Journal of Visual Communication and Image Representation 49 (2017) 204–212

Contents lists available at ScienceDirect

Journal of Visual Communication and Image Representation

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

An efficient and effective blind camera image quality metric via modeling quaternion wavelet coefficients☆

Lijuan Tang[a,b], Leida Li[a], Kezheng Sun[b], Zhifang Xia[c], Ke Gu[d], Jiansheng Qian[a,*]

[a] School of Info. and Cont. Eng., China Univ. of Mining and Technology, China
[b] School of Info. and Elec. Eng., Jiangsu Vocational College of Business, China
[c] The State Info. Center of P.R. China, China
[d] BJUT Faculty of Info. Tech., Beijing University of Technology, China

ARTICLE INFO

Keywords:
Blind image quality evaluation
Camera image quality
Quaternion wavelet transform
Random forest

ABSTRACT

As an extension of Discrete and Complex Wavelet Transform, Quaternion Wavelet Transform (QWT) has attracted extensive attention in the past few years, because it can provide better analytic representation for 2D images. The QWT of an image consists of four parts, i.e., one magnitude part and three phase parts. The magnitude is nearly shift-invariant, which characterizes features at any spatial location, and the three phases represent the structure of these features. This indicates that QWT is more powerful in representing image structures, and thus is suitable for image quality evaluation. In this paper, an efficient and effective Camera Image Quality Metric (CIQM) is proposed based on QWT, which is utilized to describe the intrinsic structures of an image. For an image, it is first decomposed by QWT with three scales. Then, for each scale, the magnitude and entropy of the subband coefficients, and natural scene statistics of the third phase are calculated. The magnitude is utilized to describe the generalized spectral behavior, and the entropy is used to encode the generalized information of distortions. Since the third phase of QWT is considered to be texture feature, the natural scene statistics of the third phase of QWT is used to measure structure degradations in the proposed method. All these features reflect the self-similarity and independency of image content, which can effectively reflect image distortions. Finally, random forest is utilized to build the quality model. Experiments conducted on three camera image databases and two multiply distorted image databases have proved that CIQM outperforms the relevant state-of-the-art models for both authentically distorted images and multiply distorted images.

1. Introduction

In many image processing fields, such as image restoration, image compression and image recommend systems, image quality is the key factor that determines the performance of the whole systems [1]. Furthermore, it is important to evaluate image quality due to its ability to guide and optimize various cloud and remote computing systems [6]. The most trustworthy method of evaluating image quality is subjective rating because human is the end consumer. However, this is difficult to operate in many real applications. Hence, objective image quality assessment has got a huge development [2–5]. In the light of prior knowledge on image distortions, no reference (NR) image quality assessment (IQA) models can be classified into two classes, namely distortion-specific models and general purpose models [1]. Distortion-specific models are designed for known distortion types, such as noise, blockiness [7,10] and sharpness [9], so their usage is limited. By contrast, general purpose models do not assume any prior distortion types, so they can be widely used in practical scenarios. In this paper we focus on NR general purpose IQA.

In the past few years, several blind general purpose image quality metrics have been proposed [8,12–15,21,22]. Moorthy et al. used the wavelet transform to construct the representative feature vector. Then support vector regression (SVR) [46,47] was used to bulid the Blind Image Quality Index (BIQI) [12]. Mittal et al. addressed the Natrual Image Qualtiy Evaluator (NIQE) [13] model, which used a set of features and fitted them to a multivariate Gaussian model (MVG). Then the distance between MVG model and the statistical regularities model of natural scene statistics (NSS) was used to produce the quality score. In [16], the authors utilized the Gaussian Scale mixture model to model the joint statistics of the wavelet coefficients of natural images to produce the Distortion Identification-based Image Verity and INtegrity Evaluation (DIIVINE) metric. In [14], Saad et al. presented the BLind

☆ This paper has been recommended for acceptance by Zicheng Liu.

* Corresponding author.
E-mail address: qianzhangiqa@163.com (J. Qian).

http://dx.doi.org/10.1016/j.jvcir.2017.09.010
Received 26 April 2017; Received in revised form 10 September 2017; Accepted 19 September 2017
Available online 21 September 2017

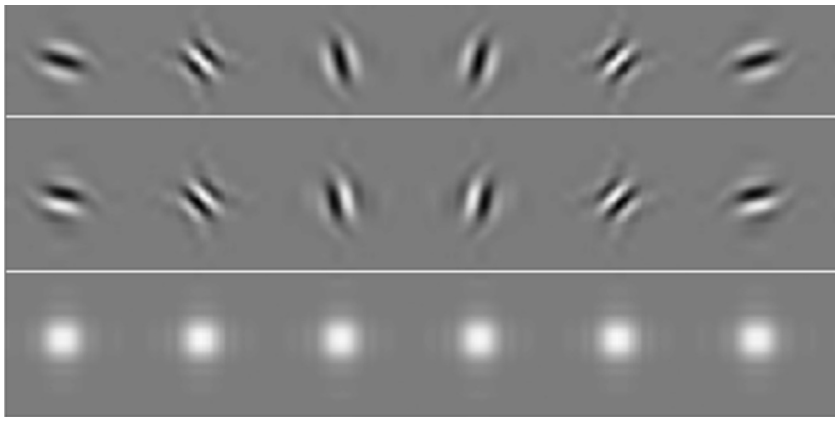

Fig. 1. 2D Dual-Tree Complex Wavelets [26,29,30,34]. Six wavelets are shown in first row are the real part, and the imaginary part of complex wavelets are shown in the second row. The third row shows the magnitude of the six complex.

Image Integrity Notator using DCT Statistic (BLIINDS-II) model based on the features from the NSS model of the image Discrete Cosine Transform coefficients (DCT). Then Bayesian inference model was used to produce the objective score. In [15], the authors addressed the Blind Image Spatial QUality Evaluator (BRISQUE) based on spatial NSS of locally normalized luminance coefficients, which worked under the principle that distortions change the regular statistical properties of natural images. Gu et al. proposed the NR Free Energy based Robust Metric (NFERM) [17] model based on classical human visual system (HVS) properties and free energy. Zhang et al. addressed the Integrated Local Natural Image Quality Evaluator (IL-NIQE) [11], which integrated the features of natural image statistics derived from multiple cues. A Bhattacharyya-like distance was used to measure the quality of each distorted image patch. The Spatial Spectral Entropy-based Quality (SSEQ) model was proposed by Liu et al. [18]. It utilized image entropy features to measure the image structure distortions.

Developing a generalized blind image quality assessment metric is still an open problem. On the one hand, few attempts have been made to estimate camera images distortions [23]. On the other hand, most existing general purpose blind IQA models have limitations, for example, DIIVINE and BLIINDS-II metrics have high computing complexity, BIQI metric generally produces poor performance and so on. Furthermore, although most of aforementioned general purpose IQA models work well on the conventional image database, they do not work consistently well on authentically distorted images. Because most of these image quality models have been devoted to images containing simulated or single distortions, which are not designed for images suffering from different distortion types in practical real scenarios [24].

