistics of gradient magnitude and Laplacian features*, IEEE Transactions on Image Processing **23**(11), 2014, pp. 4850–4862, DOI: [10.1109/TIP.2014.2355716](https://doi.org/10.1109/TIP.2014.2355716).
- [13] QINGBO WU, HONGLIANG LI, FANMAN MENG, NGAN K.N., *Q-DNN: a quality-aware deep neural network for blind assessment of enhanced images*, [In] *2016 Visual Communications and Image Processing (VCIP)*, 2017, pp. 1–4, DOI: [10.1109/VCIP.2016.7805579](https://doi.org/10.1109/VCIP.2016.7805579).
- [14] NARWARIA M., DA SILVA M.P., LE CALLET P., PEPION R., *Tone mapping-based high-dynamic-range image compression: study of optimization criterion and perceptual quality*, Optical Engineering **52**(10), 2013, article ID 102008, DOI: [10.1117/1.OE.52.10.102008](https://doi.org/10.1117/1.OE.52.10.102008).
- [15] KORSHUNOV P., HANHART P., RICHTER T., ARTUSI A., MANTIUK R., EBRAHIMI T., *Subjective quality assessment database of HDR images compressed with JPEG XT*, [In] *2015 Seventh International Workshop on Quality of Multimedia Experience (QoMEX)*, 2015, pp. 1–6, DOI: [10.1109/QoMEX.2015.7148119](https://doi.org/10.1109/QoMEX.2015.7148119).
- [16] YI PENG, DEYU MENG, ZONGBEN XU, CHENQIANG GAO, YI YANG, BIAO ZHANG, *Decomposable non-local tensor dictionary learning for multispectral image denoising*, [In] *2014 IEEE Conference on Computer Vision and Pattern Recognition*, 2014, pp. 2949–2956, DOI: [10.1109/CVPR.2014.377](https://doi.org/10.1109/CVPR.2014.377).
- [17] LIN ZHANG, YING SHEN, HONGYU LI, JIANWEI LU, *3D palmprint identification using block-wise features and collaborative representation*, IEEE Transactions on Pattern Analysis and Machine Intelligence **37**(8), 2015, pp. 1730–1736, DOI: [10.1109/TPAMI.2014.2372764](https://doi.org/10.1109/TPAMI.2014.2372764).
- [18] ZHOU WANG, BOVIK A.C., SHEIKH H.R., SIMONCELLI E.P., *Image quality assessment: from error visibility to structural similarity*, IEEE Transactions on Image Processing **13**(4), 2004, pp. 600–612, DOI: [10.1109/TIP.2003.819861](https://doi.org/10.1109/TIP.2003.819861).
- [19] LIN ZHANG, LEI ZHANG, XUANQIN MOU, *RFSIM: a feature based image quality assessment metric using Riesz transforms*, [In] *2010 IEEE International Conference on Image Processing*, 2010, pp. 321–324, DOI: [10.1109/ICIP.2010.5649275](https://doi.org/10.1109/ICIP.2010.5649275).
- [20] LIN ZHANG, LEI ZHANG, XUANQIN MOU, ZHANG D., *FSIM: a feature similarity index for image quality assessment*, IEEE Transactions on Image Processing **20**(8), 2011, pp. 2378–2386, DOI: [10.1109/TIP.2011.2109730](https://doi.org/10.1109/TIP.2011.2109730).
- [21] RUBINSTEIN R., PELEG T., ELAD M., *Analysis K-SVD: a dictionary-learning algorithm for the analysis sparse model*, IEEE Transactions on Signal Processing **61**(3), 2013, pp. 661–677, DOI: [10.1109/TSP.2012.2226445](https://doi.org/10.1109/TSP.2012.2226445).
- [22] MITTAL A., MOORTHY A.K., BOVIK A.C., *No-reference image quality assessment in the spatial domain*, IEEE Transactions on Image Processing **21**(12), 2012, pp. 4695–4708, DOI: [10.1109/TIP.2012.2214050](https://doi.org/10.1109/TIP.2012.2214050).
- [23] SOO-CHANG PEI, LI-HENG CHEN, *Image quality assessment using human visual DOG model fused with random forest*, IEEE Transactions on Image Processing **24**(11), 2015, pp. 3282–3292, DOI: [10.1109/TIP.2015.2440172](https://doi.org/10.1109/TIP.2015.2440172).

- [24] ITU-T P.1401, *Methods, metrics and procedures for statistical evaluation, qualification and comparison of objective quality prediction models*, International Telecommunication Union, 2012.
- [25] MOORTHY A.K., BOVIK A.C., *Blind image quality assessment: from natural scene statistics to perceptual quality*, IEEE Transactions on Image Processing **20**(12), 2011, pp. 3350–3364, DOI: [10.1109/TIP.2011.2147325](https://doi.org/10.1109/TIP.2011.2147325).
- [26] SAAD M.A., BOVIK A.C., CHARRIER C., *Blind image quality assessment: a natural scene statistics approach in the DCT domain*, IEEE Transactions on Image Processing **21**(8), 2012, pp. 3339–3352, DOI: [10.1109/TIP.2012.2191563](https://doi.org/10.1109/TIP.2012.2191563).

Received November 13, 2017