Fractal Analysis for Reduced Reference Image Quality Assessment

Yong Xu, Delei Liu, Yuhui Quan, Patrick Le Callet

Abstract—In this paper, multi-fractal analysis is adapted to reduced-reference image quality assessment (RR-IQA). A novel RR-IQA approach is proposed, which measures the difference of spatial arrangement between the reference image and the distorted image in terms of spatial regularity measured by fractal dimension. An image is first expressed in Log-Gabor domain. Then fractal dimensions are computed on each Log-Gabor subband and concatenated as a feature vector. Finally the extracted features are pooled as the quality score of the distorted image using ℓ_1 distance. Compared with existing approaches, the proposed method measures image quality from the perspective of the spatial distribution of image patterns. The proposed method was evaluated on seven public benchmark datasets. Experimental results have demonstrated the excellent performance of the proposed method in comparison with the state-of-the-art approaches.

Index Terms—image quality assessment, similarity of spatial arrangements, fractal dimension, Log-Gabor representation

◆

1 INTRODUCTION

Degradation of visual information is inevitable in our real world, *e.g.*, noise, blurring, ringing, oversaturation, *etc.*, often presents in the processes of image acquisition, compression and transmission. To quantify how much an image is affected by degradation, a metric is required to evaluate how good an image is in visual meaning for human visual system (HVS). This leads to the research on image quality assessment (IQA). The purpose of IQA is to quantitatively model the quality of images, *i.e.*, to assign a value to an image for representing its quality. The estimated image quality can be utilized as a guide for image restoration or applied to image enhancement as a constraint. IQA can be categorized into two types: subjective and objective. As the ultimate solution, subjective IQA has its advantages on the reliability and consistency with the perception of human, for the quality is directly quantified by observers. Nevertheless, the practicality of subjective IQA is very limited because it is expensive and time-consuming. Moreover, the manual process cannot be readily incorporated into the image processing and compression algorithms. Thus, objective IQA is desired and developed, in which the assess-

ment is automatically done with the demand on its consistency with subject assessment.

According to how much information about the reference image is accessible, objective IQA can be further classified into three categories: full-reference (FR), no-reference (NR), and reduced-reference (RR). FR-IQA [1–5] works on the whole reference image and hence yields promising results. However, the reference image is often unavailable in many applications, which largely reduces the applicability of the FR approaches. On the contrary, assuming nothing for reference, NR-IQA [6–8] is an extremely difficult task. To strike the balance of FR and NR, RR-IQA [9–13] requires only partial information of the original image, which is often presented in the form of a set of image features. With much less data to transmit, RR-IQA is flexible in real applications. This paper focuses on RR-IQA.

A useful RR-IQA metric is expected to achieve higher prediction accuracy while using less information of the reference image. Thus, on the way to RR-IQA metric, the challenging task is the selection of the smallest set of features that can support the quality assessment effectively. Therefore, studying and exploiting the special properties of natural images has been one of the most important tasks in RR-IQA. The recent focus of RR-IQA research mainly depends on some statistic models of natural images and has achieved impressive results; see *e.g.* [9–17]. The basic idea of these methods is to quantify image quality by quantifying the disturbance to image statistics caused by distortion. For example, Wang *et al.* [13] modeled natural images by using the marginal distributions of wavelet coefficients, and measured image distortion using the KL-distance between marginal distributions. In [12], the modeling is also done in wavelet domain but is based on the entropies of wavelet coefficients instead of marginal probability distributions. Xue *et al.* [14] employed the Weibull distribution to describe the statistics of image gradient magnitude. In [15], the generalized Gaussian density is employed to model the distribution of the discrete cosine transform coefficients. To

- *Yong Xu, Delei Liu and Yuhui Quan are with the School of Computer Science & Engineering, South China University of Technology, Guangzhou 510006, China. E-mail: yxu@scut.edu.cn, deleiliu@gmail.com, yuhui.quan@mail.scut.edu.cn*
- *Patrick Le Callet is with the IRCCyN Lab, University of Nantes, BP 50609 44306 Nantes Cedex 3, France. E-mail: patrick.lecallet@univ-nantes.fr*
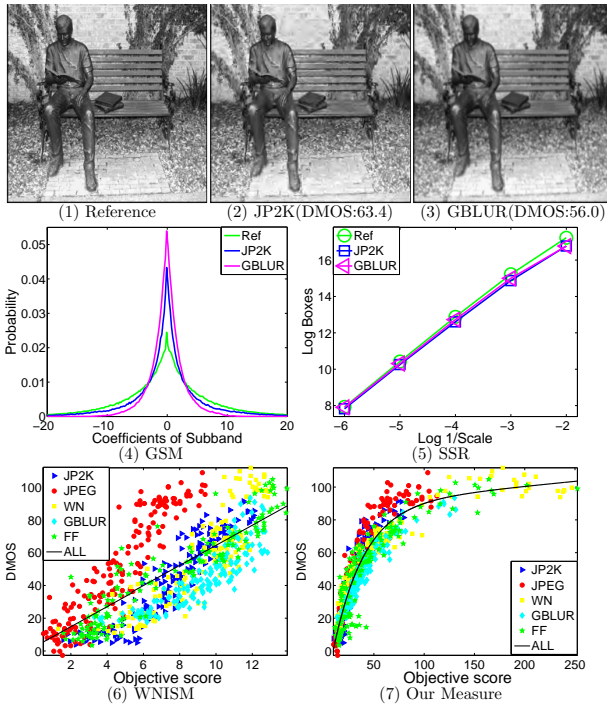- *Corresponding author: Yong Xu*

Fig. 1. The limitation of existing methods. (1)-(3) A test image of the LIVE database [19], from left to right is the reference image, the compressed image by JP2K, and the blured image by Gaussian kernel; (4) Coefficient GSM model based histogram in a wavelet subband; (5) The proposed feature computed in the same subband as (4); (6)-(7) Scatter plot of WNISM and the proposed method on the LIVE database. Note that a smaller differential mean opinion score (DMOS) means a better visual quality of the image. It can be seen that the GSM histogram varies with different distortion types, while the proposed method are quite consistent across different distortion types. It can also be seen that WNISM is biased by the degradation type. In comparison, our method has less bias and can be generalized to a broad spectrum of distortion types.

model the perceptual sensitivity of biological vision, Li *et al.* [11] proposed the so-called divisive normalization transformation (DNT) for image representation. The image statistic is based on the Gaussian scale mixtures (GSM) model. To adapt SSIM [18] to RR-IQA, Rehman *et al.* [1] combined the GSM-based statistics in multi-resolution and multi-orientation DNT domains, following the philosophy in the construction of SSIM. In [9], natural image statistics in transform domain are exploited based on multiscale geometric analysis including curvelets, bandlets, wavelets, and contourlets.

Although the aforementioned methods have achieved good performances in several IQA tasks, some limitations can still be figured out [12]. One limitation is the lack of generality to a broad spectrum of distortion types. Many existing methods either work well only for specific types of distortions or succeed when individual distortion type is tested respectively. As a result, the performance of these methods cannot be guaranteed when these method are applied to dealing with multiple types of distortions. For instance, although WNIS-M [2] works well on individual distortion type, as shown in Fig. 1(6), its performance degrades significantly when images

with different distortion types are evaluated together. The other limitation is that most existing methods are based on histogram-based statistics, and disregard spatial distribution of elements which are helpful to IQA. Figure 1 shows the performance comparison of the histogram-based method and the proposed method. It can be seen that the histogram-based method cannot correctly reflect the degree of the distortion.

We demonstrate in this paper that the above challenges can be mitigated by using Log-Gabor representation and multi-fractal analysis. The former aims to create a complete basis of visual perception, while the latter can encode spatial information in the form of the geometrical distribution of visual data [20–22]. Specifically, the proposed RR-IQA feature, called *spectrum of spatial regularity* (SSR), characterizes the spatial distribution of image structures based on fractal geometry. The spatial-frequency components of image are first extracted by Log-Gabor filtering. Then fractal dimension is used to measure the spatial regularity of the arrangements in each Log-Gabor subband. Finally all the computed fractal dimensions are collected as a feature vector. By using fractal analysis that has a strong correlation with HVS [23], the image structures are well encoded and the difference of their spatial arrangements between images can be well characterized. Unlike previous work based on fractal analysis [24, 25], in our method fractal analysis is working on the visual perceptive space instead of image space due to the fact that HVS is a limited-bandwidth system which is sensitive to specific spatial frequencies [9, 26, 27]. Our approach was evaluated on seven public IQA benchmark databases using five evaluation criteria. The competitive results achieved demonstrate that our method performs on par with the state-of-the-art approaches.