In this paper, we propose a blind Camera Image Quality Metric (CIQM) via modeling quaternion wavelet coefficients. There are few relevant methods that use QWT for image quality evaluation. Chen et al. proposed a full reference IQA method based on hybrid phase congruency map, where QWT was exploited to extract the quaternion phase congruency map to represent the essential image structures. And complex phase congruency map was used as complementary visual effects of detailed structure on IQA [31]. In [32,33], the authors addressed a reduced-reference metric based on the QWT coefficients from information criteria. The probability density functions was used to model the QWT coefficients as reduced-references. Information criteria was used to obtain the number of distribution. Reduced-reference IQA metric was obtained by comparing the probability density functions of the reference image and the distributions of the distorted image of the QWT subbands. Different from the previous works, which mainly deal with synthetic distortions, the proposed method analyzes the QWT coefficients to evaluate the quality of authentically distorted images. Furthermore, the substantial high performance of the proposed model is evaluated on both real distorted image databases and multiply distorted image databases.

The remainder of this paper is organized as follows. We provide the theoretical basis of the QWT in Section 2. Section 3 presents details of the proposed blind camera quality evaluation algorithm. Results of the CIQM algorithm on authentically distorted image databases and multiply distorted image databases are shown in Section 4. Section 5 gives the general conclusions.

2. Quaternion wavelet transform

Real-world images typically consist of smooth regions and edges, and they can be well characterized by their singularity structure. The Discrete Wavelet Transform (DWT) is a multi-resolution analysis tool that can be utilized to process singularity-rich signals [25].

Since DWT is lack of shift invariance and signal location information, researchers have made a lot of efforts to develop Complex Wavelet Transforms (CWT) to remedy these drawbacks. In [27], the authors proposed the 2D dual-tree CWT using the 1D Hilbert Transform of conventional 2D real DWT wavelets in different directions. Six complex wavelets generated by 2D Dual-Tree complex wavelet transform are shown in Fig. 1. However, each 2D dual-tree basis coefficient has only one angle phase, which is insufficient for analyzing 2D signals [28]. To overcome this drawback, quaternion wavelet transform was proposed [26]. There are several ways to implement QWT. There are great differences between the traditional quaternion transform and the used QWT. The traditional quaternion is an expansion of complex, namely hyper complex. And pure quaternion quaternion is widely used to represent color image, where the three channels of the color image (red, green and blue) are represented by the three imaginary parts. Hence, it is mainly used to handle color images. However, in our proposed method, we adopt the dual-tree QWT proposed by Chan [36]. An extension of Hilbert Transform and analytic signal is used to extend the 1D Complex Wavelet Transforms to 2D. The imaginary and real parts of the signal are a linear combination of the original signal and its partial and complete Hilbert transforms. Let $y = (y1,y2)$ be the real 2D signal, the pair of complex signals generated by the complete 2D complex analytic signal in space domain is defined as [26]:

$$f_1(y) = [f(y) - f_{Hi}(y)] + [f_{Hi1}(y) + f_{Hi2}(y)]j \tag{1}$$

$$f_2(y) = [f(y) + f_{Hi}(y)] + [f_{Hi1}(y) - f_{Hi2}(y)]j \tag{2}$$

where

$$f_{Hi}(y) = f(y) * \frac{1}{\pi^2 y_1 y_2} \tag{3}$$

$$f_{Hi1}(y) = f(y) * \frac{\delta(y_2)}{\pi y_1} \tag{4}$$

$$f_{Hi2}(y) = f(y) * \frac{\delta(y_1)}{\pi y_2} \tag{5}$$

where f is a 2-D real value function, f_{Hi} is the Hilbert transform, f_{Hi1} and

f_{Hi2} are the partial Hilbert transforms, $*$ deNotes 2D convolution. According to [26], in the spatial domain, the quaternionic analytic signal of a real signal y can be defined as:

$$f_A^q(y) = f(y) + if_{Hi1}(y) + jf_{Hi2}(y) + kf_{Hi}(y) \tag{6}$$

The dual-tree QWT is a 4× redundant tight frame with horizontal, vertical, and diagonal subband. The components of QWT wavelets can be organized as the following matrix form:

$$F = \begin{bmatrix} \varphi_h(x)\varphi_h(y) & \varphi_h(x)\psi_h(y) & \psi_h(x)\varphi_h(y) \\ \varphi_g(x)\varphi_h(y) & \varphi_g(x)\psi_h(y) & \psi_g(x)\varphi_h(y) \\ \varphi_h(x)\varphi_g(y) & \varphi_h(x)\psi_g(y) & \psi_h(x)\varphi_g(y) \\ \varphi_g(x)\varphi_g(y) & \varphi_g(x)\psi_g(y) & \psi_g(x)\varphi_g(y) \end{bmatrix} \tag{7}$$

where $\varphi_h(x)$ and $\psi_h(x)$ denote the wavelet and scaling functions, respectively. The wavelet functions $\varphi_h(x)$ and $\varphi_g(x)$ from two trees paly the role of the real and imaginary parts of a complex analytic wavelet. The imaginary wavelet $\varphi_g(x)$ is the 1-D HT of the real wavelet $\varphi_h(x)$. Every quaternion wavelet comprises four components that are 90^0 phase shifts of each other in the vertical, horizontal, and both directions. Each column of F corresponds to the four parts of QWT. Three quaternion wavelets generated from the 2D dual-tree quaternion wavelet transform are shown in Fig. 2. The dual-tree QWT has three phases which encode the shifts of image features in horizontal and vertical direction, and edge orientation mixtures and texture information [26]. Hence, QWT has been widely used in multi-scale image analysis and processing [31–33,35,36].

3. The proposed blind quality metric

In our work, a novel dual-tree QWT based blind camera image quality assessment metric is proposed. Fig. 3 shows the flowchart of the proposed metric. First, the input image I is decomposed by the quaternion wavelet transform into wavelet subband coefficients W_i. According to a large number of experiment results, the test image is decomposed into 3 scales which can acquire the best performance and meantime maintain low complexity (The details will be discussed in Section 4.1). Hence, we use QWT with 3 scales decomposition. Then, for each scale, the magnitude, entropy and the natural scene statistics of the third phase of the subbands are calculated to encode the image. The magnitude is utilized to describe the generalized spectral behavior. And the entropy and natural scene statistics of third phase are used to encode the generalized information of distortions. All of these properties reflect the self-similarity and independency of image content. Finally, random forest is utilized to predict the objective score of the input image.

It has been demonstrated that QWT provides the richest scale-space analysis with one near shift-invariant magnitude and three phases, which can separate the information contained in the image better than the classical DWT [31–33]. The QWT of an image at each point has one magnitude and three phases. We first analyze the magnitude and entropy of the QWT coefficients. The following equations are used to compute the magnitude and entropy:

$$Mag_s = \frac{1}{M_s N_s} \sum_{j=1}^{M_s} \sum_{i=1}^{N_s} \log_2(1 + |W_s(i,j)|), \tag{8}$$

$$Ent_s = \sum_{j=1}^{M_s} \sum_{i=1}^{N_s} p|W_s(i,j)||\ln p|W_s(i,j)|, \tag{9}$$

where M_s,N_s stand for the width and length of the s-th subband, respectively. $W_s(i,j)$ denotes the QWT coefficient of (i,j), and $p(\cdot)$ is the probability function of the QWT coefficients.

