

No-reference Image Quality Assessment With A Gradient-induced Dictionary

Leida Li¹, Dong Wu¹, Jinjian Wu², Jiansheng Qian^{1,*} and Beijing Chen³

¹ School of Information and Electrical Engineering, China University of Mining and Technology, China
[e-mail: reader1104@hotmail.com, wudong@cumt.edu.cn, memerry507@163.com]

² School of Electronic Engineering, Xidian University, China
[e-mail: jinjian.wu@mail.xidian.edu.cn]

³ School of Computer and Software, Nanjing University of Information Science and Technology, China
[e-mail: nbutimage@126.com]

*Corresponding author: Jiansheng Qian

*Received August 10, 2015; revised October 22, 2015; accepted November 6, 2015;
published January 31, 2016*

Abstract

Image distortions are typically characterized by degradations of structures. Dictionaries learned from natural images can capture the underlying structures in images, which are important for image quality assessment (IQA). This paper presents a general-purpose no-reference image quality metric using a Gradient-Induced Dictionary (GRID). A dictionary is first constructed based on gradients of natural images using K-means clustering. Then image features are extracted using the dictionary based on Euclidean-norm coding and max-pooling. A distortion classification model and several distortion-specific quality regression models are trained using the support vector machine (SVM) by combining image features with distortion types and subjective scores, respectively. To evaluate the quality of a test image, the distortion classification model is used to determine the probabilities that the image belongs to different kinds of distortions, while the regression models are used to predict the corresponding distortion-specific quality scores. Finally, an overall quality score is computed as the probability-weighted distortion-specific quality scores. The proposed metric can evaluate image quality accurately and efficiently using a small dictionary. The performance of the proposed method is verified on public image quality databases. Experimental results demonstrate that the proposed metric can generate quality scores highly consistent with human perception, and it outperforms the state-of-the-arts.

Keywords: Image quality assessment, no-reference, dictionary, K-means clustering, SVM

This research was supported by the Fundamental Research Funds for the Central Universities (2015XKMS032), A Project Funded by the Priority Academic Program Development of Jiangsu Higher Education Institutions (PAPD) and Jiangsu Collaborative Innovation Center on Atmospheric Environment and Equipment Technology (CICAEET).

1. Introduction

Digital images are easily subject to distortions during their acquisition, processing and transmission, such as noise, blurring, blocking artifacts, etc. Image quality assessment (IQA) is of great importance in real-world applications, such as image/video coding [1, 2, 3], image recognition [4], image forensics [5, 6,7] and image enhancement [8]. Objective IQA aims to build computational models for measuring image quality, and meantime keeps consistent with the human visual system (HVS) [9]. According to the necessity and amount of reference information, the existing IQA metrics can be categorized into full-reference (FR), reduced-reference (RR) and no-reference (NR). FR/RR metrics require complete/partial information of a reference image, which is usually not available in practice. By contrast, NR metrics evaluate the quality of an image without any reference [10], so they are more useful in real-world scenarios. Most of the current NR-IQA metrics are designed for specific distortions, such as blockiness [11, 12, 13] and blur [14, 15, 16]. Generally, these distortion-specific metrics are limited in application scopes. By contrast, general-purpose NR quality metrics have much wider applications, because they can evaluate image quality without any prior information. Recently, general-purpose NR-IQA has attracted extensive attention.

Most of the existing general-purpose NR-IQA metrics are based on Natural Scene Statistics (NSS). Moorthy *et al.* first extracted NSS features in the wavelet domain [17, 18]. Then support vector machine (SVM) was utilized for model training and the subsequent quality evaluation. Saad *et al.* extracted NSS features in DCT domain [19]. The generalized Gaussian distribution (GGD) was used to characterize the distribution of DCT coefficients, and the estimated parameters of GGD were combined to produce the quality score. In [20], Mittal *et al.* modeled the distortion-induced decrease in image “naturalness” by locally normalized luminance in the spatial domain. Then image quality was estimated by quantifying the decreased naturalness. In [21], Mittal *et al.* first extracted NSS features for measuring the statistical regularities in undistorted natural images. The parameters of the multivariate Gaussian (MVG) model were used to characterize the naturalness. For a distorted image, the MVG parameters were computed, and the distance between the MVG parameters was computed as the quality score. Liu *et al.* extracted NSS features based on image entropy in both spatial and spectral domains [22]. The SVM was then employed for model training and quality evaluation. Besides the NSS-based approaches, machine learning-based approach has also been investigated recently. In [23], Ye *et al.* addressed the unsupervised feature learning framework for NR-IQA. Specifically, a visual codebook was first trained using raw image patches. Image features were then extracted based on the codebook. The quality score of an image was generated based on support vector regression (SVR). This metric performs consistently well on a variety of databases. However, it is sensitive to the size of the codebook. Typically, a 10000-dimensional codebook is needed to achieve satisfactory results, which requires much memory.

Unsupervised feature learning has been proved to be able to extract highly discriminative features for image analysis [24, 25]. This paper presents a new feature learning-based general-purpose NR-IQA metric using a Gradient-Induced Dictionary (GRID). A dictionary is first trained from gradients of natural image patches based on K-means clustering. With the dictionary, image features are then extracted using Euclidean-norm coding and max-pooling. By combining these features with image distortion types and subjective scores respectively, a distortion classification model and several distortion-specific quality regression models are

trained using the SVM. In the quality evaluation stage, features are first extracted from distorted images. Then they are input to the two kinds of models, producing the probabilities that the image belongs to different distortions and the corresponding distortion-specific quality scores. The final quality score is defined as the probability-weighted distortion-specific scores. The proposed method can evaluate image quality accurately using a small dictionary with low computational complexity. Experiments on public image quality databases demonstrate the advantages of the proposed method.

2. Proposed Quality Model

The diagram of the proposed method is shown in Fig. 1. The proposed quality metric is a two-stage approach, which consists of distortion classification and distortion-specific quality assessment. To achieve this goal, a distortion classification model and several quality regression models are trained based on image features, which are extracted with a pre-trained dictionary using Euclidean-norm coding and max-pooling. SVM is utilized for the model training and quality evaluation. In the following subsections, we will introduce key steps of the proposed method in detail.

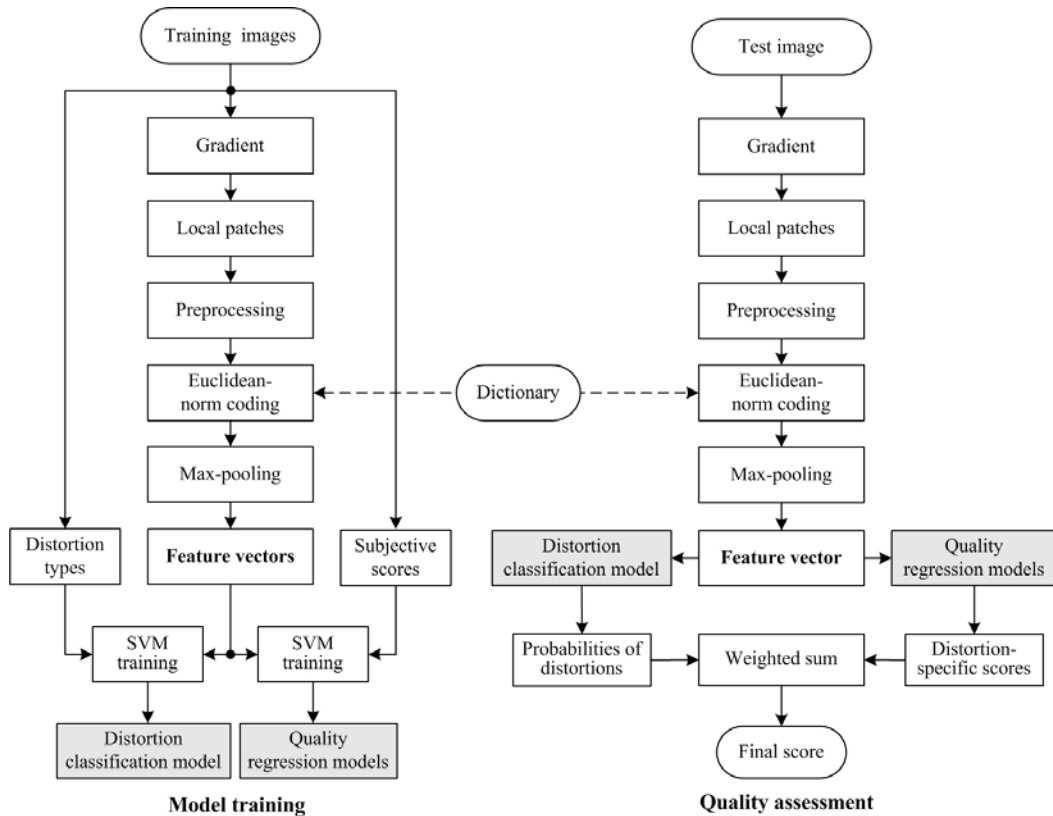


Fig. 1. Diagram of the proposed quality metric.