The rest of this paper is organized as follows: Section 2 gives a brief review on the Log-Gabor representation and fractal analysis. A detailed description of the proposed method is given in Section 3. Experimental evaluation is presented in Section 4. Finally, Section 5 concludes the paper.

2 PRELIMINARIES

Before presenting the detailed description of our approach, we first give an introduction to the two mathematical tools upon which our approach is built.

2.1 Log-Gabor Representation

The Log-Gabor filters defined in frequency domain are as follows [28],

$$H(\rho, \theta) = \exp\left(-\frac{(\log(\rho/\rho_0))^2}{2(\log(\sigma_\rho/\rho_0))^2}\right) \cdot \exp\left(-\frac{(\theta - \theta_0)^2}{2\sigma_\theta^2}\right), \quad (1)$$

where (ρ, θ) are the polar coordinates in frequency domain, (ρ_0, θ_0) and $(\sigma_\rho, \sigma_\theta)$ denote the shift and the bandwidth respectively in the polar coordinates of frequency domain. The term σ_ρ/ρ_0 should be held constant to obtain filters with constant shape ratios and its value determines the filter bandwidth. In our implementation, the parameters for generating the Log-Gabor filters are set as: $\rho_0 = 1/5$, $\theta_0 = 0$, $\sigma_\rho/\rho_0 = 0.75$, and $\sigma_\theta = 0.6$. An example of a 2D Log-Gabor filter in the frequency domain is shown in Fig. 2. By applying Log-Gabor
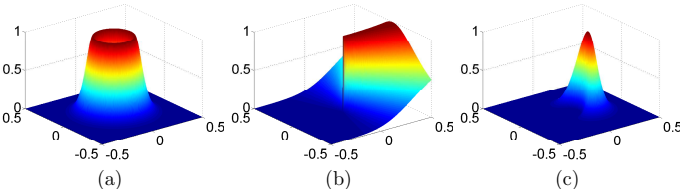
Fig. 2. An example of a Log-Gabor filter in the frequency domain. (a) The radial component of the filter; (b) The angular component of the filter; (c) The Log-Gabor filter, which is the product of the radial component and the angular component.
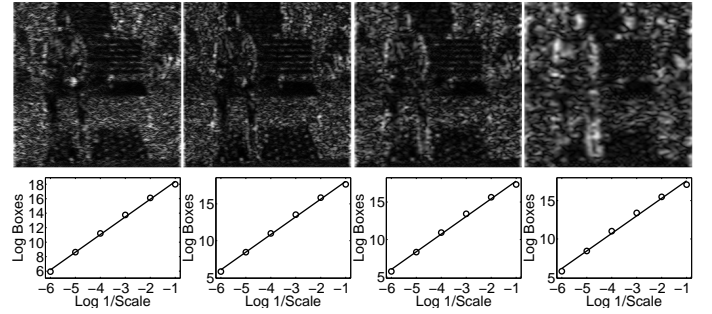

Fig. 3. Log-log plot of box number versus box scale for four Log-Gabor subbands. The upper row shows the coefficients in four Log-Gabor subbands of the reference image in Fig. 1. The bottom row shows the corresponding log-log fittings whose slops are 2.4374, 2.3990, 2.3529, and 2.2883 respectively.

filters to an image, the Log-Gabor representation is obtained. Figure 3 shows the example of Log-Gabor representation.

There are two major reasons to adopt the Log-Gabor representation in this paper. Firstly, the receptive fields of simple cells in the primary visual cortex (V1) of primates can be well modeled by the Log-Gabor functions [28, 29]. In other words, the Log-Gabor functions form a complete basis of visual perception, which provides convenience for defining perceptive features for RR-IQA. Secondly, HVS is a system of limited bandwidth, which is sensitive to specific spatial frequencies [9, 26, 27]. By Log-Gabor filtering, we can well localize different spatial frequency components of image. In comparison with Gabor filters and other wavelet filters, the Log-Gabor filters better fit the statistics of natural images.

2.2 Fractal Analysis

Fractal analysis first developed by Mandelbrot [30] provides a powerful mathematical framework to study the irregular, complex and self-similar objects in nature. Fractals can be viewed as the objects with statistical self-similarities. Most of the intensity surfaces of natural images can be modeled by isotropic fractals [23]. One fundamental concept in fractal analysis is the so-called fractal dimension, denoted by d, which summarizes the irregularity and statistical self-similarity of a given point set E in some measurement space $m(\cdot)$ by measuring its power-law behavior according to the scale δ:

$$m_\delta(E) \propto \delta^{-d},$$

where $m_\delta(E)$ is some measurement of the given point set E at scale δ. For images, the measurement could be intensity, gradients, or other local image features.

There are many techniques to estimate the fractal dimension of image surface. An efficient approach is the differential box counting (DBC) method, which has advantages in terms of efficiency, accuracy and generality [31]. The DBC method considers an image $I(x, y)$ of size $M \times M$ as a 3D point set $\{(x, y, z)|z = I(x, y)\}$, where (x, y) denotes the 2D position and z denotes the gray level of the image. Suppose the image is scaled down to a size $s \times s$, where s is an integer bounded by $1 < s \le M/2$. Let $r = s/M$. The (x, y) space is partitioned into grids of size $s \times s$. A column of boxes of size $s \times s \times h$ are placed on each grid respectively, where h denotes the height of a single box. We generally set the values of h and s to satisfy $G/h = M/s$, where G is the total number of gray levels. Suppose the minimum gray value and the maximum

gray value in the (i, j)-th grid fall in the k-th box and the l-th box respectively. We compute the contribution $n_r(i, j)$ in the (i, j)-th grid as follows,

$$n_r(i, j) = l - k + 1. \tag{2}$$

Summing the contributions from all grids, we have

$$N_r = \sum_{i,j} n_r(i, j), \tag{3}$$

where N_r is counted for different values of r. Then the DBC fractal dimension of the image I is defined as

$$d_{\text{DBC}}(I) = \lim_{r \to 0} \frac{\log N_r}{-\log r}. \tag{4}$$

In practice, d_{DBC} can be estimated by the least squares linear fitting in the $\log N_r$ against $\log 1/r$ coordinates system [31].

Fractal analysis can be adopted to encode Log-Gabor representation. To demonstrate this, we plot the behaviors of the coefficients in four Log-Gabor subbands by log-log fitting in Fig. 3. It can be seen that the Log-Gabor representations of image do behave according to some power law. The advantages of using fractal analysis are as follows. Firstly, fractal dimension has a strong correlation with HVS [23]. Secondly, compared with other statistical approaches, fractal dimension can encode spatial information in the form of the geometrical distribution of image [20, 22]. Specifically, the histogram only counts the number of pixels in a category under one resolution, while the fractal dimension counts the number of patterns in a category under multiple resolutions and estimates the exponential changing ratio of the number of patterns with respect to the resolution. This multi-resolution analysis can significantly mitigate the effect caused by image distortions and enhance the predictive precision in IQA. See an example studied in Fig. 1.

3 OUR METHOD

In this section, we present the proposed method, which is built upon the developed feature called *spectrum of spatial regularity* (SSR). The outline of the proposed method is illustrated in Fig. 4. There are mainly two steps. In the first
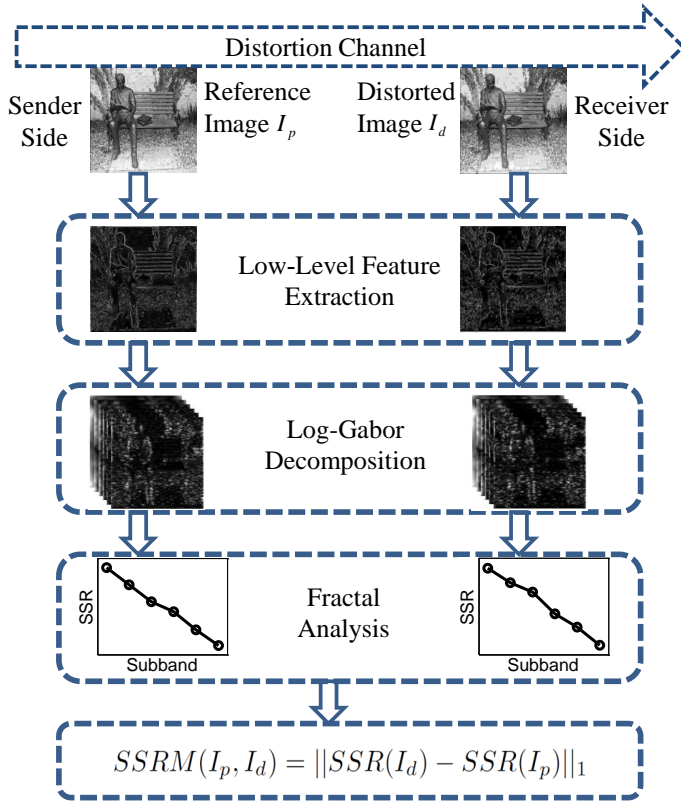
Fig. 4. Outline of the proposed method.