Fig. 4 shows distortion free images from CSIQ database [39]. And Fig. 5 shows the magnitudes and entropies of the QWT coefficients of the images in Fig. 4. It is observed from Fig. 5 that the magnitudes and entropies of distortion free images have similar characteristics.

To illustrate the impact of magnitude and entropy of the QWT coefficients on image quality, we give an example in Fig. 6. Fig. 6 shows four different distorted images and their magnitudes and entropies of the QWT coefficients. From Fig. 6, it is easily found that with the increase of distortion degree, the greater the curve deviates from that of the original image. Hence, the magnitude and entropy of QWT can be used to characterize the quality degradation of an image.

For an image, Dual-Tree based QWT coefficients are approximately analytic, so the magnitude $|q|$ and three phases (ϕ,θ,ψ) are approximately analytic [33]. The phase ϕ and θ provides small shifts of the structure in the horizontal and vertical directions, respectively. From the measure of divergences, it can be found that these two phases give a poor detection of degradation [33]. Because very little degradation can introduce these shifts, the measure is instable. However, the phase ψ describes image texture details and achieves a more robust description of an image. It is very useful for image structure analysis. It is well known that image structure plays a key role in image quality evaluation. These findings are similar to the study in [32,33]. Hence, we conduct extensive experiments for the phase ψ.

According to the study of the phase ψ on the whole CSIQ image database [39], we find that the phase ψ of undistorted natural images shows Gaussian characteristic. Furthermore, the distribution changes when the images encounters different types of distortions, as illustrated in Fig. 7. It was found that a Generalized Gaussian Distribution (GGD) can capture the distorted image statistics [15]. In this work, we utilize the GGD to model the coefficients of these phases. The GGD function is

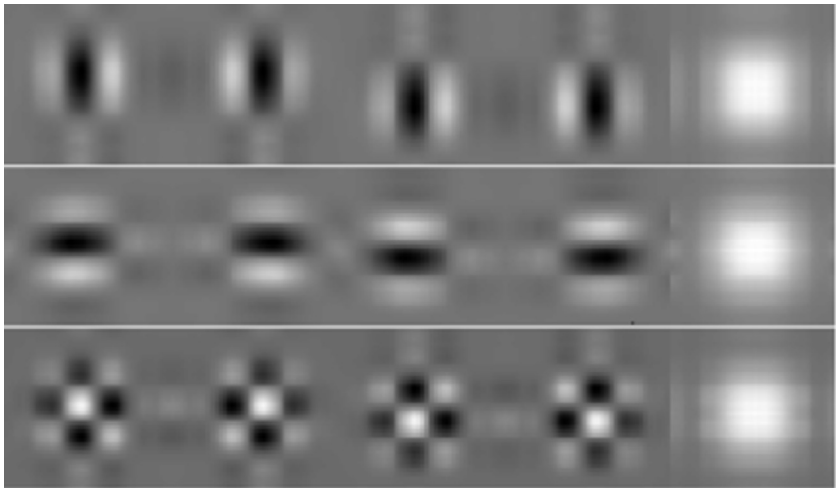

Fig. 2. Quaternion wavelets from the 2D dual-tree quaternion wavelet transform frame [36]. The first row shows the vertical subband, and the second row shows the horizontal subband, and the third row shows the diagonal subband. The last column is the quaternion wavelet magnitude of each subband.

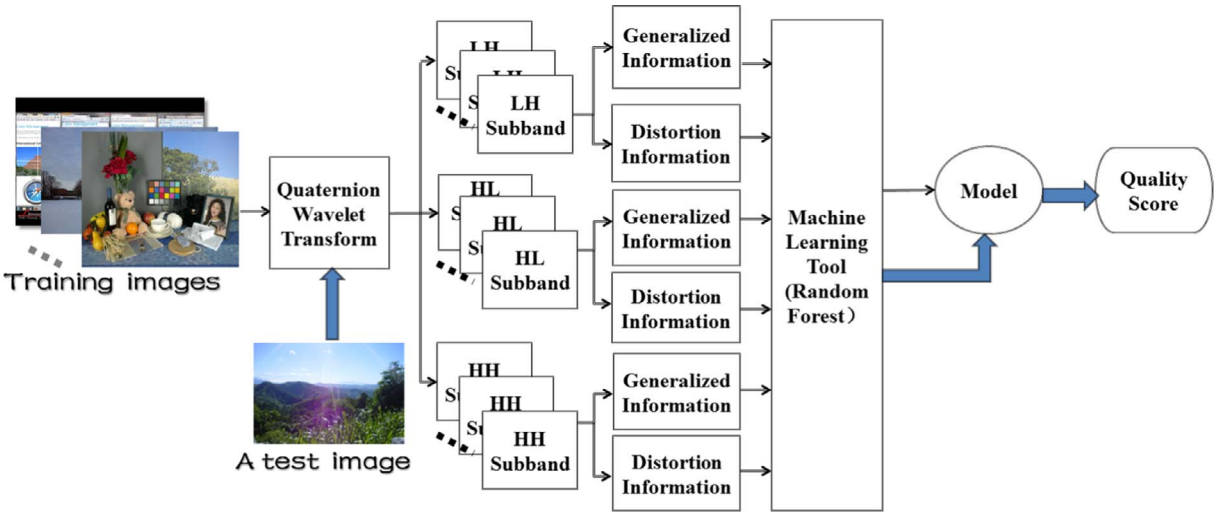

Fig. 3. Flowchart of the proposed blind camera image quality metric.

Fig. 4. The Distortion free 30 images from the CSIQ database [39].

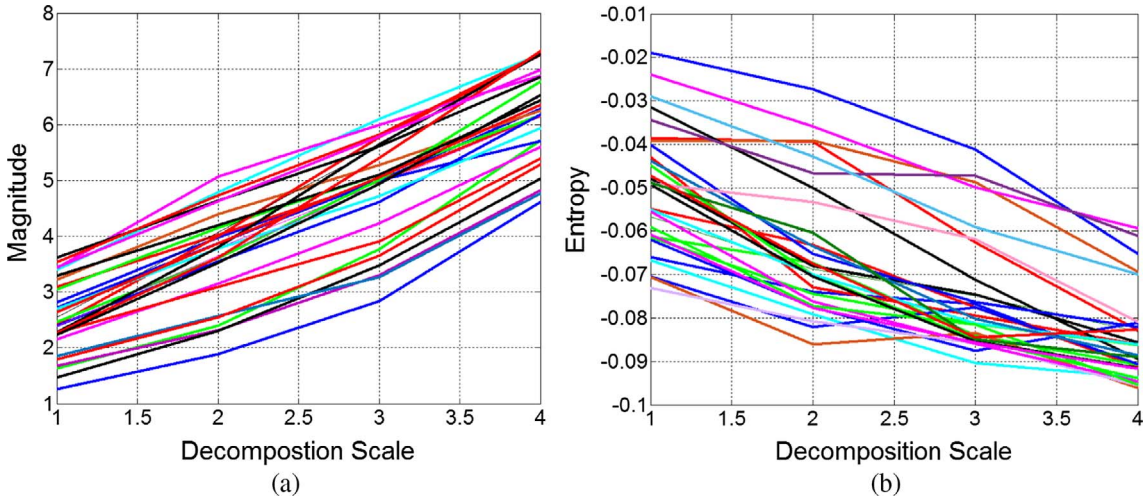

Fig. 5. Magnitudes and entropies of the QWT coefficients of distortion free images from CSIQ database [39]. (a) The magnitudes of 30 original images from CSIQ database. (b) The entropies of 30 original images from CSIQ database.