2.1 Gradient-induced Dictionary

It has been widely acknowledged that natural images share underlying structures, which can be represented using a dictionary [26]. With such a dictionary, image features can be extracted

with appropriate coding and pooling. Traditionally, raw patches from natural images are used to train such a dictionary. However in practice, image structures are usually present in the form of edges and gradients. Therefore, training the dictionary using raw image patches may not be sufficient. In order to highlight structure components in the dictionary, we propose to train the dictionary in the gradient domain instead of the spatial domain. This gradient-induced dictionary is then employed to design the quality model in this work.

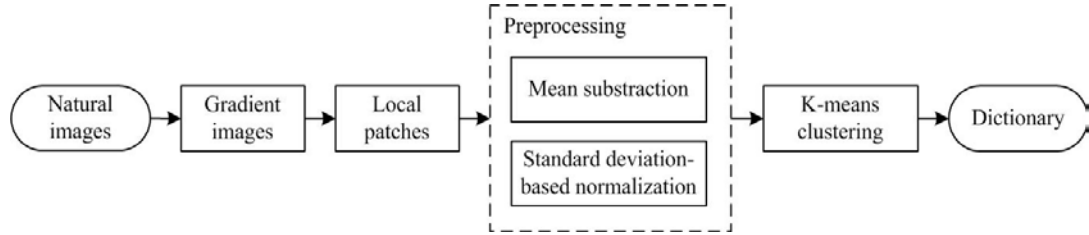


Fig. 2. Diagram of dictionary learning.

Fig. 2 illustrates how the dictionary is constructed in this paper, which consists of several key steps, namely computing image gradient, preprocessing and K-means clustering. For natural images, the gradients are first computed using the Scharr operator [27]:

$$h_x = \frac{1}{16} \begin{bmatrix} 3 & 0 & -3 \\ 10 & 0 & -10 \\ 3 & 0 & -3 \end{bmatrix}, \quad h_y = \frac{1}{16} \begin{bmatrix} 3 & 10 & 3 \\ 0 & 0 & 0 \\ -3 & -10 & -3 \end{bmatrix}, \quad (1)$$

where h_x and h_y denote the Scharr operators in horizontal and vertical directions, respectively. The gradient of an image \mathbf{I} is defined as:

$$\mathbf{G} = \sqrt{(\mathbf{I} \otimes h_x)^2 + (\mathbf{I} \otimes h_y)^2}, \quad (2)$$

where \otimes denotes the convolution.

Once the gradient of an image is obtained, a large number of local patches are randomly selected for training the dictionary. For each patch, it is normalized by first subtracting the mean value and then dividing by its standard deviation [28]:

$$\tilde{\mathbf{G}}(x, y) = \frac{\mathbf{G}(x, y) - \mu}{\sigma + 1}, \quad (3)$$

where $\tilde{\mathbf{G}}(x, y)$ is the normalized patch, $\mathbf{G}(x, y)$ denotes the gradient patch, μ denotes the mean of the patch and σ denotes the standard deviation. By conducting the normalization, the produced dictionary can represent image structures better. In implementation, 10,000 patches with size 7×7 are selected from each gradient image to train the dictionary. It should be noted that since these patches are randomly selected from the image, they may overlap.

In implementation, the patches with size $d = \sqrt{d} \times \sqrt{d}$ (7×7 in this work) are first randomly selected. Then they are normalized and rearranged into row vectors, producing the

training signals $\mathbf{X} = [\mathbf{x}_1; \mathbf{x}_2; \dots; \mathbf{x}_N] \in \mathbf{R}^{N \times d}$, where N denotes the number of patches. Then K-means clustering is employed to generate the dictionary $\mathbf{C} = [\mathbf{c}_1; \mathbf{c}_2; \dots; \mathbf{c}_K] \in \mathbf{R}^{K \times d}$, where $\mathbf{c}_i, i = 1, 2, \dots, K$, denote the K clustering centers.

In this paper, ten natural images are used for dictionary training, and 100,000 local patches (10,000 from each image) are selected to generate the dictionary. Fig. 3 shows the training images and an 800-dimensional dictionary, which has been visualized for better understanding. It is observed from the figure that the atoms in the dictionary are typically edge patterns. These atoms have been proved effective in representing image structures [26], which are important for image quality assessment.

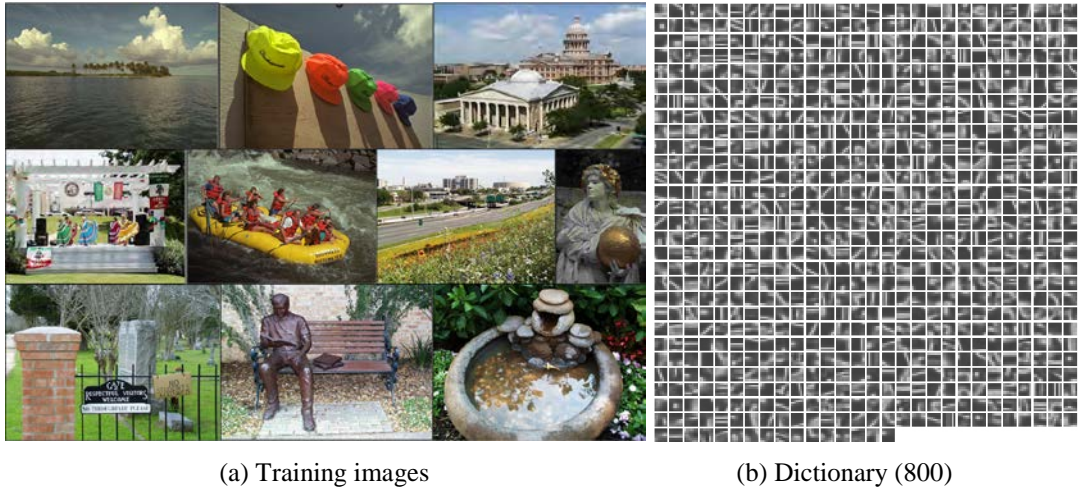


Fig. 3. Training images and the dictionary.

2.2 Feature Generation

In this work, an image is represented by a feature vector, which is extracted from local patches of the image itself using the dictionary. For an image, the gradient is first computed and a large number of local patches are selected followed by normalization, producing the signals $\mathbf{Y} = [\mathbf{y}_1; \mathbf{y}_2; \dots; \mathbf{y}_n] \in \mathbf{R}^{n \times d}$, where \mathbf{y}_i denotes the processed signal of a local patch, and n denotes the patch number. Then with the dictionary $\mathbf{C} \in \mathbf{R}^{K \times d}$, $\mathbf{Y} \in \mathbf{R}^{n \times d}$ is converted into a feature vector $\mathbf{F} = [f_1, f_2, \dots, f_K] \in \mathbf{R}^{1 \times K}$ using Euclidean-norm coding and max-pooling.