Fig. 5. The SROCC curves of the SSR features computed on seven databases in global and local manners.

step, the SSR features are computed from the reference image and distorted image based on the Log-Gabor representation and fractal analysis. In the second step, the computed features are pooled into a single index using ℓ_1 distance. We will give the details of each step in the following subsections.

3.1 Quality Feature Regarding Spatial Regularity

Given an image I, we first extract low-level visual features from the image. A feature map M is computed from I by

$$M = m(I), \qquad (5)$$

based on some measurement $m(\cdot)$. An accepted measurement should be sensitive to a wide range of image distortion types. In our method, the intensity measurement m_{int} and the gradient measurement m_{grad} are used, which are defined as follows,[1]

$$m_{\text{int}}(I) = I, \qquad (6)$$

$$m_{\text{grad}}(I) = |\nabla I| = \sqrt{\frac{\partial I}{\partial x}^2 + \frac{\partial I}{\partial y}^2}. \qquad (7)$$

The gradient operator ∇ can effectively capture image local structures, to which HVS is highly sensitive. The most commonly encountered image distortions, such as noise corruption, blur and compression artifacts, will lead to highly visible structural changes in the gradient domain. In practice, using image gradient to design IQA models is popular; see *e.g.* [5, 32–34]. By using these two measurements, our approach is able

---

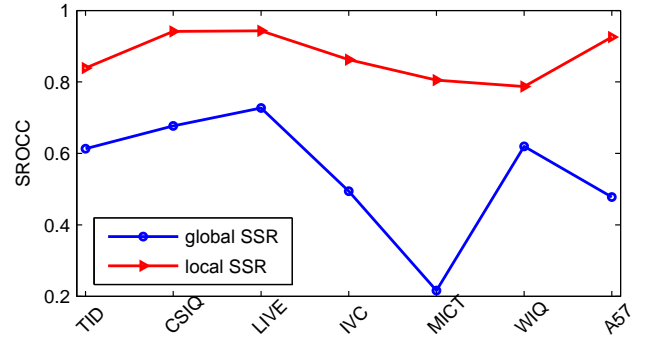1. In our implementation, the Scharr gradient operator [5] is used.

to explore different structures of natural images from different aspects.

In order to simulate the perception of human to image via the V1 cortex of primates, the Log-Gabor filters $\{H(\rho, \theta)\}_{\rho,\theta}$ are run on the feature image M to localize the image structures of different orientations and frequencies. This procedure results in multiple response images $\{G_{\rho,\theta}\}_{\rho,\theta}$ from the feature image:

$$G_{\rho,\theta}(M) = \mathcal{F}(M, H(\rho, \theta)), \qquad (8)$$

where $\mathcal{F}(\cdot, H)$ denotes filtering with the Log-Gabor filter H. By using Log-Gabor filtering, we convert the extracted image features from image space into visual perceptive space. Then, parallel to characterizing the spatial distribution of the visual responses in the V1 area of HVS, we extract features from the Log-Gabor response images via fractal analysis. The fractal dimension is computed on each response image using (4) and all the fractal dimensions are concatenated as the proposed SSR feature, *i.e.*,

$$SSR(I) = \biguplus_{\rho,\theta} d_{\text{DBC}}(G_{\rho,\theta}(m(I))), \qquad (9)$$

where $\biguplus$ denotes the concatenation of all the scalars into a vector. The SSR feature globally summarizes the spatial regularity of image structures of different orientations and frequencies. Thus, it can efficiently reflect the changes of spatial layout of local image features caused by different types of distortions.

In practice, the SSR features can be extracted globally through an image or locally in small blocks. We found that local SSR gives better prediction performance than global SSR, as shown in Fig. 5. This is because local SSR provides a richer description. Using local SSR also facilitate the generation of a quality map for estimating quality in local regions [36]. Figure 6 shows the quality map of a block-distorted image, which describes how a distorted image is deviated from its reference. Inspired by the advantages of local SSR, we adapt the global SSR scheme to a local version as follows. A given image is first divided into non-overlapped subimages, and then the SSR features defined in (9) are computed on each subimage and concatenated as output. In other words, the modified SSR
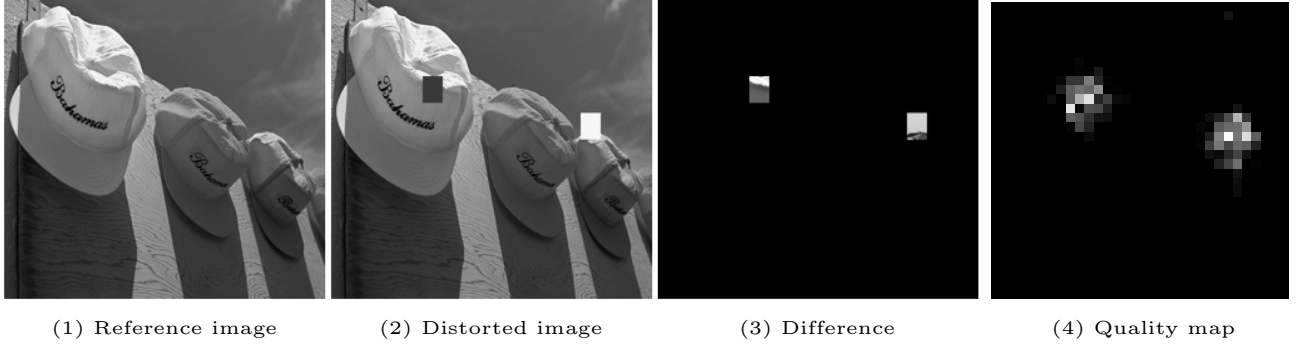
(1) Reference image   (2) Distorted image   (3) Difference   (4) Quality map

Fig. 6. The quality map by the proposed method for an block-distorted image from the TID2008 dtabase [35]. (1) The reference image; (2) The distorted images with a local distortion caused by a block damage; (3) The difference images between the reference image and the distorted image; (4) The quality map, magnified for better reading, where brighter pixels indicate lower quality.

feature is defined as

$$SSR(I) = \biguplus_{i,\rho,\theta} d_{\text{DBC}}(G_{\rho,\theta}(m(P_i(I)))), \qquad (10)$$

where P_i denotes the operation that extracts the i-th subimage. In the remainder of the paper, we refer to the SSR feature as the modified version.

Note that using different measurements results in different SSR features. The SSR features built upon the intensity measurement m_{int} encodes spatial distribution of image brightness while the SSR features built upon the gradient measurement m_{grad} focuses on image edges. It should also be noted that in the scenario where the data transmission rate is vital, the length of a RR-IQA feature should be considered. For this case, we define a reduced version of SSR, denoted as SSR*, by summing the local fractal dimensions of all filter responses over orientations as follows,

$$SSR^*(I) = \biguplus_{i,\rho} \sum_{\theta} d_{\text{DBC}}(G_{\rho,\theta}(m(P_i(I)))) \qquad (11)$$

The summation is based on the fact that the fractals of most natural images are isotropic [23].

3.2 Similarity Index of Spatial Regularity

Once the SSR features of both the pristine image I_p and the distorted image I_d have been obtained, we compute the proposed *similarity of spatial regularity measure* (denoted by SSRM) as the quality score of the distorted image, which is done by calculating the ℓ_1 distance of the two SSR feature vectors as follows,

$$SSRM(I_p, I_d) = ||SSR(I_p) - SSR(I_d)||_1. \qquad (12)$$

Correspondingly, we denote SSRM_INT and SSRM_GRAD as the SSRM computed based on the intensity measurement m_{int} and the gradient measurement m_{grad} respectively. The SSRM measures the difference between the distorted image and the pristine image in the meaning of the regularity of the spatial arrangements of local image features. The SS-RM_INT estimates such difference of regularities regarding image brightness while SSRM_GRAD regarding image edges.