defined as follows:

$$f(y;\mu,\alpha,\beta) = \frac{\alpha}{2\beta\Gamma(\frac{1}{\alpha})}\exp\left(-\left(\frac{|y-\mu|}{\beta}\right)^{\alpha}\right) \qquad \alpha > 0 \tag{10}$$

where α is the shape parameter that controls the distribution, μ denotes the mean and

$$\beta = \sigma\sqrt{\frac{\Gamma\left(\frac{1}{\alpha}\right)}{\Gamma\left(\frac{3}{\alpha}\right)}} \tag{11}$$

and the gamma function $\Gamma(\cdot)$ is given by:

$$\Gamma(z) = \int_0^{\infty} t^{z-1}e^{-t}dt \tag{12}$$

where the parameter stands for standard deviation. In this model, we utilize the zero mean GGD model to fit the ψ phase of QWT coefficients. The zero mean GGD is given by:

$$f(y;\alpha,\beta) = \frac{\alpha}{2\beta\Gamma\left(\frac{1}{\alpha}\right)}\exp\left(-\left(\frac{|y|}{\beta}\right)^{\alpha}\right). \tag{13}$$

For each decomposition scale, a pair of parameters (α,σ^2) from a GGD fit of the ψ phase of QWT coefficients can be obtained.

The input image is initially decomposed by QWT into subband wavelet coefficients. High-high, low-high and high-low subbands are

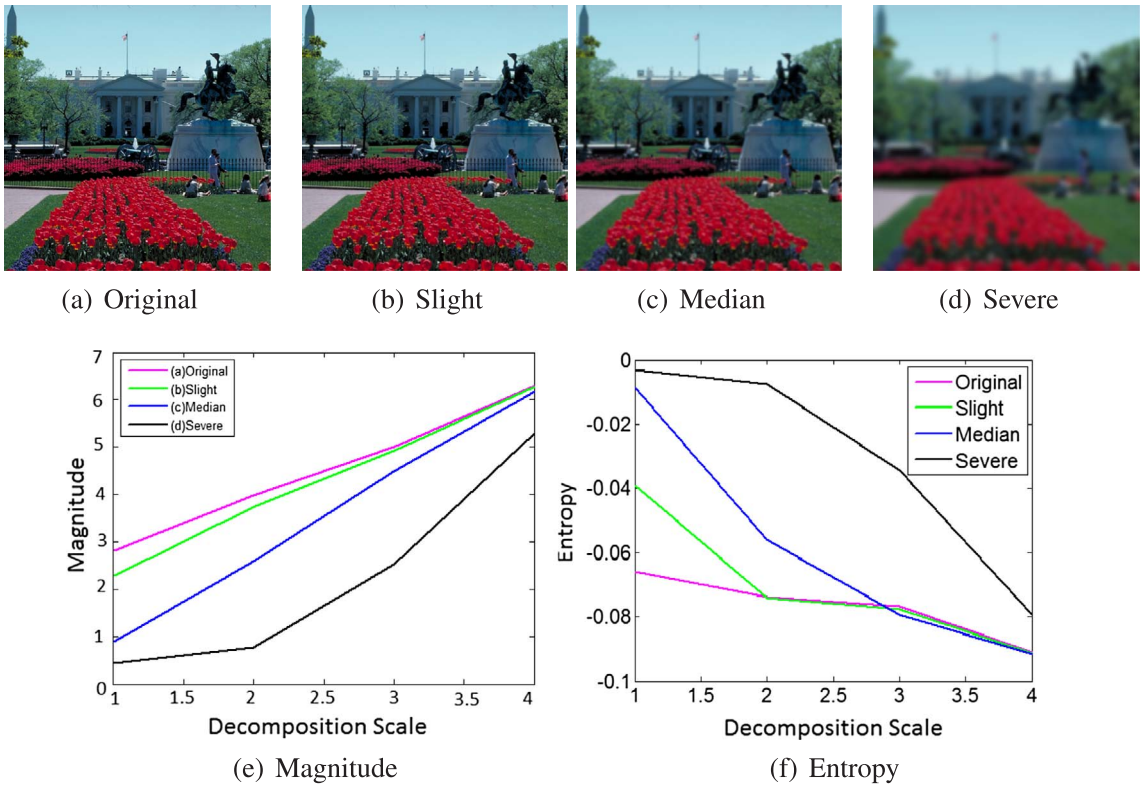

Fig. 6. Four images with different levels distortion together with their magnitudes and entropies of the QWT coefficients. (a)–(d) Denote original image and its different levels of distortion images, (e) denotes their magnitudes and (f) denotes their entropies, respectively.

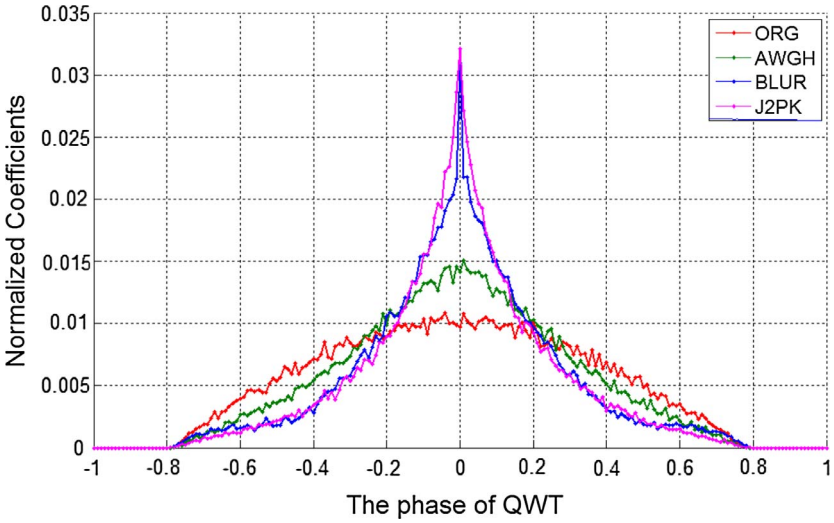

Fig. 7. Histogram of the ψ phase for the natural undistorted image, child_swimming, and its different distorted versions. ORG stands for the undistorted image. AWGN denotes additive white Gaussian noise. BLUR stands for Gaussian blur. J2PK denotes JPEG2000 compression.

used to calculate the features. Due to the similarity in the statistics of low-high and high-low subbands at the same scale, we do not distinguish low-high and high-low subbands at the same scale. By decomposing an image with 3 scales, we have 6 subbands in total. For each subband, the magnitude and entropy are calculated. The phase ψ describes image texture details and it gives a robust description of an image. So we use the phase ψ as feature to describe the image. For each decomposition scale, a pair of parameters (α,σ^2) from a GGD fit of the ψ phase of QWT coefficients is generated. In our algorithm, we use 3 scales decomposition for each image to describe degradation characteristics. Hence, three ψ phase features are used to describe image texture details and three pair of parameters (α,σ^2) are used to capture the ψ phase distortion information for an image. Hence, 21 features in total, including 6 magnitude features, 6 entropy features and 9 phase statistics features, are used to train the quality model.