In [28], Euclidean-norm coding has been proved effective in generating discriminative features in K-means based dictionary. In this paper, we employ Euclidean-norm coding to generate the feature vector. To this end, the distance between $\mathbf{Y} = [\mathbf{y}_1; \mathbf{y}_2; \dots; \mathbf{y}_n] \in \mathbf{R}^{n \times d}$ and the dictionary $\mathbf{C} = [\mathbf{c}_1; \mathbf{c}_2; \dots; \mathbf{c}_K] \in \mathbf{R}^{K \times d}$ is computed :

$$\mathbf{D} = \|\mathbf{Y} - \mathbf{C}\|_2 = \sqrt{\mathbf{Y}^2 - 2\mathbf{Y} * \mathbf{C}^T + \mathbf{C}^2}, \quad (4)$$

where $\mathbf{D} = [\mathbf{d}_1, \mathbf{d}_2, \dots, \mathbf{d}_K] \in \mathbf{R}^{n \times K}$ denotes the Euclidean distance between $\mathbf{Y} \in \mathbf{R}^{n \times d}$ and

$\mathbf{C} \in \mathbf{R}^{K \times d}$. Then the average distance $\mu(\mathbf{D}) \in \mathbf{R}^{n \times 1}$ between each patch and the dictionary is computed. Finally, the difference between $\mathbf{D} \in \mathbf{R}^{n \times K}$ and $\mu(\mathbf{D}) \in \mathbf{R}^{n \times 1}$ is computed and denoted by $\mathbf{Z} = [\mathbf{z}_1, \mathbf{z}_2, \dots, \mathbf{z}_K] \in \mathbf{R}^{n \times K}$.

With $\mathbf{Z} \in \mathbf{R}^{n \times K}$, the feature vector of an image is generated based on pooling. In IQA, pooling is very useful for making the predicted quality scores consistent with subjective scores [29]. Generally, minimum-pooling, average-pooling and maximum-pooling can be used. In this paper, we employ max-pooling in this stage, which can produce better results. Specifically, each element of $\mathbf{Z} \in \mathbf{R}^{n \times K}$ is first compared with zero, and the larger one is selected and denoted by $\tilde{\mathbf{Z}} = [\tilde{\mathbf{z}}_1, \tilde{\mathbf{z}}_2, \dots, \tilde{\mathbf{z}}_K] \in \mathbf{R}^{n \times K}$. Then the maximum value of each column vector $\tilde{\mathbf{z}}_i = (\alpha_{1i}, \alpha_{2i}, \dots, \alpha_{ni})$, $i = 1, 2, \dots, K$ in $\tilde{\mathbf{Z}} \in \mathbf{R}^{n \times K}$ is obtained, which is denoted by $f_i = \max\{\tilde{\mathbf{z}}_i\} = \max\{\alpha_{1i}, \alpha_{2i}, \dots, \alpha_{ni}\}$, $i = 1, 2, \dots, K$. Then $\mathbf{F} = [f_1, f_2, \dots, f_K] \in \mathbf{R}^{1 \times K}$ is the final feature vector, which is used for the subsequent quality assessment.

2.3 Training and Testing

The proposed method consists of two stages, namely model training and quality assessment. In the first stage, a distortion classification model and several quality-specific regression models are trained using the SVM. The distortion classification model is used to predict the probabilities that an image belongs to different kinds of distortions, and the quality regression models are used to predict the corresponding distortion-specific quality scores. In the quality assessment stage, the probabilities and distortion-specific quality scores are first generated using the trained models. Then an overall quality score of the test image is computed as the probability-weighted distortion-specific scores.

For a group of k training images $\{\mathbf{I}_i, i = 1, 2, \dots, k\}$ together with their distortion types $\{l_i, i = 1, 2, \dots, k\}$ and subjective scores $\{s_i, i = 1, 2, \dots, k\}$, the feature vectors are first extracted, which are denoted by $\{\mathbf{F}_i, i = 1, 2, \dots, k\}$. Then $\{\mathbf{F}_i, i = 1, 2, \dots, k\}$ and $\{l_i, i = 1, 2, \dots, k\}$ are used to train the distortion classification model, while $\{\mathbf{F}_i, i = 1, 2, \dots, k\}$ and $\{s_i, i = 1, 2, \dots, k\}$ are used to train the distortion-specific regression models. In this paper, both kinds of models are trained using the SVM [30], and the Radial Basis Function (RBF) is used as the kernel. The distortion classification model is trained using the standard SVM classifier, and the regression models are trained using e-SVR [30].

For a test image, the feature vector is first extracted. Then it is input to the trained models, producing the probabilities of distortions $\{p_j, j = 1, 2, \dots, m\}$ and the corresponding distortion-specific quality scores $\{q_j, j = 1, 2, \dots, m\}$. Finally, an overall quality score is computed as follows:

$$\text{Score} = \sum_{j=1}^m p_j \cdot q_j \quad (5)$$

where m denotes the number of distortion types considered.

3. Experimental Results and Analysis

3.1 Experiment Settings

Four commonly used image quality databases are used to test the performance of the proposed method, including LIVE [31], CSIQ [32], TID2008 [33] and TID2013 [34]. The LIVE database contains 29 reference images and 779 distorted images with five kinds of distortions: JPEG compression, JPEG2000 compression (JP2K), additive Gaussian white noise (WN), Gaussian blur (GB) and Fast Fading Rayleigh (FF). The CSIQ database consists of 30 reference images and 866 distorted images with six distortion types. In TID2008, there are 25 reference images and the corresponding 1700 distorted images with 17 distortion types. TID2013 is an extended version of TID2008, which contains 3000 distorted images with 24 distortion types. **Table 1** summarizes the information of the four databases. In the table, MOS denotes the mean opinion score, and DMOS denotes the difference MOS, which are used as ground truth of subjective qualities. In our experiments, we focus on JPEG, JK2K, WN and GB, which are the most common distortion types and also shared by the four databases.

Table 1. Database information for performance evaluation

Database	Reference images	Distorted images	Distortion types	Subjects	Subjective scores
LIVE	29	779	5	161	DMOS
CSIQ	30	866	6	35	DMOS
TID2008	25	1700	17	838	MOS
TID2013	25	300	24	971	MOS

Three quantitative criteria are employed to measure the performance of the proposed method. Specifically, Pearson Linear Correlation Coefficient (PLCC) and Root Mean Squared Error (RMSE) are used to measure the prediction accuracy, and Spearman Rank Order Correlation Coefficient (SRCC) is used to measure the prediction monotonicity. Before computing these criteria, a five-parameter logistic mapping is conducted between the subjective and objective scores [35]:

$$f(x) = \beta_1 \left(\frac{1}{2} - \frac{1}{1 + e^{\beta_2(x - \beta_3)}} \right) + \beta_4 x + \beta_5, \quad (6)$$

where β_i , $i = 1, 2, 3, 4, 5$ are the parameters to be fitted. Let the subjective and mapped objective scores of the i^{th} image be denoted by s_i and q_i , then PLCC, SRCC and RMSE are defined as follows:

$$\text{PLCC} = \frac{\sum_{i=1}^N (s_i - \bar{s})(q_i - \bar{q})}{\sqrt{\sum_{i=1}^N (s_i - \bar{s})^2} \sqrt{\sum_{i=1}^N (q_i - \bar{q})^2}}, \quad (7)$$

$$\text{SRCC} = 1 - \frac{6 \sum_{i=1}^N d_i^2}{N(N^2 - 1)}, \quad (8)$$

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{i=1}^N (s_i - q_i)^2}, \quad (9)$$

where \bar{s} and \bar{q} denote the average values of the subjective and mapped objective scores, d_i denotes the difference between the i^{th} image's ranks in subjective and objective evaluations, and N denotes the total number of images

In experiments, several parameters should be determined first, including the patch size, gradient operator, pooling method and the SVM kernels. In this work, these parameters are set as: patch size 7×7 , Scharr gradient operator, max-pooling and RBF kernel, which are tuned based on LIVE database. A combination of them have achieved the best performance, and the corresponding PLCC and SRCC values are 0.9735 and 0.9722, respectively.

In order to have an intuitionistic view of the influences of these parameters on the performances of the proposed method, we test the performances when different parameters are used. To be specific, for patch size, we test 5×5 , 9×9 , 7×7 and 11×11 . For the gradient operator, we test the commonly used Roberts, Sobel, Prewitt and Scharr. For pooling, we test min-pooling, average-pooling and max-pooling. For SVM kernel, we test linear, polynomial, Sigmoid and RBF. In our experiment, 80% images in LIVE database are used for training and the other 20% images are used for test. The training-test procedure is repeated by 1,000 times, and the median result is employed for performance evaluation. It should be noted that in each iteration, the 80% training images are selected randomly. Since the LIVE database contains 29 reference images, each time 23 reference images and the corresponding distorted images are used for model training, while the other 6 reference images and the corresponding distorted images are used for test. The experimental results in terms of PLCC and SRCC are summarized in [Tables 2-5](#).