To give an insight on why the proposed method works, we computed the SSRM values with respect to five types of distortion of the reference image in Fig. 1 as well as the average SSRM values over all the reference images in the LIVE database. The results are illustrated in Fig. 7. Note that an IQA metric can be viewed as an excellent metric as long as it monotonously changes with distortion increasing. From Fig. 7 it is seen that the estimated quality score of our SSRM method inclines to rise when the degree of distortion increases. In other words, the SSRM are quite consistent with the image quality. We can see that the SSRM_GRAD method yields a larger range of quality score than the SSRM_INT method for blurring and compression which eliminate edgels. This is not surprising because SSRM_GRAD based on gradients is more sensitive than SSRM_INT to the changes of image edges.

4 EXPERIMENT

We evaluated the proposed method on seven benchmark datasets in terms of five performance measures. This section is devoted to the performance evaluation and experimental analysis. The experiments show that our method is very competitive against the state-of-the-art approaches.

4.1 Benchmark Databases and Test Methodology

There are seven public benchmark databases widely used in the IQA community, including the TID2008 database [35], the CSIQ database [37], the LIVE database [19], the IVC database [38], the MICT database [39], the WIQ database [40] and the Cornel_A57 database [41]. All of them are used for the evaluation of our method. The characteristics of these seven databases, in terms of the number of reference images, the number of distorted images, the number of quality distortion types, and the typical image size, are summarized in Table 1.

The performance is evaluated in terms of five measures, which have been widely used in previous studies:

- the Pearson linear correlation coefficient (PLCC)
- the Spearman rank-order correlation coefficient (SROCC)
- the Kendall rank-order correlation coefficient (KROCC)
- the Root mean square error (RMSE)
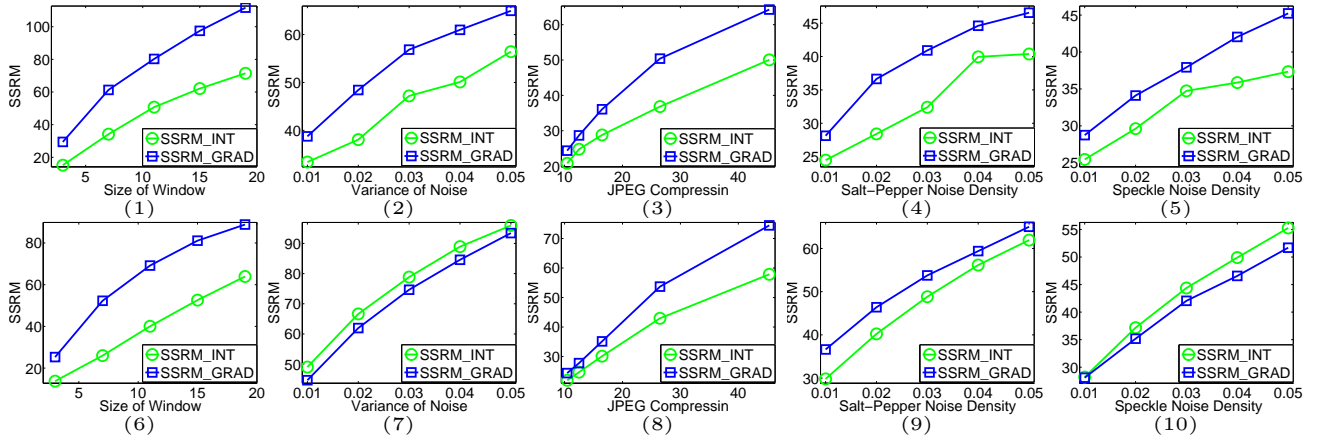- the Mean absolute error (MAE).

Fig. 7. SSRM curves with respective to five distortion types. The first row shows the SSRM curves of the reference image in Fig. 1, and the second row shows the average the SSRM curves of all the reference images in the LIVE database.

TABLE 1
Major characteristics of seven benchmark databases

Database	Ref. Img	Dist. Img	Dist. Types	Typical Size
TID2008	25	1700	17	512×384
CSIQ	30	866	6	512×512
LIVE	29	779	5	768×512
IVC	10	185	4	512×512
MICT	14	168	2	768×512
WIQ	7	80	1	512×512
A57	3	54	6	512×512

Generally speaking, the PLCC, RMSE and MAE criteria are used to measure the prediction accuracy, while the SROCC and KROCC criteria are used to measure the monotonicity. A desirable objective RR-IQA metric is expected to produce high values for the SROCC, KROCC and PLCC criteria, and meanwhile yield low values for the RMSE and MAE criteria.

Four representative RR-IQA metrics with competitive results were selected for comparison, including

- WNISM [2], one of the most successful RR-IQA metrics for general distortions, in which the KL-divergence is used for pooling scores;
- RR-SSIM [1], an extension of WNISM, with an additional DNT step before computing KL-divergence;
- RRED [12], which computes the average difference of scaled local entropies in the wavelet domain between the reference and distorted images;
- HWD2 [9], which decomposes images based on multigeometry analysis and extracts features based on multichannel structures.

Furthermore, we compared our RR-IQA methods with five representative FR-IQA approaches, including

- FSIM [5], a powerful method using phase congruency and gradient magnitude to compute a local similarity maps as well as the weights for pooling quality scores;
- IW-SSIM [42], in which the SSIM values are weighted using the weights derived from image contents;
- VIF [3], which relates image quality to the amount of information shared by the reference and test images;

- SSIM [18], a popular method which exploits the structural sensitivity of human eyes under a mild assumption;
- PSNR, the simplest and most classic FR metric, in which the pixel-wise signal fidelity is measured.

It is noted that not all of these approaches were used for the comparison on all the benchmark databases. We only refer the results that are available.

The objective scores x estimated by the tested method is first mapped to the subjective score $f(x)$ with a nonlinear mapping f, as there is a nonlinear relationship between the objective scores and the subjective ratings; see *e.g.* [43]. Following the protocols of [44], the logistic function is used as the mapping, which is defined as

$$f(x) = \beta_1 \left(\frac{1}{2} - \frac{1}{1 + exp(\beta_2(x - \beta_3))} \right) + \beta_4 x + \beta_5, \quad (13)$$

where the parameters β_i for $i = 1, 2, \ldots, 5$ are determined by least squares fitting.

4.2 Implementation Details

The performance of an IQA metric is related to the scale of image [45]. By a proper normalization on image size, RR-IQA method including SSRM can use less information of the reference image while achieving higher prediction accuracy. However, the image scale is dependant on both the image resolution and the viewing distance and hence is difficult to tune up for optimality. In our implementation, we follow the empirical scale proposed in [45] and normalize all the images to 256×256. In Table 2, we compare the SROCC scores estimated by SSRM using non-normalization and normalization (normalized to 256×256 and normalized to 128×128 respectively) on the CSIQ and TID2008 databases. It can be seen that empirical scale does affect the performance of the proposed approach. Similar conclusions could be drawn by using other databases.

There are three parameters to be determined in our method, *i.e.*, the scale of Log-Gabor wavelet decomposition, the orientations of Log-Gabor wavelet decomposition, and the size of sub-images. In the literature, two strategies are often used for
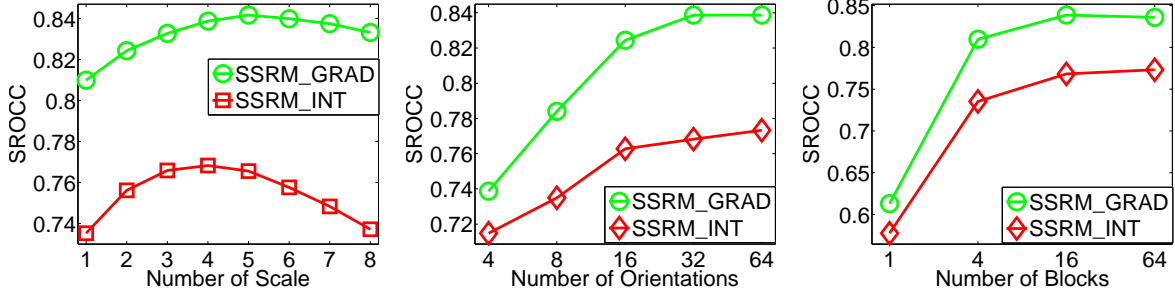
Fig. 8. The performance of SSRM in terms of SROCC vs. number of scales/orientations/blocks on the TID2008 database.