After the feature extraction, random forest (RF) is utilized to learn the regression model to predict quality score [37]. Considering a set of extracted feature vector $w = \{w_1,\cdots,w_n\}$ and s is the predicted score of the test image, the training objective function of the i-th node of the t-th decision tree $t \in \{1,\cdots,T\}$ is defined as:

$$\theta_i^* = \underset{\theta_i \in T_i}{\operatorname{argmax}}\, G_i \tag{14}$$

where T_i controls the amount of randomness to train node i, and the G_i is defined as:

$$G_i = \sum_{w \in P_i} \log(|\Lambda_s(w)|) - \sum_{j \in \{L,R\}} \Big(\sum_{w \in P_i^j} \log(|\Lambda_s(w)|) \Big) \tag{15}$$

where Λ_s is the conditional covariance matrix computed from probabilistic linear fitting, P_i denotes the amount of training data for node i, and P_i^L, P_i^R stands for the left and right partition sets, respectively. Then, the predicted score $\hat{s}$ is computed by averaging the outputs of T regression trees as:

$$\hat{s} = \frac{1}{T} \sum_{t=1}^{T} p_t(s|w) \tag{16}$$

4. Experimental results and analysis

4.1. Experimental settings

The experiments are conducted on the recently released Camera Image Database (CID) [41], Realistic Blurred Image Database (RBID) [40] and LIVE in the Wild Image Quality Challenge Database (LIVEW) [42], which contain authentically distorted images. The CID database contains 474 distorted images from 6 image sets under 36 scenarios, and these images are captured by 79 different digital cameras. The images in CID contain the typical distortions which depend on optics, built in image processing pipeline and camera sensor type. The mean opinion scores of the images in CID database are ranging from 0 to 100. The RBID database consists of 586 images with resolutions ranging from 2816 × 2112 to 640 × 480 pixels, which contains not only typical and easy to model blur cases, but also complex and authentic ones. The images in RBID database are captured by a various of camera apertures, scenes, varying exposure times and lighting conditions. And mean opinion scores of the images in RBID database are ranging from 0 to 5. The LIVEW database contains 1162 authentically distorted images captured by modern mobile cameras. The images in LIVEW contains various realistic distortions, such as underexposure, overexposure, compression errors, motion blur, blur and noise, and their mixtures. And mean opinion scores of the images in LIVEW database are ranging from 0 to 100.

In this paper, three criterions are adopted for performance measure, including Pearson's correlation coefficient (PLCC), Spearman's rank ordered correlation coefficient (SRCC) and root mean square error (RMSE) [48]. The prediction monotonicity is measured by SRCC. And the prediction accuracy is measured by PLCC and RMSE. Before computing these values, a nonlinear fitting function with five parameters is employed to map objective quality scores to ground truth subjective scores [48]:

$$f(y) = \lambda_1 \left(\frac{1}{2} - \frac{1}{1 + e^{\lambda_2 (y - \lambda_3)}} \right) + \lambda_4 y + \lambda_5 \tag{17}$$

where y denotes the predicted objective score; $f(y)$ stands for the truth subjective score; λ_i $\{i = 1,\cdots,5\}$ are the parameters to be fitted.

Since the proposed method is based on QWT, it is necessary to decide the optimal decomposition scale of QWT. Hence, we conducted experiments on blur subset of CSIQ database [39] to check how the method performs with the change of QWT with different decomposition scales. We test QWT with the decomposition scale from 1 to 7 and Fig. 8 shows that the SRCC values increase with the decomposition scale from 1 to 3. Then with the decomposition scale increased, SRCC values decrease inversely. Based on this information, the proposed method employs QWT with 3 scales, which is applied in the subsequent experiments.

4.2. Comparison with authentically distorted and general purpose blind IQA models

Limited works have been dedicated to assess the quality of

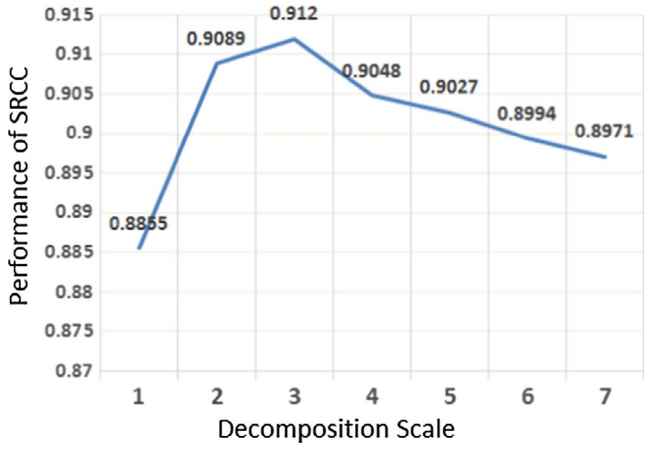

Fig. 8. The computed curves of SRCC with different decomposition scales of QWT.

naturally-distorted image so far. Saad et al. proposed an objective consumer device photo quality evaluation metric, which utilized NSS modeling and the consumer-centric, quality aware interpretable features for image quality prediction [20]. Zhu et al. conducted a study on common, camera-specific kinds of distortions and proposed a blind IQA metric for photographic images produced by consumer devices [19].

In this section, the proposed CIQM model is compared with the above two metrics [19,20], which are designed for authentically distorted image quality assessment, as well as eight state-of-the-art general purpose blind IQA metrics, including BIQI [12], NIQE [13], BLIINDS-II [14], BRISQUE [15], DIIVINE [16], NFERM [17], SSEQ [18] and IL-NIQE [11]. Since CIQM adopts RF regression model for quality prediction, each database is randomly split into two parts: 80% images for model training, 20% for testing. To avoid bias, this process is conducted 1000 times and the median values are reported. All the source codes of compared blind IQA metrics were obtained from the authors or web sites. Since most of the compared metrics are based on SVR and they use LIVE database [38] to train the model, for fairness, the compared metrics are all retrained on CID, RBID and LIVEW databases before testing. Table 1 lists the experimental results on the three databases that contain realistic distortion images, and the top two performance values are marked in boldface. From Table 1, it can be observed that the proposed CIQM method achieves the best prediction monotonicity and accuracy on CID and RBID database. On LIVEW database, the proposed model is slightly worse than the BRISQUE metric. From the results, we can draw the conclusion that the proposed CIQM is effective for authentically distorted images.