Table 2. Performances of the proposed metric using different patch sizes on LIVE database.

Patch size	5×5	7×7	9×9	11×11
PLCC	0.9727	0.9735	0.9722	0.9733
SRCC	0.9719	0.9722	0.9712	0.9717

Table 3. Performances of the proposed metric using different gradient operators on LIVE database.

Gradient	Roberts	Sobel	Prewitt	Scharr
PLCC	0.9669	0.9711	0.9681	0.9735
SRCC	0.9645	0.9701	0.9668	0.9722

Table 4. Performances of the proposed metric using different pooling methods on LIVE database.

Pooling	Min-pooling	Average-pooling	Max-pooling
PLCC	0.7503	0.9490	0.9735
SRCC	0.7129	0.9498	0.9722

Table 5. Performances of the proposed metric using different SVM kernels on LIVE database.

SVM kernel	Linear	Ploynomial	Sigmoid	RBF
PLCC	0.5298	0.8234	0.2650	0.9735
SRCC	0.4119	0.8258	0.1824	0.9722

It is observed from the [Tables 2-5](#) that patch size and gradient operator have little influences on the performance, while patch size 7×7 and Scharr operator produce slightly

better results. By contrast, the pooling method and SVM kernel have great influences on the performances, while max-pooling and RBF kernel produce the best results.

3.2 Performance on LIVE Database

In this subsection, the performance of the proposed method on LIVE database is evaluated. For comparison, eight state-of-the-art general-purpose NR quality metrics are also tested, including Blind Image Quality Index (BIQI) [17], Distortion Identification-based Image Verity and INtegrity Evaluation (DIIVINE) [18], BLind Image Integrity Notator using DCT Statistics (BLIINDS-II) [19], Blind/Referenceless Image Spatial Quality Evaluator (BRISQUE) [20], Natural Image Quality Evaluator (NIQE) [21], Spatial-Spectral Entropy based Quality (SSEQ) [22], COdebook Representation for No-reference Image quality Assessment (CORNIA) [23], and Quality-Aware Clustering (QAC) [36]. The source codes of these metrics are downloaded from the authors' websites and the default parameters are used.

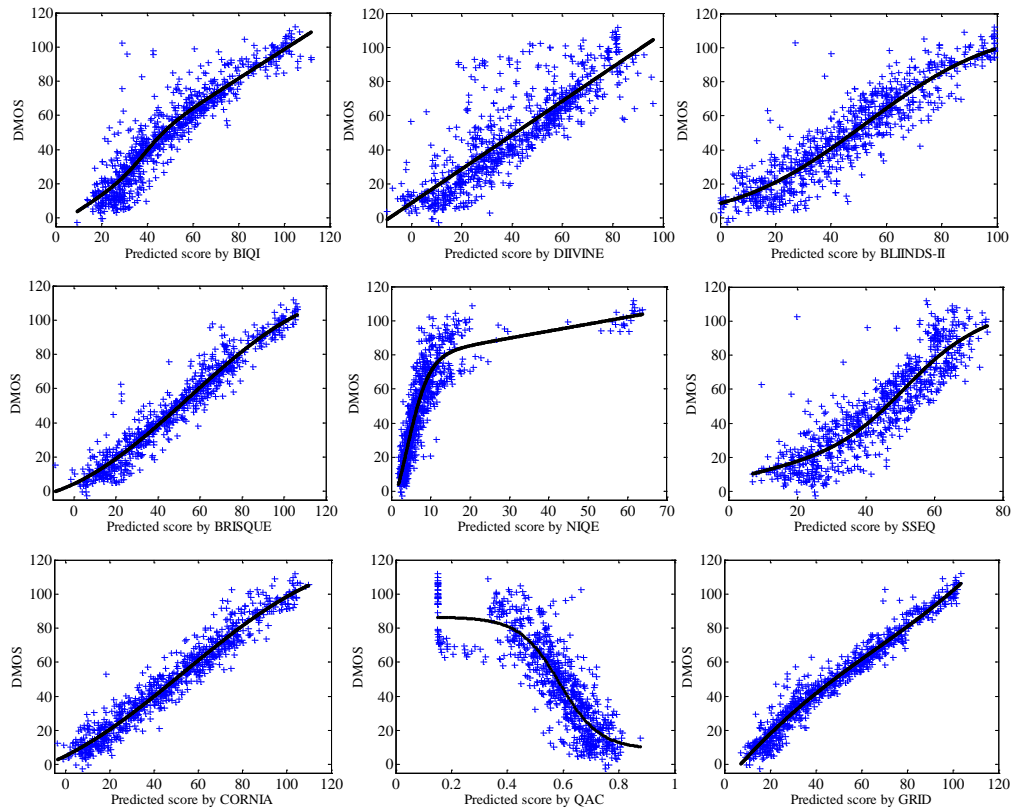


Fig. 4. Scatter plots of subjective scores versus predicted scores of different metrics on LIVE database.

Fig. 4 shows the fitting results of the nine metrics on LIVE database. A good quality metric is expected to produce a fitting curve with sample points gathering closely around it. It is observed from **Fig. 4** that BRISQUE, CORNIA and the proposed GRID produce better fitting results than the other metrics. Furthermore, the fitting result of GRID is slightly better than those of BRISQUE and CORNIA.

Based on the fitting results, PLCC, RMSE and SRCC can be calculated based on equations (7)-(9). **Tables 6-8** summarize the experimental results. In the tables, the results on each distortion type as well as the overall performance are listed, and the best two results are marked in boldface.

Table 6. Performances in terms of PLCC on LIVE database.

Metric	JP2K	JPEG	WN	GB	FF	All
BIQI [17]	0.9270	0.9631	0.9928	0.9613	0.8197	0.9219
DIIVINE [18]	0.9267	0.8041	0.9913	0.9600	0.8741	0.8443
BLIINDS-II [19]	0.9352	0.9471	0.9654	0.9175	0.8453	0.9144
BRISQUE [20]	0.9553	0.9866	0.9926	0.9720	0.9352	0.9670
NIQE [21]	0.9261	0.9514	0.9763	0.9433	0.8792	0.9020
SSEQ [22]	0.9222	0.9460	0.9812	0.9376	0.8704	0.8925
CORNIA [23]	0.9684	0.9750	0.9823	0.9856	0.9545	0.9691
QAC [36]	0.8648	0.9435	0.9180	0.9105	0.8248	0.8625
GRID	0.9736	0.9834	0.9933	0.9771	0.9455	0.9735

Table 7. Performances in terms of RMSE on LIVE database.

Metric	JP2K	JPEG	WN	GB	FF	All
BIQI [17]	9.4604	8.5675	3.3598	5.0897	16.3154	10.5832
DIIVINE [18]	9.4794	18.9345	3.6818	5.1691	13.8372	14.6397
BLIINDS-II [19]	8.9341	10.2213	7.2910	7.3470	15.2184	11.0587
BRISQUE [20]	7.4622	5.1958	3.3897	4.3370	10.0873	6.9641
NIQE [21]	9.5169	9.8124	6.0537	6.1307	13.5714	11.7949
SSEQ [22]	9.7572	10.7414	5.3965	6.4237	14.0245	12.3245
CORNIA [23]	6.2929	7.0799	5.2334	3.1236	8.4940	6.7445
QAC [36]	12.6665	10.5550	11.0981	7.6359	16.1059	13.8259
GRID	5.7854	5.7635	3.5623	3.8725	9.2315	6.3204

Table 8. Performances in terms of SRCC on LIVE database.

Metric	JP2K	JPEG	WN	GB	FF	All
BIQI [17]	0.9221	0.9547	0.9903	0.9545	0.8199	0.9198
DIIVINE [18]	0.9185	0.8141	0.9878	0.9581	0.8586	0.8560
BLIINDS-II [19]	0.9299	0.9471	0.9597	0.9103	0.8348	0.9115
BRISQUE [20]	0.9464	0.9816	0.9911	0.9744	0.9254	0.9654
NIQE [21]	0.9186	0.9412	0.9718	0.9328	0.8635	0.9062
SSEQ [22]	0.9213	0.9275	0.9815	0.9246	0.8602	0.8970
CORNIA [23]	0.9612	0.9691	0.9730	0.9827	0.9486	0.9706
QAC [36]	0.8621	0.9362	0.9508	0.9134	0.8231	0.8683
GRID	0.9680	0.9759	0.9907	0.9751	0.9369	0.9722

It is observed from **Tables 6-8** that the proposed GRID outperforms BIQI, DIIVINE, BLIINDS-II, NIQE, SSEQ and QAC, in all kinds of distortion types. For JPEG and WN, BRISQUE produces slightly better results than GRID. For GB and FF, CORNIA produces slightly better results than GRID. For all distortions, the PLCC and SRCC values of the

proposed method are always among the top two. Furthermore, the overall performance of the proposed method is advantageous over all these state-of-the-art metrics. Therefore, the proposed method achieves the best performance on LIVE database.