TABLE 2
The results of PLCC and SROCC by SSRM versus image size
on the CSIQ and TID2008 Databases

Model	Img.size	512×512	256×256	128×128
SRRM_GRAD	PLCC	0.8626	0.9442	0.8927
	SROCC	0.8556	0.9416	0.8751
SRRM_INT	PLCC	0.8507	0.8818	0.8515
	SROCC	0.7883	0.8376	0.8014
			CSIQ	

Model	Img.size	512×384	256×256	128×128
SRRM_GRAD	PLCC	0.7684	0.8371	0.7346
	SROCC	0.7527	0.8387	0.7186
SRRM_INT	PLCC	0.7686	0.7825	0.6711
	SROCC	0.7376	0.7683	0.6359
			TID2008	

determining parameters. One is to choose the parameters depending on how well the resulting model fits the physiological or psychophysical data, *e.g.* [28]. The other strategy is to select the parameters which optimize the performance in terms of predicting subjective ratings, *e.g.* [5, 9]. We adopted the second strategy. More precisely, we tuned up the parameters based on the TID2008 database. The performance of our method with respect to the aforementioned parameters is summarized in Fig. 8. According to Fig. 8, the parameters are set as follows. The Log-Gabor decomposition is done with 32 orientations at 4 scales. Thus there are a total of 128 subbands after the decomposition. The local SSR feature is computed on the blocks of size 64×64 in each Log-Gabor subband.

4.3 Performance Comparison

The experimental results of the proposed approach and the compared approaches on the test databases are listed in Table 3. The FR-IQA methods which perform the best on each database are marked in boldface, while the RR-IQA methods with best performances are underlined. In addition, in order to provide an overall performance evaluation, Table 3 gives the average PLCC, SROCC, and KROCC values over all the test databases. The average values are computed in two cases: direct average and weighted average. The direct average is about setting the weights of each database equal, while the

TABLE 4
The statistical significance tests on several IQA methods, which
is done on the (a) TID2008, (b) CSIQ and (c) LIVE databases.
The value '1' indicates that the method in the row is
significantly better than the method in the column.

TID2008	SSRM_GRAD	SSRM_INT	RRED [12]	FSIM[5]	IW-SSIM[41]	VIF[3]	SSIM[18]	PSNR
SSRM_GRAD	0	1	0	0	0	1	1	1
SSRM_INT	0	0	0	0	0	0	0	1
RRED [12]	0	1	0	0	0	1	1	1
FSIM[5]	1	1	1	0	1	1	1	1
IW-SSIM[41]	1	1	1	0	0	1	1	1
VIF[3]	0	1	0	0	0	0	1	1
SSIM[18]	0	0	0	0	0	0	0	1
PSNR	0	0	0	0	0	0	0	0

(a)

CSIQ	SSRM_GRAD	SSRM_INT	RRED [12]	FSIM[5]	IW-SSIM[41]	VIF[3]	SSIM[18]	PSNR
SSRM_GRAD	0	1	1	1	1	1	1	1
SSRM_INT	0	0	0	0	0	0	1	1
RRED [12]	0	1	0	0	0	1	1	1
FSIM[5]	0	1	0	0	0	1	1	1
IW-SSIM[41]	0	1	0	0	0	1	1	1
VIF[3]	0	1	1	1	1	0	1	1
SSIM[18]	0	0	0	0	0	0	0	1
PSNR	0	0	0	0	0	0	0	0

(b)

LIVE	SSRM_GRAD	SSRM_INT	RRED [12]	FSIM[5]	IW-SSIM[41]	VIF[3]	SSIM[18]	PSNR
SSRM_GRAD	0	1	0	0	0	0	0	1
SSRM_INT	0	0	0	0	0	0	0	1
RRED [12]	0	1	0	0	0	0	0	1
FSIM[5]	1	1	1	0	1	0	1	1
IW-SSIM[41]	1	1	1	0	0	0	1	1
VIF[3]	1	1	1	0	1	0	1	1
SSIM[18]	0	1	0	0	0	0	0	1
PSNR	0	0	0	0	0	0	0	0

(c)

TABLE 3

Performance comparison on seven benchmark databases. The 'img.size*' denotes the image size after normalization.

DataBase	Criteria	SSRM_GRAD	SSRM_INT	RRED [12]	FSIM[5]	IW-SSIM[41]	VIF[3]	SSIM[18]	PSNR
Type		**RR**	**RR**	**RR**	FR	FR	FR	FR	FR
Length		**img.size*/32**	**img.size*/32**	**img.size/36**	img.size	img.size	img.size	img.size	img.size
TID 2008	PLCC	**0.8371**	0.7825	0.8255	**0.8738**	0.8579	0.8084	0.7732	0.5734
	SROCC	**0.8387**	0.7683	0.8237	**0.8805**	0.8559	0.7491	0.7749	0.5531
	KROCC	**0.6480**	0.5719	0.6346	**0.6946**	0.6636	0.5860	0.5768	0.4027
	RMSE	**0.7342**	0.8356	0.7573	**0.6525**	0.6895	0.7899	0.8511	1.0994
	MAE	**0.5564**	0.6480	0.5641	**0.4926**	0.5276	0.6000	0.6547	0.8327
CSIQ	PLCC	**0.9442**	0.8818	0.9121	0.9120	0.9144	**0.9277**	0.8613	0.8000
	SROCC	**0.9416**	0.8376	0.9184	0.9242	0.9213	**0.9195**	0.8756	0.8058
	KROCC	**0.7820**	0.6559	0.7429	0.7567	0.7529	**0.7537**	0.6907	0.6084
	RMSE	**0.0865**	0.1238	0.1077	0.1077	0.1063	**0.0980**	0.1334	0.1575
	MAE	**0.0692**	0.0972	0.0820	0.0797	0.0801	**0.0743**	0.0991	0.1195
LIVE	PLCC	**0.9423**	0.9210	0.9385	0.9597	0.9522	**0.9604**	0.9449	0.8723
	SROCC	**0.9432**	0.9200	0.9429	0.9634	0.9567	**0.9636**	0.9479	0.8756
	KROCC	**0.7907**	0.7501	0.7888	**0.8337**	0.8175	0.8282	0.7963	0.6865
	RMSE	**9.1434**	10.6465	9.4317	7.6780	8.3473	**7.6137**	8.9455	13.3597
	MAE	**7.0767**	8.3466	7.2976	**5.9468**	6.4702	6.1070	6.9325	10.5093
IVC	PLCC	0.8688	0.8549	**0.9050**	**0.9376**	0.9231	0.9028	0.9119	0.7196
	SROCC	0.8622	0.8393	**0.8987**	**0.9262**	0.9125	0.8964	0.9018	0.6884
	KROCC	0.6780	0.6514	**0.7175**	**0.7564**	0.7339	0.7158	0.7223	0.5218
	RMSE	0.6033	0.6322	**0.5183**	**0.4236**	0.4686	0.5239	0.4999	0.8460
	MAE	0.4461	0.4734	**0.3971**	**0.3388**	0.3694	0.4104	0.3777	0.6677
MICT	PLCC	0.8059	0.7922	**0.8272**	0.9078	**0.9248**	0.9138	0.8887	0.6429
	SROCC	0.8048	0.7958	**0.8228**	0.9059	**0.9202**	0.9077	0.8794	0.6132
	KROCC	0.6102	0.5951	**0.6306**	0.7302	**0.7537**	0.7315	0.6939	0.4443
	RMSE	0.7409	0.7638	**0.7033**	0.5248	**0.4761**	0.5084	0.5738	0.9585
	MAE	0.5686	0.6107	**0.5465**	0.4021	**0.3677**	0.4038	0.4386	0.7761
A57	PLCC	**0.9458**	0.8876	0.8547	**0.9393**	0.9034	0.6915	0.8017	0.7073
	SROCC	**0.9253**	0.8794	0.8399	**0.9181**	0.8709	0.6223	0.8066	0.6189
	KROCC	**0.7769**	0.7042	0.6483	**0.7639**	0.6842	0.4589	0.6058	0.4309
	RMSE	**0.0798**	0.1132	0.1276	**0.0844**	0.1054	0.1784	0.1469	0.1737
	MAE	**0.0615**	0.0921	0.1051	**0.0721**	0.0892	0.1329	0.1209	0.1417
WIQ	PLCC	0.8483	**0.8944**	0.8367	**0.8546**	0.8329	0.7605	0.7980	0.7939
	SROCC	0.7871	**0.8638**	0.7626	**0.8006**	0.7865	0.6918	0.7261	0.6257
	KROCC	0.6196	**0.6867**	0.5864	**0.6215**	0.6038	0.5246	0.5569	0.4626
	RMSE	12.1285	**10.2441**	12.5448	**11.8949**	12.6765	14.8731	13.8046	14.1381
	MAE	9.4520	**8.1653**	9.8095	**9.0496**	9.9121	12.2465	10.9873	11.2027
Weighted Average	PLCC	**0.8846**	0.8408	0.8726	**0.9050**	0.8960	0.8728	0.8407	0.7020
	SROCC	**0.8830**	0.8230	0.8719	**0.9094**	0.8955	0.8423	0.8430	0.6874
	KROCC	**0.7083**	0.6362	0.6934	**0.7409**	0.7215	0.6827	0.6593	0.5161
Direct Average	PLCC	**0.8846**	0.8592	0.8714	**0.9121**	0.8542	0.8522	0.8542	0.7299
	SROCC	**0.8718**	0.8435	0.8584	**0.9027**	0.8446	0.8215	0.8446	0.6830
	KROCC	**0.7008**	0.6593	0.6784	**0.7367**	0.6632	0.6570	0.6632	0.5082