4.3. Extension to evaluate multiple distortions

As a method specifically designed for authentically distorted images, the proposed model is expected to perform well on multiply distorted images. To this end, the proposed CIQM model is further tested on multiply distorted images. Recently, several multiply distorted image databases have been created, such as Waterloo exploration database [43], multiply distorted image database (MDID) [45] and LIVEMD database [44]. The Waterloo exploration database [43] contains 4744 pristine natural images and 94880 distorted images created from them. And the Waterloo exploration database uses three innovative evaluation criteria, the pristine/distorted image discriminability test (D-test), the listwise ranking consistency test (L-test) and the pairwise preference consistency test (P-test) to evaluate the relative performance of IQA models instead of collecting the mean opinion score for each image via subjective testing. It is impossible to compare the proposed metric with the other image quality assessment algorithms using PLCC, SRCC and RMSE on Waterloo exploration database. Therefore, in this part, the proposed metric is only tested on MDID [45] and LIVEMD [44] multiply distorted databases.

The LIVEMD database contains 15 pristine images and 450 multiply distorted images of two types, including blur followed by JPEG and blur followed by noise. Hence, the distortions in LIVEMD database are simulated. But the distortions in the other three camera image quality databases are authentic, which are generally more complex, diverse and multipartite, and thus the visual quality assessment requires more efforts in the community. The difference mean opinion scores of the images in LIVEMD database are ranging from 0 to 100.

The MDID database contains 20 reference images and 1600 distorted images. The distorted images are obtained by contaminations of the reference images with multiple distortions of random types and levels, so multiple types of distortions are present in each distorted image. Table 2 summarizes the experimental results on the LIVEMD and MDID databases, where the best performance values are marked in boldface. It is known from Table 2 that the proposed CIQM model significantly outperforms the general purpose IQA metrics. From Tables 1 and 2, we can draw the conclusion that the proposed CIQM is effective for both authentically distorted images and multiply distorted images.

4.4. Effects of training image numbers and machine learning tools

To test the impacts of training image numbers and regression methods on the performance of the proposed metric, experiments are conducted using different numbers of images and different training methods. In this part, we use 80%, 60%, 50% and 40% four different percentage of training images, and SVR, RF two different regression models for training model. The experimental results are listed in Table 3. It is observed from Table 3 that when using different regression methods, the performances are similar. However, with the decreasing number of training images, RF has better stability than SVR. For CID database, the SRCC values of SVR change from 0.8007 to 0.7369, but

Table 1
The SRCC, PLCC and RMSE results of the proposed CIQM model and the compared metrics on authentically distorted image databases. The top two metrics are highlighted in boldface.

Database		CID [41]			RBID [40]			LIVEW [42]		
IQA Metric	Type	PLCC	SRCC	RMSE	PLCC	SRCC	RMSE	PLCC	SRCC	RMSE
Zhu [19]	Realistic	**0.8025**	**0.7795**	**13.4658**	0.4266	0.4013	1.1323	0.4162	0.3512	18.4553
Saad [20]	Realistic	0.6353	0.6146	17.4836	0.3733	0.4013	1.1614	0.2633	0.2642	19.6090
BIQI [12]	Simulated	0.7741	0.7460	13.8275	0.5983	0.5732	0.9951	0.5324	0.5077	17.1349
NIQE [13]	Simulated	0.6640	0.6517	16.8732	0.4601	0.4566	1.1116	0.4966	0.4510	17.6171
BLIINDS-II [14]	Simulated	0.7152	0.7010	15.2091	0.5477	0.5285	1.0434	0.4973	0.4619	17.5425
BRISQUE [15]	Simulated	0.7512	0.7502	14.1831	**0.6083**	**0.5834**	**0.9927**	**0.6346**	**0.6107**	**15.6238**
DIIVINE [16]	Simulated	0.6089	0.5783	17.2475	0.4787	0.4369	1.0913	0.3511	0.2267	18.9278
NFERM [17]	Simulated	0.7677	0.7572	14.0283	0.5642	0.5455	1.0255	0.5877	0.5489	16.325
SSEQ [18]	Simulated	0.7036	0.6903	15.3965	0.5999	0.5754	0.9946	0.5390	0.4962	17.0352
IL-NIQE [11]	Simulated	0.4199	0.3023	20.4806	0.5032	0.4861	1.0819	0.5040	0.4392	17.5149
CIQM (Pro.)	Realistic	**0.817**	**0.8049**	**12.8871**	**0.6373**	**0.6304**	**0.9579**	**0.6186**	**0.5831**	**15.907**

Table 2
The SRCC, PLCC and RMSE results of the proposed CIQM model on the multiply distorted image databases LIVEMD [44] and MDID [45]. We highlight the top one metric in boldface.

Database	LIVEMD [44]			MDID [45]		
IQA Metric	PLCC	SROCC	RMSE	PLCC	SRCC	RMSE
BIQI [12]	0.8863	0.8808	8.4077	0.7088	0.7085	1.5429
NIQE [13]	0.8377	0.7725	10.3292	0.6712	0.6503	1.6333
BLIINDS-II [14]	0.8922	0.8717	8.4645	0.7608	0.7481	1.4339
BRISQUE [15]	0.9048	0.884	7.9915	0.7924	0.7756	1.3491
NFERM [17]	0.9242	0.8992	7.0985	0.8003	0.7958	1.3202
DIIVINE [16]	0.7901	0.822	11.5036	0.4986	0.4991	1.9048
SSEQ [18]	0.8541	0.817	9.7414	0.7689	0.7606	1.4131
IL-NIQE [11]	0.8365	0.8777	10.3624	0.7240	0.6884	1.5200
CIQM (Pro.)	**0.9259**	**0.9063**	**7.0857**	**0.8825**	**0.8744**	**1.0321**

the values of RF only change from 0.8049 to 0.7620. Hence, in our proposed method, we adopt RF to train the regression model. Even only 40% images are used to train the model, the performance of the proposed method is still very good. For example, the SRCC values of CID database can reach 0.7620 and 0.7369, and the SRCC values of RBID database can reach 0.5701 and 0.5439. These results are even better than most of existing general purpose bind IQA metrics.

4.5. Contribution of each component

The proposed CIQM method consists of three kinds of features, including magnitude, entropy, and the phase statistics. It is important to test the contribution of each kind of features. To this end, we conduct experiments on these three kinds of features, respectively. The training and testing processes are similar as before. The values of PLCC, SRCC and RMSE are listed in Table 4, which indicates how well each kind of features correlate with the subjective scores.

It is easy to draw important conclusions from Table 4. Each kind of features utilized in CIQM performs well. Moreover, it should be noted that three kinds of features used in CIQM method considering different aspects. The first kind of features is utilized to describe the generalized spectral behavior. The second kind is used to encode the generalized information of the distortion. And third kind is considered to be texture feature, and natural scene statistics of the third phase of QWT is a measure of structure degradation. Hence, better performance can be acquired by integrating all three kinds of features.

Table 4
Contribution of each component used in the CIQM algorithm.