3.3 Cross-database Evaluation

For learning-based image quality models, it is important to test the generalization ability, i.e., training the model in one database and testing it in other databases. In this experiment, the model is first trained using the whole LIVE database. Then the trained model is used to test the performances of the proposed method on TID2008, TID2013 and CSIQ databases. For the compared learning-based metrics, their models are also trained using the whole LIVE database so that the results are fair for comparison. In order to evaluate the overall performances, the average results across the three databases are also computed, including the direct average and database-size-weighted average. **Tables 9-11** show the cross-database performances, where the best two results are marked in boldface.

Table 9. Performances in terms of PLCC on TID2008, TID2013 and CSIQ databases.

Metric	TID2008 (400 images)	TID2013 (500 images)	CSIQ (600 images)	Direct Average	Weighted Average
BIQI [17]	0.7370	0.7772	0.8225	0.7789	0.7846
DIIVINE [18]	0.7532	0.7859	0.8543	0.7978	0.8045
BLIINDS-II [19]	0.8613	0.8354	0.9026	0.8664	0.8692
BRISQUE [20]	0.8801	0.8701	0.9245	0.8916	0.8945
NIQE [21]	0.7978	0.8075	0.8886	0.8313	0.8374
SSEQ [22]	0.8706	0.8611	0.8870	0.8729	0.8740
CORNIA [23]	0.8916	0.8901	0.9251	0.9023	0.9045
QAC [36]	0.7948	0.8051	0.8736	0.8245	0.8298
GRID	0.9095	0.9131	0.9104	0.9110	0.9111

Table 10. Performances in terms of RMSE on TID2008, TID2013 and CSIQ databases.

Metric	TID2008 (400 images)	TID2013 (500 images)	CSIQ (600 images)	Direct Average	Weighted Average
BIQI [17]	1.0698	0.8777	0.1607	0.7027	0.6421
DIIVINE [18]	1.0325	0.8798	0.1452	0.6858	0.6267
BLIINDS-II [19]	0.8416	0.7531	0.1217	0.5721	0.5241
BRISQUE [20]	0.7516	0.6875	0.1077	0.5156	0.4727
NIQE [21]	0.9543	0.8228	0.1296	0.6356	0.5806
SSEQ [22]	0.7788	0.7091	0.1305	0.5395	0.4962
CORNIA [23]	0.7168	0.6357	0.1073	0.4866	0.4460
QAC [36]	0.9825	0.8273	0.1375	0.6491	0.5928
GRID	0.6428	0.6125	0.1169	0.4574	0.4223

Table 11. Performance in terms of SRCC on TID2008, TID2013 and CSIQ databases.

Metric	TID2008 (400 images)	TID2013 (500 images)	CSIQ (600 images)	Direct Average	Weighted Average
BIQI [17]	0.7532	0.7886	0.7975	0.7798	0.7827
DIIVINE [18]	0.7630	0.7878	0.8284	0.7931	0.7974
BLINDS-II [19]	0.8523	0.8165	0.8787	0.8492	0.8509
BRISQUE [20]	0.8719	0.8472	0.9007	0.8733	0.8752
NIQE [21]	0.7874	0.7964	0.8717	0.8185	0.8241
SSEQ [22]	0.8314	0.8078	0.8484	0.8292	0.8303
CORNIA [23]	0.8866	0.8747	0.8981	0.8865	0.8872
QAC [36]	0.7980	0.8054	0.8416	0.8150	0.8179
GRID	0.9005	0.8993	0.8639	0.8879	0.8855

From the results, we know that in TID2008 and TID2013, the PLCC and SRCC values of GRID are higher than all the compared metrics. In CSIQ, the performance of GRID is slightly worse than CORNIA and BRISQUE. For the weighted results, the proposed method achieves the best overall prediction accuracy, while the overall prediction monotonicity is very close to CORNIA. It should be noted that the performance of CORNIA is tested using a codebook with 10,000 dimensions. By contrast, the dictionary size of the proposed method is only 800, which is much smaller than that of CORNIA. This is highly desired in practice, because the proposed method requires much less memory during the quality evaluation. From this perspective, the proposed method is advantageous over CORNIA.

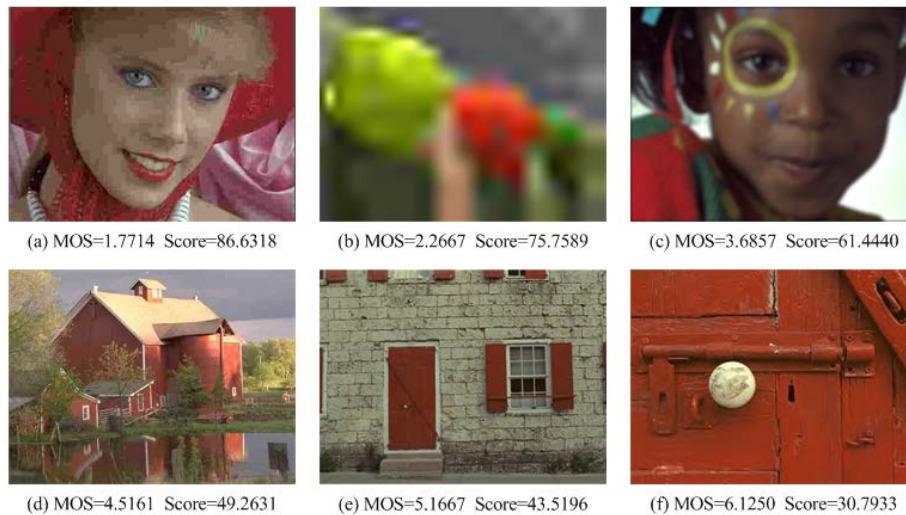
**Fig. 5.** Example images, their subjective scores, and objective scores predicted by GRID. Images are taken from the TID2008 database.

Fig. 5 shows six images from the TID2008 database, together with their subjective and objective scores predicted by GRID. It is observed that with the increase of subjective scores (MOS), the corresponding objective scores (predicted DMOS) decrease accordingly. This indicates that the proposed metric can evaluate the qualities of the images consistently with human subjective ratings.

3.4 Statistical Performance Analysis

In order to know the statistical significance of the state-of-the-art metrics against the proposed method, F-test is conducted [37]. F-test is conducted based on the prediction errors between the predicted scores and the corresponding subjective scores. Let the prediction error of a compared metric be denoted by σ_X^2 and that of the proposed method be denoted by σ_{GRID}^2 , then F-score is defined as:

$$F_{\text{score}} = \sigma_X^2 / \sigma_{\text{GRID}}^2. \quad (10)$$

The significance of a metric can be determined by comparing F_{score} with a threshold F_{critical} , which is related to a confidence level and the number of images in a database. Specifically, if F_{score} is bigger than F_{critical} , the performance of the proposed method is significantly better than the compared metric. If F_{score} is smaller than $1/F_{\text{critical}}$, the performance of the proposed method is significantly worse than the compared metric. Otherwise, their performances are competitive. In this paper, the confidence level is set to 99%, and the thresholds of F_{critical} are 1.1802, 1.2095, 1.2624 and 1.2317, respectively for LIVE, CSIQ, TID2008 and TID2013 databases.