weights in the weighted average are determined by the number of distorted images in each database.

From Table 3, it can be seen that the proposed method performs consistently well across all the databases. In particular, SSRM_INT performs the best among the compared RR-IQA methods on the WIQ database, and SSRM_GRAD performs better than RRED on all the databases except MICT. Considering the scales of the databases (*i.e.* the number of images, the number of distortion types, and the number of observers), we believe that the results obtained on TID2008, CSIQ, and LIVE are much more convincing than those obtained on MICT. Moreover, SSRM_GRAD performs the best on average, no matter which kind of averaging and which performance measure are used. Compared with the FR-IQA methods, SRRM_GRAD performs the best on average among

all the compared methods except FSIM. Even compared with FSIM, SRRM_GRAD shows noticeable performance improvement on the CSIQ and A57 databases.

In order to make statistically meaningful conclusions on the performance of the compared methods, we further conducted a series of hypothesis tests based on the prediction residuals of each compared method after nonlinear regression. Smaller the residuals are, better the metric is. By assuming that the model residuals follow a Gaussian distribution, we apply the left-tailed F-test to the residuals of every pair of the compared methods. A value of 1 for the left-tailed F-test at a significance level of 0.05 means that the first model has better performance than the second model with a confidence greater than 95%. A value of 0 means that the first model is not significantly better than the second one. Table 4 shows the results of the

significance test on the TID2008, CSIQ and LIVE databases. It can be seen that on the CSIQ database, SSRM_GRAD is significantly better than all other compared methods. On the CSIQ and TID2008 databases, SSRM_GRAD is very competitive to RRED.

As discussed in Section 3.1, the data transmission rate is a key factor in some scenarios. In such cases, the feature length should be considered. Thus, we evaluated our SSR* feature using the so-called *minimum feature* protocol proposed in [12]. In this setting, we selected three low-data-rate RR-IQA metrics for comparison, including RR-SSIM, HWD2 and WNISM. The results are shown in Table 5. The SSR-M_GRAD* outperforms WNISM by a large margin across all the benchmark databases and on all the performance measures. Compared with RR-SSIM, when the feature length is set 16, the superiority of SSRM_GRAD* is obvious on the TID2008, CSIQ and A57 databases. On average, the SSRM_GRAD* method has advantages on data rate.

4.4 Evaluation on Individual Distortion Type

The satisfactory performance of an IQA metric evaluated on multiple distortion types does not necessarily guarantee good results by the metric for individual distortion type, and vice versa. To examine how the proposed SSRM method behaves on different distortion types, we show the performance of SSRM with respect to each type of distortion on the TID2008, CSIQ and LIVE databases in Table 6. For clarity, only the SROCC values are listed. Note that SROCC is chosen as it is suitable for measuring a small number of data points and its value will not be affected by an unsuccessful monotonic nonlinear mapping. There are a total of 28 groups of distorted images in the three databases. We use boldface font to highlight the best results in each group. It is seen that RRED is marked 19 times, followed by SSRM_GRAD which is only marked 7 times. However, SSRM_GRAD is better than RRED in terms of the average performance on the three databases. From Table 3 and Table 6, it can be seen that WNISM just performs well on some specific distortion types.

For visualization, we show the scatter plots of the predicted quality scores against subjective DMOS scores for several compared method on the CSIQ database with six types of distortions in Fig. 9, where each point represents one test image and the curves are obtained by the nonlinear fitting via (13). It can be observed that for SSRM_INT, the distribution of the predicted scores on the CTD distortion deviates much from the distributions on other types of distortions, which degrades its overall performance. It can also be seen that the objective scores predicted by SSRM_GRAD are more correlated with the subjective ratings than other compared methods.

Figure 9 also gives another interesting phenomena. The scatter plot of SSRM_GRAD is concentrated across different types of distortions. For example, its points corresponding to JPEG, APN and CTD distortions are very close to each other. However, the points of RRED corresponding to JPEG, APN and CTD are relatively far from each other, which explains why some RR-IQA methods perform well for many individual types of distortions but their performance decreases on the entire database.

4.5 Computational Cost

We evaluated the computational efficiency of our method by comparing its average running time with other methods. The test was performed on a Dell Inspiron 530s PC embedded with an Intel T5850 processor and 2GB RAM. The software platform is Matlab R2012b. The size of the test images is 768×512. All the Matlab source codes of the compared methods were obtained from the original authors. The results are shown in Table 7. It can be seen that our approach has moderate computational cost - the running time of our approach is about two times as the RRED method and 1.4 times as the WNISM method. Note that WNISM is about two times faster than RR-SSIM [1]. The computational complexities of WNISM, RRED and RR-SSIM mainly come from the highly overcomplete steerable pyramid decomposition, while ours comes from the multi-orientation and multi-scale wavelet decomposition in Log-Gabor representation.

5 CONCLUSION

Image quality assessment can facilitate many image-related tasks including image processing and image recognition. However, image quality assessment is a challenging task, especially when only partial information of the reference image is known. Moreover, the quality of an image depends on not only the content of the image but also the perception ability of humans. To simulate the humans visual perception, Log-Gabor is employed to preprocess and decompose the image, which extracts the spatial-frequency components of images in parallel to the V1 cortex of HVS.

For characterizing the distortion of image content, many traditional RR-IQA approaches have been proposed based on counting and comparing the numbers of local elements of the reference image and the distorted image. Such approaches may lose the details of the spatial distribution of image elements. To overcome this problem, the regularity of spatial arrangements of local image structure is accounted in this paper. Fractal analysis is employed to characterize the spatial distributions of image structure expressed in Log-Gabor subbands of both the reference image and the distorted image.

The proposed method has several advantages: competitive performance among state-of-the-arts, consistent performance across different types of distortions, high ratio of accuracy over data rate, moderate and acceptable computational cost. The proposed approach was evaluated on seven public benchmark datasets using five performance metrics. The results show that our method is competitive against the state-of-the-art approaches. In the future, we would like to study the application of our approach to video quality assessment.