Component	Criterion	CID [41]	RBID [40]	LIVEW [42]	LIVEMD [44]
Magnitude	SRCC	0.6009	0.5260	0.4601	0.8543
	PLCC	0.6130	0.5390	0.4802	0.8894
	RMSE	17.6571	1.0441	17.6936	8.5346
Entropy	SRCC	0.7900	0.6059	0.5463	0.9024
	PLCC	0.8036	0.6346	0.5840	0.9208
	RMSE	13.3510	0.9615	16.4186	7.3230
Phase_NSS	SRCC	0.7289	0.5497	0.3753	0.8180
	PLCC	0.7401	0.5656	0.4072	0.8513
	RMSE	15.0862	1.0244	18.3897	9.8443
All	SRCC	0.8049	0.6204	0.5831	0.9063
	PLCC	0.8170	0.6373	0.6186	0.9259
	RMSE	12.8871	0.9579	15.9070	7.0857

4.6. Computational complexity analysis

In practical application, computational complexity is an important factor for many real image processing systems. Hence, we compare the computational complexity of our proposed methods with the state-of-the-art general purpose blind IQA models. Experiments are conducted using a desktop with Intel(R) Pentium(R) CPU G3250 3.20 GHZ and 10 GB RAM. Table 5 summarizes the feature numbers used by the compared methods, and the average time consumed by each IQA metric and the proposed CIQM model for extracting the features when applied to 100 realistic distortion camera images from the LIVEW database [42]. It is known from Table 5 that the proposed model CIQM uses relatively few features and it performs the best with very low computational complexity.

5. Conclusions

In this paper, we have presented a novel blind camera image quality metric via modeling quaternion wavelet coefficients. The proposed metric solves the problem of quality evaluation of authentically distorted images from three aspects, namely magnitude, entropy and the third phase of QWT. The magnitude is utilized to describe the generalized spectral behavior. The entropy and natural scene statistics of third phase are used to encode the generalized information of distortions. We have compared our CIQM with IQA models designed for authentically distorted images, as well as the state-of-the-art general purpose blind IQA metrics on CID, RBID and LIVEW databases. The

Table 3
Performances of the CIQM algorithm when using different regression models and image numbers.

Train-Test Partition		80–20%		60–40%		50–50%		40–60%	
Database	Criterion	SVR	RF	SVR	RF	SVR	RF	SVR	RF
CID [41]	PLCC	0.8141	0.8170	0.7831	0.7971	0.7671	0.7868	0.7424	0.7731
	SRCC	0.8007	0.8049	0.7793	0.7868	0.7632	0.7760	0.7369	0.7620
	RMSE	12.6411	12.8871	13.8791	13.6203	14.3296	13.9136	14.9979	14.3236
RBID [40]	PLCC	0.6490	0.6373	0.6037	0.6143	0.5845	0.5979	0.5552	0.5810
	SRCC	0.6364	0.6204	0.5939	0.5988	0.5742	0.5861	0.5439	0.5701
	RMSE	0.9466	0.9579	0.9973	0.9863	1.0159	1.0013	1.0394	1.0173
LIVEW [42]	PLCC	0.6200	0.6186	0.5955	0.6013	0.5884	0.5904	0.5734	0.5786
	SRCC	0.6020	0.5831	0.5811	0.5687	0.5740	0.5596	0.5598	0.5492
	RMSE	15.8661	15.907	16.2551	16.2145	16.395	16.3921	16.6360	16.5629
LIVEMD [44]	PLCC	0.8952	0.9259	0.8717	0.9140	0.8648	0.9043	0.8450	0.8953
	SRCC	0.8663	0.9063	0.8392	0.8970	0.8313	0.8886	0.8115	0.8795
	RMSE	8.0523	7.0857	9.2535	7.6858	9.4779	8.0637	10.1027	8.4086

Table 5
The feature numbers and the computational complexity of compared algorithms and our proposed CIQM metric.

IQA Model	Feature No.	Time (s)
BIQI [12]	18	0.047
NIQE [13]	36	0.3329
BLIINDS-II [14]	24	55.894
BRISQUE [15]	36	0.107
NFERM [17]	23	42.879
DIIVINE [16]	88	16.867
SSEQ [18]	12	1.160
IL-NIQE [11]	468	1036.56
CIQM (Pro.)	21	0.265

experiment results have confirmed that our proposed metric can achieve superior performance on images containing realistic distortions. Furthermore, it is also worthy to emphasize that our metric also effective for evaluating the quality of multiply distorted images.

Acknowledgment

This work was supported in part by the Fundamental Research Funds for the Central Universities under Grant 2017XKQY084.