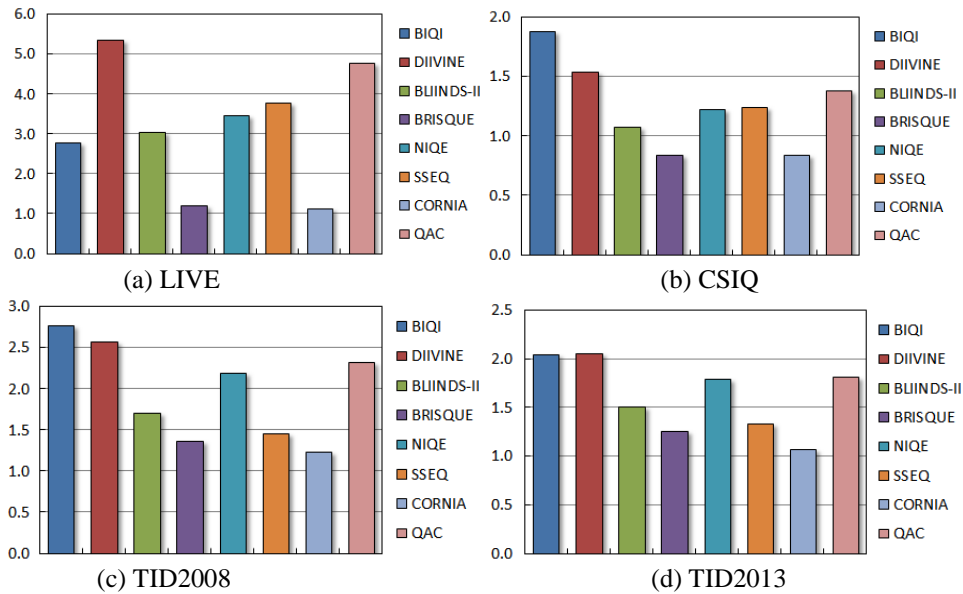


Fig. 6. F statistics of the compared metrics against GRID on four databases.

The results of F statistics are shown in Fig. 6 and Table 12. In Table 12, “1” indicates that the performance of GRID is significantly better than the compared metric, and “0” indicates that performance of a compared metric is competitive to GRID. It is observed from Fig. 6 that in LIVE, TID2008 and TID2013, the values of the bars are all higher than 1, which indicates that the compared metrics all produce bigger prediction errors than the proposed method. In CSIQ, only BRISQUE and CORNIA produce slightly smaller prediction errors than the proposed method. From Table 12, we know that the performance of GRID is significantly

better than BIQI, DIIVINE, BLIINDS-II, BRISQUE, NIQE, SSEQ and QAC in LIVE, TID2008 and TID2013 databases. In CSIQ database, the performances of BLIINDS-II, BRISQUE, and CORNIA are competitive to GRID, and GRID significantly outperforms BIQI, DIIVINE, NIQE, SSEQ and QAC. From this experiment, we know that the statistical performance of GRID is similar to CORNIA, and it is significantly better than the other state-of-the-arts. As aforementioned, CORNIA needs a 10,000-dimensional codebook for image representation, and the proposed method uses an 800-dimensional dictionary, which is much smaller. This indicates that the proposed method does not require too much memory in implementation, which is desired in real applications. From these results, we can conclude that the proposed method outperforms the state-of-the-art general-purpose NR quality metrics.

Table 12. Statistical significance of the state-of-the-art metric against the proposed method.

Metric	LIVE	TID2008	TID2013	CSIQ
BIQI [17]	1	1	1	1
DIIVINE [18]	1	1	1	1
BLIINDS-II [19]	1	1	1	0
BRISQUE [20]	1	1	1	0
NIQE [21]	1	1	1	1
SSEQ [22]	1	1	1	1
CORNIA [23]	0	0	0	0
QAC [36]	1	1	1	1

3.5 Impact of Dictionary Size

In this part, we investigate the impact of dictionary size on the performance of the proposed method. For comparison, we also provide the results of CORNIA, which is also a feature learning-based general-purpose NR quality metric with the state-of-the-art performance. In this experiment, the dictionary sizes are 200, 400, 800, 1200, 2500, 5000 and 10000, respectively. Fig. 7 shows the experimental results in terms of PLCC and SRCC on the LIVE database.

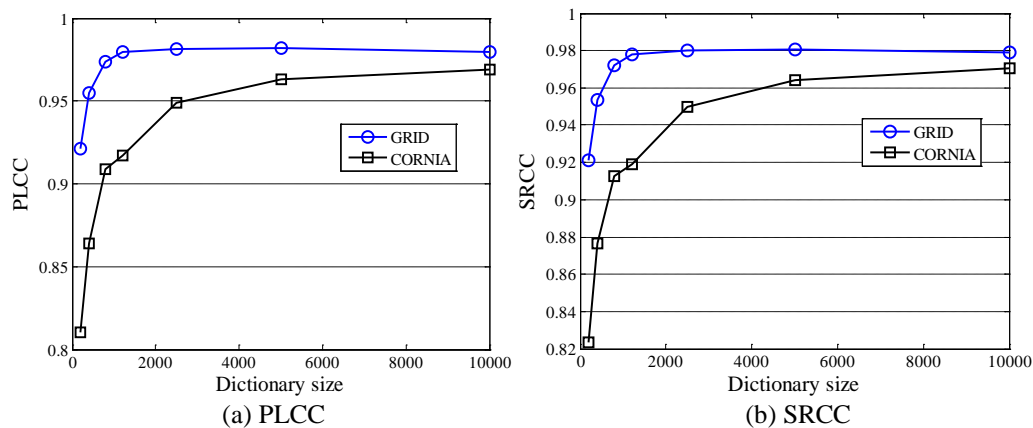


Fig. 7. Impacts of dictionary sizes on GRID and CORNIA (on LIVE database).

It is observed from [Fig. 7](#) that for both methods, their performances improve with the increase of dictionary size. For the proposed GRID, its performance improves rapidly with the increase of dictionary size, and the best performance can be achieved when a small dictionary (around 800) is used. By contrast, the performance of CORNIA improves slowly with the dictionary size, and the best performance is achieved when the dictionary size is 10000. Therefore, the proposed method can produce very promising results with a small dictionary, but CORNIA requires a very large dictionary to produce satisfactory result. In this paper, an 800-dimensional dictionary is used in the proposed method. This further demonstrates that the proposed gradient-induced dictionary is more effective in representing quality-related image structures, thus facilitate image quality assessment.

3.6 Computational Complexity

For NR-IQA metrics, computational complexity is important, especially for real-time applications. In this part, we evaluate the computation costs of the proposed method and the state-of-the-art metrics. To be specific, the average time for generating the quality score of an image from the LIVE database is computed and used for measuring the computational complexity. This experiment is conducted on a 64-bit workstation with 8 Intel Xeon 3.7GHz CPUs and 8.0 GB RAM. The experimental results are listed in [Table 13](#).

Table 13. Average computational time for generating quality score of an image in LIVE database.

Metric	BIQI [17]	DIIVINE [18]	BLINDS-II [19]	BRISQUE [20]	
Time (s)	0.4080	19.9023	103.0710	0.1652	
Metric	NIQE [21]	SSEQ [22]	CORNIA [23]	QAC [36]	GRID
Time (s)	0.2469	1.3773	2.8240	0.0858	0.3108

It is observed from [Table 13](#) that QAC [36] is the fastest method, followed by BRISQUE [20] and NIQE [21]. The computational complexity of the proposed method ranks the fourth, and it is faster than the remaining metrics. DIIVINE [18] and BLINDS-II [19] are very time consuming, which are mainly due to the feature extraction stages. The computational cost of CORNIA [23], another feature learning-based metric with competitive performance, is about nine times higher than the proposed method. From these results, we know that the complexity of the proposed method is low.

3.7 Performances on Multiply-Distorted and Contrast-Changed Images

The performance of the proposed method is further tested on two multiply-distorted image databases and a contrast-distorted image database, namely LIVE Multiply-Distorted image database (LIVEMD) [38], Multiply-Distorted Image Database (MDID2014) [39] and Contrast-Changed Image Database (CCID2014) [40]. [Table 14](#) lists the experimental results and comparisons with the general-purpose NR metrics.

It is observed from [Table 14](#) that the proposed method performs well on multiply distorted images. Specifically, among the 9 metrics, the proposed method ranks the second on LIVEMD database, and on MDID2014 it ranks the third. In LIVEMD, the images are distorted by either “blur and JPEG compression” or “blur and white Gaussian noise”. In MDID2014, the images are simultaneously distorted by blur, JPEG compression and noise injection. Since these distortions have been considered when training our quality model, the proposed method can efficiently evaluate the quality of multiply-distorted images in LIVEMD and MDID2014.