REFERENCES

[1] A. Rehman and Z. Wang. "Reduced-reference image quality assessment by structural similarity estimation". *IEEE Trans. Image Process.*, 21(8):3378–3389, 2012. 1, 2, 6, 9

[2] Z. Wang, G. Wu, H. R. Sheikh, E. P. Simoncelli, E. H. Yang, and A. C. Bovik. "Quality-aware images". *IEEE Trans. Image Process.*, 15(6):1680–1689, 2006. 2, 6

[3] H. R. Sheikh and A. C. Bovik. "Image information and visual quality". *IEEE Trans. Image Process.*, 15(2):430–444, 2006. 6

TABLE 5
Results of the minimum feature test on seven benchmark databases

DataBase	Criteria	SSRM_GRAD*				SSRM_INT*				RR-SSIM [1]	HWD2 [9]	WNISM [2]
Length	$Scale \times block$	1×16	2×16	3×16	4×16	1×16	2×16	3×16	4×16	36	16	18
TID 2008	PLCC	0.7661	0.7831	0.7973	**0.8047**	0.6845	0.7105	0.7278	0.7374	0.7231	N/A	0.5891
	SROCC	0.7734	0.7936	0.8101	**0.8174**	0.6825	0.7092	0.7245	0.7326	0.7210	N/A	0.5119
	KROCC	0.5708	0.5926	0.6094	**0.6174**	0.4844	0.5072	0.5210	0.5285	0.5236	N/A	0.3589
	RMSE	0.8625	0.8345	0.8100	**0.7966**	0.9783	0.9444	0.9203	0.9064	0.9270	N/A	1.0843
	MAE	0.6655	0.6337	0.6137	**0.6031**	0.7770	0.7486	0.7297	0.7171	0.7190	N/A	0.8666
CSIQ	PLCC	0.8961	0.9045	0.9104	**0.9142**	0.8205	0.8344	0.8404	0.8423	0.8426	N/A	0.7124
	SROCC	0.8876	0.8964	0.9043	**0.9074**	0.7819	0.7990	0.8047	0.8015	0.8527	N/A	0.7431
	KROCC	0.6962	0.7098	0.7212	**0.7256**	0.5825	0.6043	0.6114	0.6084	0.6540	N/A	0.5457
	RMSE	0.1165	0.1119	0.1086	**0.1064**	0.1501	0.1447	0.1423	0.1415	0.1413	N/A	0.1842
	MAE	0.0935	0.0895	0.0865	**0.0852**	0.1205	0.1156	0.1138	0.1132	0.1092	N/A	0.1492
LIVE	PLCC	0.9067	0.9144	0.9183	0.9213	0.8702	0.8831	0.8912	0.8938	0.9194	**0.9624**	0.7365
	SROCC	0.9052	0.9147	0.9188	0.9217	0.8689	0.8817	0.8901	0.8924	0.9129	**0.9418**	0.7472
	KROCC	0.7247	0.7388	0.7459	0.7515	0.6699	0.6858	0.6983	0.7029	0.7349	N/A	0.5577
	RMSE	11.5242	11.0600	10.8181	10.6261	13.4605	12.8199	12.3960	12.2557	11.3026	**6.3657**	18.4814
	MAE	9.3972	8.9267	8.7825	8.5453	10.6940	10.3576	9.9702	9.8502	9.1889	**4.8445**	14.6352
IVC	PLCC	0.7502	0.8035	0.8385	**0.8390**	0.7461	0.7846	0.8098	0.8258	0.8177	N/A	0.5311
	SROCC	0.7341	0.7866	**0.8199**	0.8185	0.7289	0.7686	0.7957	0.8082	0.8154	N/A	0.4114
	KROCC	0.5411	0.5874	**0.6197**	0.6164	0.5372	0.5754	0.6018	0.6109	0.6164	N/A	0.2907
	RMSE	0.8056	0.7253	0.6638	**0.6630**	0.8113	0.7554	0.7148	0.6870	0.7014	N/A	1.0322
	MAE	0.6373	0.5611	0.5168	**0.5120**	0.6520	0.5904	0.5526	0.5284	0.5619	N/A	0.8550
MICT	PLCC	0.6362	0.7342	0.7266	0.7272	0.6593	0.6697	0.6883	0.6927	**0.8051**	N/A	0.6542
	SROCC	0.6416	0.7389	0.7254	0.7269	0.6413	0.6573	0.6782	0.6878	**0.8003**	N/A	0.6322
	KROCC	0.4575	0.5470	0.5341	0.5368	0.4633	0.4850	0.5014	0.5099	**0.6090**	N/A	0.4570
	RMSE	0.9656	0.8497	0.8599	0.8591	0.9410	0.9294	0.9079	0.9026	**0.7423**	N/A	0.9464
	MAE	0.7965	0.6775	0.6778	0.6806	0.7517	0.7256	0.7032	0.6967	**0.5648**	N/A	0.7742
A57	PLCC	0.7498	0.7919	0.8134	**0.8396**	0.5440	0.6837	0.7565	0.7791	0.7044	N/A	0.5125
	SROCC	0.7397	0.7953	0.8105	**0.8276**	0.5252	0.6818	0.7290	0.7711	0.7301	N/A	0.3140
	KROCC	0.5531	0.6007	0.6189	**0.6399**	0.3517	0.4692	0.5168	0.5559	0.5345	N/A	0.2210
	RMSE	0.1626	0.1501	0.1430	**0.1335**	0.2063	0.1794	0.1607	0.1541	0.1744	N/A	0.2317
	MAE	0.1359	0.1257	0.1176	**0.1077**	0.1713	0.1481	0.1289	0.1275	0.1433	N/A	0.1971
WIQ	PLCC	0.8522	0.8415	0.8427	0.8383	0.8322	0.8494	0.8689	**0.8742**	N/A	N/A	0.3401
	SROCC	0.7768	0.8008	0.8061	0.8016	0.7446	0.8091	0.8290	**0.8368**	N/A	N/A	0.2156
	KROCC	0.5892	0.6126	0.6177	0.6107	0.5651	0.6164	0.6380	**0.6500**	N/A	N/A	0.1561
	RMSE	11.9849	12.3734	12.3307	12.4897	12.7025	12.0881	11.3383	**11.1206**	N/A	N/A	21.5404
	MAE	9.2227	9.6357	9.7240	9.8264	9.5498	9.7289	**9.1873**	9.2003	N/A	N/A	16.9682

TABLE 6
SROCC values of several RR-IQA methods with respect to individual distortion type on the TID2008, CSIQ, and LIVE databases. The 'img.size*' denotes the image size after normalization.