References

[1] W. Lin, C.-C. Jay Kuo, Perceptual visual quality metrics: a survey, J. Vis. Commun. Image Represent. 22 (4) (2011) 297–312.
[2] L. Zhang, Y. Shen, H. Li, VSI: a visual saliency induced index for perceptual image quality assessment, IEEE Trans. Image Process. 23 (10) (2014) 4270–4281.
[3] L. Zhang, H. Li, SR-SIM: a fast and high performance IQA index based on spectral residual, in: Proc. IEEE Int. Conf. Image Process. (ICIP)), 2012, pp. 1473–1476.
[4] L. Zhang, L. Zhang, X. Mou, D. Zhang, A comprehensive evaluation of full reference image quality assessment algorithms, in: Proc. IEEE Int. Conf. Image Process. (ICIP), 2012. pp. 1477–1480.
[5] L. Zhang, L. Zhang, X. Mou, D. Zhang, FSIM: a feature similarity index for image quality assessment, IEEE Trans. Image Process. 20 (8) (2011) 2378–2386.
[6] K. Gu, D. Tao, J. Qiao, W. Lin, Learning a no-reference quality assessment model of enhanced images with big data, IEEE Trans. Neural Netw. Learn. Syst. (2017).
[7] L. Li, Y. Zhou, W. Lin, J. Wu, X. Zhang, B. Chen, No-reference quality assessment of deblocked images, Neurocomputing 177 (2016) 572–584.
[8] J. Qian, D. Wu, L. Li, D. Cheng, X. Wang, Image quality assessment based on multi-scale representation of structure, Digital Signal Process. 33 (2014) 125–133.
[9] L. Li, W. Lin, X. Wang, G. Yang, K. Bahrami, A.C. Kot, No-reference image blur assessment based on discrete orthogonal moments, IEEE Trans. Cybern. 46 (1) (2016) 39–50.
[10] L. Li, H. Zhu, G. Yang, J. Qian, Referenceless measure of blocking artifacts by Tchebichef kernel analysis, IEEE Sig. Process. Lett. 21 (1) (2014) 122–125.
[11] L. Zhang, L. Zhang, A.C. Bovik, A feature-enriched completely blind image quality evaluator, IEEE Trans. Image Process. 24 (8) (2015) 2579–2591.
[12] A.K. Moorthy, A.C. Bovik, A two-step framework for constructing blind image quality indices, IEEE Sig. Process. Lett. 17 (5) (2010) 513–516.
[13] A. Mittal, R. Soundararajan, A.C. Bovik, Making a completely blind Image quality analyzer, IEEE Sig. Process. Lett. 20 (3) (2013) 209–212.
[14] M.A. Saad, A.C. Bovik, Blind image quality assessment: a natural scene statistics approach in the DCT domain, IEEE Trans. Image Process. 21 (8) (2012) 3339–3352.
[15] A. Mittal, A.K. Moorthy, A.C. Bovik, No-reference image quality assessment in the spatial domain, IEEE Trans. Image Process. 21 (12) (2012) 4695–4708.
[16] A.K. Moorthy, A.C. Bovik, Blind image quality assessment: from natural scene statistics to perceptual quality, IEEE Trans. Image Process. 20 (12) (2011) 3350–3364.
[17] K. Gu, G. Zhai, X. Yang, W. Zhang, Using free energy principle for blind image quality assessment, IEEE Trans. Multimedia 17 (1) (2015) 50–63.
[18] L. Liu, B. Liu, H. Huang, A.C. Bovik, No-reference image quality assessment based on spatial and spectral entropies, Sig. Process. Image Commun. 29 (8) (2014) 856–863.
[19] Y. Zhu, G. Zhai, K. Gu, Z. Che, No-reference image quality assessment for photographic images of consumer device, in: Proc. IEEE Int. Conf. Acoustics, Speech and Signal Process. (ICASSP), 2016. pp. 1085–1089.
[20] M.A. Saad, P. Corriveau, R. Jaladi, Objective consumer device photo quality evaluation, IEEE Sig. Process. Lett. 22 (10) (2015) 1516–1520.
[21] K. Gu, V. Jakhetiya, J. Qiao, X. Li, W. Lin, D. Thalmann, Model-based referenceless quality metric of 3D synthesized images using local image description, IEEE Trans. Image Process. 99 (2017) 1.
[22] K. Gu, J. Zhou, J. Qiao, G. Zhai, W. Lin, A.C. Bovik, No-reference quality assessment of screen content pictures, IEEE Trans. Image Process. 26 (8) (2017) 4005–4018.
[23] L. Tang, L. Li, K. Gu, Xingming Sun, J. Zhang, Blind quality index for camera images with natural scene statistics and patch-based sharpness assessment, J. Vis. Commun. Image Represent. 40 (2016) 335–344.
[24] M.A. Saad, P. Corriveau, R. Jaladi, Objective consumer device photo quality evaluation, IEEE Signal Process. Lett. 22 (10) (2015) 1516–1520.
[25] S. Mallat, A Wavelet Tour of Signal Processing, Academic, New York, 1998.
[26] T. Bülow, G. Sommer, A novel approach to the 2-D analytic signal, in: Proc. Int. Conf. Computer Analysis of Images and Patterns, 1999, pp. 25–32.
[27] N. Kingsbury, Complex wavelets for shift invariant analysis and filtering of signals, Appl. Comput. Harmonic Anal. 10 (3) (2001) 234–253.
[28] P. Perona, Steerable-scalable kernels for edge detection and junction analysis, Lect. Notes Comput. Sci. 588 (92) (1992) 3–18.
[29] A.F. Abdelnour, I.W. Selesnick, Design of 2-band orthogonal near-symmetric CQF, in: Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing, 2001, pp. 3693–3696.
[30] N.G. Kingsbury, A dual-tree complex wavelet transform with improved orthogonality and symmetry properties, in: Proc. IEEE Int. Conf. Image Process., vol. 2, 2000, pp. 375–378.
[31] Q. Chen, Y. Xu, C. Li, N. Liu, X. Yang, An image quality assessment metric based on quaternion wavelet transform, in: Proc. IEEE Int. Conf. Multimedia and Expo Workshops, vol. 370, 2013, pp. 1–6.
[32] A. Traore, P. Carré, C. Olivier, Quaternionic wavelet coefficients modeling for a reduced-reference metric, Sig. Process. Image Commun. 36 (2015) 127–139.
[33] A. Traore, P. Carré, C. Olivier, Reduced-reference metric based on the quaternionic wavelet coefficients modeling by information criteria, in: Proc. IEEE Int. Conf. Image Process., vol. 36, 2014, pp. 127–139.
[34] E. Bayro-Corrochano, The theory and use of the quaternion wavelet transform, J. Math. Imag. Vision 24 (1) (2006) 19–35.
[35] M. Kadiri, M. Djebbouri, P. Carré, Magnitude-phase of the dual-tree quaternionic wavelet transform for multispectral satellite image denoising, Eurasip J. Image Video Process. 41 (2014) 1–16.
[36] W.L. Chan, H. Choi, R.G. Baraniuk, Coherent multiscale image processing using dual-tree quaternion wavelets, IEEE Trans. Image Process. 17 (7) (2008) 1069–1082.
[37] A. Criminisi, J. Shotton, E. Konukoglu, Decision forests for classification, regression, density estimation, manifold learning and semi-supervised learning, Microsoft Research Technical Report MSR-TR-2011-114 (28).
[38] H.R. Sheikh, Z. Wang, L. Cormack, A.C. Bovik, LIVE image quality assessment database release 2, 2005.
[39] E.C. Larson, D. Chandler, Categorical image quality (CSIQ) database, 2010. < http://vision.okstate.edu/csiq > .
[40] A. Ciancio, A.L.N.T. Da Costa, E.A. da Silva, A. Said, R. Samadani, P. Obrador, No-reference blur assessment of digital pictures based on multifeature classifiers, IEEE Trans. Image Process. 20 (1) (2011) 64–75.
[41] T. Virtanen, M. Nuutinen, M. Vaahteranoksa, P. Oittinen, J. Hakkinen, CID2013: a database for evaluating no-reference image quality assessment algorithms, IEEE Trans. Image Process. 24 (1) (2015).
[42] D. Ghadiyaram, A.C. Bovik, Massive online crowdsourced study of subjective and objective picture quality, IEEE Trans. Image Process. 25 (1) (2016) 372–387.
[43] K. Ma, Z. Duanmu, Q. Wu, Z. Wang, H. Yong, H. Li, L. Zhang, Waterloo exploration database: new challenges for image quality assessment models, IEEE Trans. Image Process. 26 (2) (2017) 1004–1016.
[44] D. Jayaraman, A. Mittal, A. Moorthy, A. Bovik, Objective quality assessment of multiply distorted images, in: Proc. IEEE Int. Conf. Record of the Forty Sixth Asilomar Conf. on Sig., Syst. and Computers (ASILOMAR), 2012, pp. 1693–1697.
[45] W. Sun, F. Zhou, Q. Liao, MDID: a multiply distorted image database for image quality assessment, Pattern Recogn. 61 (12) (2017) 153–168.
[46] B. Schokopf, A.J. Smola, R.C. Williamson, P.L. Bartlett, New support vector algorithms, Neural Comput. 12 (5) (2000) 1207–1245.
[47] C-C. Chang, C-J. Lin, LIBSVM: a library for support vector machines, ACM Trans. Intell. Symp. Technol. 2 (3) (2011) Available: <http://www.csie.ntu.edu.tw/cjlin/libsvm> .
[48] VQEG, Final report from the video quality experts group on the validation of objective models of video quality assessment, Mar. 2000. Available: < http://www.vqeg.org/ > .