Table 14. Performances in terms of PLCC on LIVE MD, MDID2014 and CCID2014 databases.

Database	LIVE MD [38]		MDID2014 [39]		CCID2014 [40]	
Criterion	PLCC	SRCC	PLCC	SRCC	PLCC	SRCC
BIQI [17]	0.7389	0.6109	0.3369	0.3108	0.1589	0.1692
DIIVINE [18]	0.7183	0.6585	0.3707	0.4248	0.3346	0.2956
BLINDS-II [19]	0.2370	0.2503	0.2221	0.2057	0.5561	0.4014
BRISQUE [20]	0.6049	0.6003	0.5901	0.5960	0.3574	0.2118
NIQE [21]	0.8377	0.7725	0.5634	0.5446	0.4307	0.3655
SSEQ [22]	0.3699	0.3021	0.4762	0.4349	0.3686	0.2908
CORNIA [23]	0.8676	0.8298	0.7727	0.7819	0.1453	0.2088
QAC [36]	0.3764	0.3579	0.1732	0.2270	0.3072	0.1451
GRID	0.8389	0.8106	0.5657	0.5829	0.1369	0.1304

Another observation from **Table 14** is that the proposed method is not effective in evaluating the quality of contrast-distorted images. This is mainly because that when training our quality model, contrast-distorted images are not included. By including contrast-distorted images in the training the quality model, the performance on CCID2014 database can be improved. It should be noted that almost all of the tested quality metrics perform bad on CCID2014 database.

4. Conclusion

In this paper, we have presented a new general-purpose NR image quality metric based on feature learning. A gradient-induced dictionary, trained from gradients of natural images, is adopted to represent images and generate the feature vectors for quality assessment. The proposed method is a two-stage approach, which consists of distortion classification and distortion-specific quality assessment. Euclidean-norm coding and max-pooling have been adopted to generate image features using the dictionary. We have evaluated the performance of the proposed method on four popular image quality databases and compared it with the state-of-the-art general-purpose NR quality metrics. Experimental results have demonstrated that our method outperforms the state-of-the-art metrics with a relatively small dictionary and low computational cost. We have also tested it on multiply-distorted image databases, and very promising results are obtained.

In this work, the proposed method operates on gray-scale images. This means that for color images, they must be converted into gray scale before quality evaluation. In real-world scenarios, almost all images are in color format and color distortions also have significant influence on the perceived quality. Incorporating color in image quality modeling is expected to produce quality scores more consistent with subjective evaluations. Quaternion-based color image processing has attracted extensive attention recently [41]. As future work, we will try to incorporate quaternion-based color processing in image quality assessment.

References

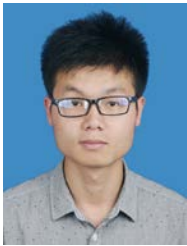
- [1] Y. H. Kim, J. Shin and H. Kim, "Lightweight quality metric based on no-reference bitstream for H.264/AVC video," *KSII Transactions on Internet and Information Systems*, vol. 6, no. 5, pp. 1388-1399, May, 2012. [Article \(CrossRef Link\)](#)
- [2] X. G. Liu, M. Chen, T. Wan and C. Yu, "Hybrid no-reference video quality assessment focusing on codec effects," *KSII Transactions on Internet and Information Systems*, vol. 5, no. 3, pp. 592-606, March, 2011. [Article \(CrossRef Link\)](#)
- [3] Z. Q. Pan, Y. Zhang and S. Kwong, "Efficient motion and disparity estimation optimization for low complexity multiview video coding," *IEEE Transactions on Broadcasting*, vol. 61, no. 2, pp. 166-176, June, 2015. [Article \(CrossRef Link\)](#)
- [4] D. T. Nguyen, Y. H. Park, K. Y. Shin and K. R. Park, "New finger-vein recognition method based on image quality assessment," *KSII Transactions on Internet and Information Systems*, vol. 7, no. 2, pp. 347-365, February, 2013. [Article \(CrossRef Link\)](#)
- [5] J. Li, X. L. Li, B. Yang and X. M. Sun, "Segmentation-based image copy-move forgery detection scheme," *IEEE Transactions on Information Forensics and Security*, vol. 10, no. 3, pp. 507-518, March, 2015. [Article \(CrossRef Link\)](#)
- [6] Z. H. Xia, X. H. Wang, X. M. Sun and B. W. Wang, "Steganalysis of least significant bit matching using multi-order differences," *Security and Communication Networks*, vol. 7, no. 8, pp. 1283-1291, 2014. [Article \(CrossRef Link\)](#)
- [7] Z. H. Xia, X. H. Wang, X. M. Sun, Q. S. Liu and N. X. Xiong, "Steganalysis of LSB matching using differences between nonadjacent pixels," *Multimedia Tools and Applications*, DOI 10.1007/s11042-014-2381-8, 2015. [Article \(CrossRef Link\)](#)
- [8] K. Gu, G. T. Zhai, X. K. Yang, W. J. Zhang and C. W. Chen, "Automatic contrast enhancement technology with saliency preservation," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 25, no. 9, pp. 1480-1494, 2015. [Article \(CrossRef Link\)](#)
- [9] W. S. Lin and C-C. Jay Kuo, "Perceptual visual quality metrics: a survey," *Journal of Visual Communication and Image Representation*, vol. 22, no. 4, pp. 297-312, 2011. [Article \(CrossRef Link\)](#)
- [10] K. Gu, G. T. Zhai, X. K. Yang and W. J. Zhang, "Using free energy principle for blind image quality assessment," *IEEE Transactions on Multimedia*, vol. 17, no. 1, pp. 50-63, 2015. [Article \(CrossRef Link\)](#)
- [11] L. D. Li, H. C. Zhu, G. B. Yang and J. S. Qian, "Referenceless measure of blocking artifacts by Tchebichef kernel analysis," *IEEE Signal Processing Letters*, vol. 21, no. 1, pp. 122-125, 2014. [Article \(CrossRef Link\)](#)
- [12] S. A. Golestaneh and D. M. Chandler, "No-reference quality assessment of JPEG images via a quality relevance map," *IEEE Signal Processing Letters*, vol. 21, no. 2, pp. 155-158, 2014. [Article \(CrossRef Link\)](#)
- [13] L. D. Li, W. S. Lin and H. C. Zhu, "Learning structural regularity for evaluating blocking artifacts in JPEG images," *IEEE Signal Processing Letters*, vol.21, no.8, pp. 918-922, 2014. [Article \(CrossRef Link\)](#)
- [14] R. Hassen, Z. Wang and M. M. A Salama, "Image sharpness assessment based on local phase coherence," *IEEE Transactions on Image Processing*, vol. 22, no. 7, pp. 2798-2810, July, 2013. [Article \(CrossRef Link\)](#)
- [15] L. D. Li, W. S. Lin, X. S. Wang, G. B. Yang, K. Bahrami and A. C. Kot, "No-reference image blur assessment based on discrete orthogonal moments," *IEEE Transactions on Cybernetics*, DOI: 10.1109/TCYB.2015.2392129, 2015. [Article \(CrossRef Link\)](#)
- [16] K. Gu, G. T. Zhai, W. S. Lin, X. K. Yang and W. J. Zhang, "No-reference image sharpness assessment in autoregressive parameter space," *IEEE Transactions on Image Processing*, vol. 24, no. 10, pp. 3218-3231, 2015. [Article \(CrossRef Link\)](#)
- [17] A. K. Moorthy and A. C. Bovik. "A two-step framework for constructing blind image quality indices," *IEEE Signal Processing Letters*, vol. 17, no. 5, pp. 513-516, 2010. [Article \(CrossRef Link\)](#)

- [18] A. K. Moorthy and A. C. Bovik, "Blind image quality assessment: from natural scene statistics to perceptual quality," *IEEE Transactions on Image Processing*, vol. 20, no. 12, pp. 3350-3364, 2011. [Article \(CrossRef Link\)](#)
- [19] M. A. Saad, A.C. Bovik, and C. Charrier, "Blind image quality assessment: a natural scene statistics approach in the DCT domain," *IEEE Transactions on Image Processing*, vol. 21, no. 8, pp. 3339-3352, August, 2012. [Article \(CrossRef Link\)](#)
- [20] A. Mittal, A. K. Moorthy and A. C. Bovik, "No-reference image quality assessment in the spatial domain," *IEEE Transactions on Image Processing*, vol. 21, no. 12 pp. 4695-4708, 2012. [Article \(CrossRef Link\)](#)
- [21] A. Mittal, R. Soundararajan and A. C. Bovik, "Making a "completely blind" image quality analyzer," *IEEE Signal Processing Letters*, vol. 20, no. 3, pp. 209-212, 2013. [Article \(CrossRef Link\)](#)
- [22] L. X. Liu, B. Liu, H. Huang and A. C. Bovik, "No-reference image quality assessment based on spatial and spectral entropies," *Signal Processing: Image Communication*, vol. 29, no. 8, pp. 856-863, 2014. [Article \(CrossRef Link\)](#)
- [23] P. Ye, J. Kumar, L. Kang and D. Doermann, "Unsupervised feature learning framework for no-reference image quality assessment," in *Proc. of IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 1098-1105, June 16-21, 2012. [Article \(CrossRef Link\)](#)
- [24] G. E. Hinton and S. Osindero and Y. W. Teh, "A fast learning algorithm for deep belief nets," *Neural Computation*, vol. 18, no. 7, pp. 1527-1554, 2006. [Article \(CrossRef Link\)](#)
- [25] H. L. Lee, C. Ekanadham and A. Y. Ng. "Sparse deep belief net model for visual area V2," in *Proc. of Advances in Neural Information Processing Systems*, pp. 873-880, December 8-11, 2008. [Article \(CrossRef Link\)](#)
- [26] B. A. Olshausen and D. J. Field, "Emergence of simple-cell receptive field properties by learning a sparse code for natural images," *Nature*, vol. 381, no. 13, pp. 607-609, June, 1996. [Article \(CrossRef Link\)](#)
- [27] L. Zhang, L. Zhang, X. Q. Mou and D. Zhang, "FSIM: a feature similarity index for image quality assessment," *IEEE Transactions on Image Processing*, vol. 20, no. 8, pp. 2378-2386, 2011. [Article \(CrossRef Link\)](#)
- [28] A. Coates, H. L. Lee and A. Y. Ng, "An analysis of single-layer networks in unsupervised feature learning," in *Proc. of the 14th International Conference on Artificial Intelligence and Statistics (AISTATS)*, pp. 215-223, April 11-13, 2011. [Article \(CrossRef Link\)](#)
- [29] A. K. Moorthy and A. C. Bovik, "Visual importance pooling for image quality assessment," *IEEE Journal of Selected Topics in Signal Processing*, vol. 3, no. 2, pp. 193-201, 2009. [Article \(CrossRef Link\)](#)
- [30] C. C. Chang, and C. J. Lin. "LIBSVM: a library for support vector machines," *ACM Transactions on Intelligent Systems and Technology*, vol. 2, no. 3, Article 27, 2011. [Article \(CrossRef Link\)](#)
- [31] Z. Wang, A. C. Bovik, H. R. Sheikh and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *IEEE Transactions on Image Processing*, vol. 13, no. 4, pp. 600-612, April, 2004. [Article \(CrossRef Link\)](#)
- [32] E. C. Larson and D. M. Chandler, "Most apparent distortion: full-reference image quality assessment and the role of strategy," *Journal of Electronic Imaging*, vol. 19, no. 1, pp. 001006:1-21, March, 2010. [Article \(CrossRef Link\)](#)
- [33] N. Ponomarenko, V. Lukin, A. Zelensky, K. Egiazarian, M. Carli and F. Battisti, "TID2008 - a database for evaluation of full-reference visual quality assessment metrics," *Advances of Modern Radioelectronics*, vol. 10, no. 4, pp. 30-45, 2009. [Article \(CrossRef Link\)](#)
- [34] N. Ponomarenko, O. Ieremeiev, V. Lukin, K. Egiazarian, L. Jin, J. Astola, B. Vozel, K. Chehdi, M. Carli, F. Battisti and C.-C. Jay Kuo, "Color image database TID2013: peculiarities and preliminary results," in *Proc. of the 4th European Workshop on Visual Information Processing*, pp. 106-111, June 10-12, 2013. [Article \(CrossRef Link\)](#)
- [35] VQEG, "Final report from the Video Quality Experts Group on the validation of objective models of video quality assessment-Phase II," available online: <http://www.vqeg.org/>, 2003. [Article \(CrossRef Link\)](#)

- [36] W. F. Xue, L. Zhang and X. Q. Mou, "Learning without human scores for blind image quality assessment," in *Proc. of IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 995-1002, June 23-28, 2013. [Article \(CrossRef Link\)](#)
- [37] D. M. Chandler and S. S. Hemami, "VSNR: a wavelet-based visual signal-to-noise ratio for natural images," *IEEE Transactions on Image Processing*, vol. 16, no. 9, pp. 2284-2298, 2007. [Article \(CrossRef Link\)](#)
- [38] D. Jayaraman, A. Mittal, A.K. Moorthy, A.C. Bovik, Objective quality assessment of multiply distorted images, in: *Proc. of Asilomar Conference on Signals, Systems and Computers*, Pacific Grove, CA, pp. 1693-1697, Nov. 2012. [Article \(CrossRef Link\)](#)
- [39] K. Gu, G. T. Zhai, X. K. Yang, and W. J. Zhang, "Hybrid no-reference quality metric for singly and multiply distorted images," *IEEE Transactions on Broadcasting*, vol. 60, no. 3, pp. 555-567, Sept. 2014. [Article \(CrossRef Link\)](#)
- [40] K. Gu, G. T. Zhai, W. S. Lin, and M. Liu, "The analysis of image contrast: From quality assessment to automatic enhancement," *IEEE Transactions on Cybernetics*, 2015, DOI: 10.1109/TCYB.2015.2401732, 2015. [Article \(CrossRef Link\)](#)
- [41] B. J. Chen, H. Z. Shu, G. Coatrieux, G. Chen, X. M. Sun and J. L. Coatrieux, "Color image analysis by quaternion-type moments," *Journal of Mathematical Imaging and Vision*, vol. 51, no. 1, pp. 124-144, 2015. [Article \(CrossRef Link\)](#)



Leida Li received the B.S. and Ph.D. degrees from Xidian University, Xi'an, China, in 2004 and 2009, respectively. From February to June 2008, he was a Research Assistant in the Department of Electronic Engineering, National Kaohsiung University of Applied Sciences, Taiwan. From January 2014 to January 2015, he was a Visiting Research Fellow in the School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore. He is currently an Associate Professor in the School of Information and Electrical Engineering, China University of Mining and Technology, China. His research interests include multimedia quality assessment and image forensics.



Dong Wu received the B.E. degree from Hubei Engineering University, Xiaogan, China, in 2013. He is now a master student in the School of Information and Electrical Engineering, China University of Mining and Technology, China. His research interests include image quality assessment and sparse representation.



Jinjian Wu received the B.E. and Ph.D. degrees in electronic engineering from Xidian University, Xi'an, China, in 2008 and 2014, respectively. From September 2011 to March 2013, he was a research assistant with the school of computer engineering, Nanyang Technological University, Singapore. Since August 2013, he is a postdoctoral research fellow with the school of computer engineering, Nanyang Technological University, Singapore. Currently, he is an Associate Professor with Xidian University. His research interests include visual perceptual modeling, quality assessment, saliency detection, just noticeable difference estimation.



Jiansheng Qian received the B.E. degree from Xidian University in 1985, the Ms. Degree from Xi'an Institute of Optics and Precision Mechanics of CAS in 1988, and the Ph.D. degree from China University of Mining and Technology in 2003. Currently, he is a full professor with the School of Information and Electrical Engineering, China University of Mining and Technology, China. His research interests are image and video processing.



Beijing Chen received the B.S. degree in Mathematics and Applied Mathematics in 2003 from Jiangxi Normal University, the Ms. Degree in Applied Mathematics in 2006 from Zhejiang University, and the Ph.D. degree in Computer Science in 2011 from Southeast University, China. Now he is an Associate Professor in the School of Computer & Software, Nanjing University of Information Science & Technology, China. His research interest includes color image processing, information security and pattern recognition.