| Database	Distortion Type	SSRM_GRAD	SSRM_INT	RRED [12]	SSRM_GRAD*	SSRM_INT*	RR-SSIM [1]	HWD2 [9]	WNISM [2]
Length		img.size*/32	img.size*/32	img.size/36	16	16	36	16	18
TID2008	Additive Gaussian noise	0.7921	0.7725	**0.8203**	0.7372	0.7065	N/A	N/A	0.6037
	Noise in color comp.	0.7935	0.8063	**0.8502**	0.6865	0.7209	N/A	N/A	0.6076
	Spatially corr. noise	0.8299	0.7407	**0.8417**	0.7525	0.6655	N/A	N/A	0.6008
	Masked noise	0.7548	0.7649	**0.8325**	0.6266	0.6317	N/A	N/A	0.6311
	High frequency noise	**0.9139**	0.8978	0.9088	0.7826	0.8313	N/A	N/A	0.7064
	Impulse noise	0.5949	0.6428	**0.7407**	0.4516	0.4864	N/A	N/A	0.5922
	Quantization noise	**0.8499**	0.8293	0.8308	0.7182	0.6659	N/A	N/A	0.6096
	Gaussian blur	0.9481	0.9262	**0.9573**	0.8915	0.8903	N/A	N/A	0.8723
	Image denoising	0.9487	0.9182	**0.9493**	0.8966	0.8771	N/A	N/A	0.8582
	JPEG compression	0.9271	0.9224	**0.9333**	0.8640	0.8552	N/A	N/A	0.8246
	JPEG2000 compression	0.9618	0.9587	**0.9681**	0.9245	0.9143	N/A	N/A	0.9344
	JPEG trans. error	0.8586	0.8675	**0.8704**	0.7998	0.8356	N/A	N/A	0.8774
	JPEG2000 trans. error	**0.8422**	0.7683	0.7421	0.7894	0.7270	N/A	N/A	0.6889
	Non ecc. patt. noise	0.7001	0.6982	**0.7127**	0.6166	0.6694	N/A	N/A	0.4293
	Local block-wise dist.	0.7538	0.7009	**0.8243**	0.4931	0.4326	N/A	N/A	0.6071
	Mean shift	**0.6506**	0.5326	0.5378	0.4898	0.3911	N/A	N/A	0.3204
	Contrast change	0.1816	**0.6589**	0.5424	0.0548	0.6377	N/A	N/A	0.7042
CSIQ	Additive Gaussian noise	0.9111	0.9209	**0.9353**	0.8293	0.8393	N/A	N/A	0.8188
	JPEG compression	0.9521	**0.9556**	0.9521	0.8864	0.8528	N/A	N/A	0.8955
	JPEG2000 compression	**0.9663**	0.9627	0.9628	0.9170	0.9042	N/A	N/A	0.9405
	Additive pink noise	0.9210	0.9123	**0.9362**	0.8480	0.8282	N/A	N/A	0.8002
	Gaussian blur	**0.9686**	0.9529	0.9634	0.9202	0.9176	N/A	N/A	0.9144
	Contrast change	0.9305	0.5348	**0.9383**	0.8660	0.1999	N/A	N/A	0.9122
LIVE	JPEG2000 compression	**0.9618**	0.9530	0.9580	0.9325	0.8869	N/A	0.9362	0.9330
	JPEG compression	0.9734	0.9683	**0.9759**	0.9170	0.9029	N/A	0.9543	0.9204
	Additive Gaussian noise	0.9711	0.9569	**0.9780**	0.9389	0.9320	0.9642	0.9321	0.8701
	Gaussian blur	0.9662	0.9383	**0.9675**	0.8557	0.8765	0.8692	0.8282	0.9145
	JPEG2000 trans. error	0.9214	0.8681	**0.9427**	0.8671	0.8389	0.9137	0.9386	0.9227
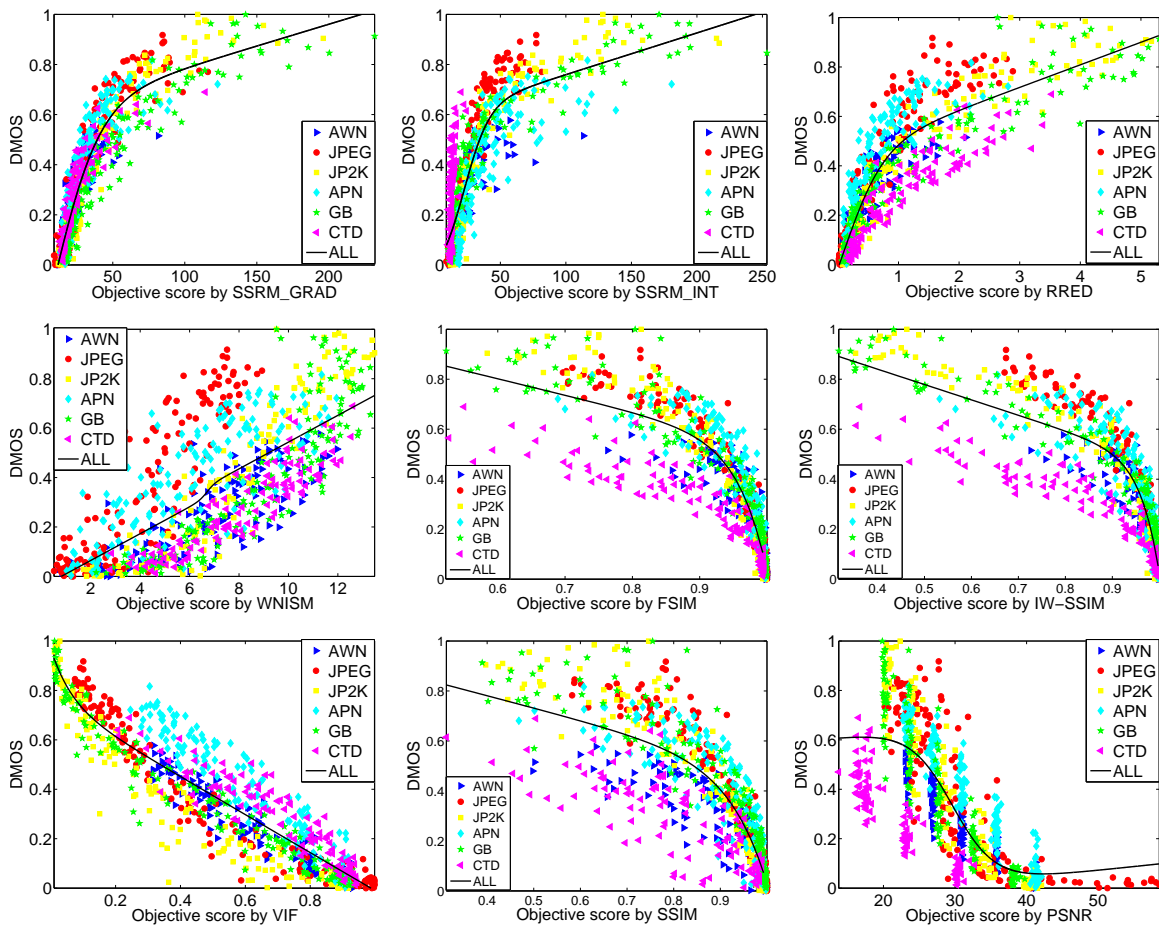
Fig. 9. Scatter plots of the predicted quality scores against the subjective quality scores (DMOS) by several IQA methods on the CSIQ database. The six types of distortions are represented by different shapes and colors.

TABLE 7
Time cost of the compared methods

Model	SSRM_GRAD	SSRM_INT	RRED [12]	RR-SSIM [1]	HWD2 [9]	WNISM [2]	FSIM[5]	IW-SSIM[41]	VIF[3]	SSIM[18]	PSNR
Type	**RR**	**RR**	**RR**	**RR**	**RR**	**RR**	FR	FR	FR	FR	FR
Time (seconds)	5.63	5.47	2.74	N/A	N/A	4.01	1.74	2.04	4.63	0.20	0.01

[4] N. Damera-Venkata, T. D. Kite, W. S. Geisler, B. L. Evans, and A. C. Bovik. "Image quality assessment based on a degradation model". *IEEE Trans. Image Process.*, 9(4):636–650, 2000.

[5] L. Zhang, L. Zhang, X. Q. Mou, and D. Zhang. "FSIM: a feature similarity index for image quality assessment". *IEEE Trans. Image Process.*, 20(8):2378–2386, 2011. 1, 4, 6, 7

[6] H. R. Sheikh, A. C. Bovik, and L. K. Cormack. "No-reference quality assessment using natural scene statistics: JPEG2000". *IEEE Trans. Image Process.*, 14(11):1918–1927, 2005. 1

[7] Z. Wang, A. C. Bovik, and B. L. Evans. "Blind measurement of blocking artifacts in images". In *Proc. IEEE Int. Conf. Image Process.*, volume 3, pages 981–984, 2000.

[8] Z. Wang, H. R. Sheikh, and A. C. Bovik. "No-reference perceptual quality assessment of JPEG compressed images". In *Proc. of IEEE Int. Conf. Image Process.*, volume 1, pages 477–480, 2002. 1

[9] X. Gao, W. Lu, D. Tao, and X. Li. "Image quality assessment based on multiscale geometric analysis". *IEEE Trans. Image Process.*, 18(7):1409–1423, 2009. 1, 2, 3, 6, 7

[10] U. Engelke, M. Kusuma, H. J. Zepernick, and M. Calder-a. "Reduced-reference metric design for objective perceptual quality assessment in wireless imaging". *SIGNAL PROCESS-IMAGE*, 24(7):525–547, 2009.

[11] Q. Li and Z. Wang. "Reduced-reference image quality assessment using divisive normalization-based image representation". *IEEE J. Sel. Topics Signal Process.*, 3(2):202–211, 2009. 2

[12] R. Soundararajan and A. C. Bovik. "RRED indices: reduced reference entropic differencing for image quality assessment". *IEEE Trans. Image Process.*, 21(2):517–526, 2012. 1, 2, 6, 9

[13] Z. Wang and E. P. Simoncelli. "Reduced-reference image quality assessment using a wavelet-domain natural image statistic model". In *Proc. SPIE Human Vision and Electronic Imaging X*, volume 5666, pages 149–159, 2005. San Jose, CA. 1

[14] W. F. Xue and X. Q. Mou. "Reduced reference image quality assessment based on Weibull statistics". In *Proc. of the Int. Workshop on Quality of Multimedia Experience*, pages 1–6, 2010. 1

[15] L. Ma, S. N. Li, F. Zhang, and K. N. Ngan. "Reduced-reference

image quality assessment using reorganized DCT-based image representation". *IEEE Trans. Multimedia*, 13(4):824–829, 2011. 1

[16] A. Ait Abdelouahad, M. El Hassouni, H. Cherifi, and D. Aboutajdine. "Image quality assessment based on IMF coefficients modeling". In *Proc. IInt. Conf. on Digital Information and Communication Technology and its Applications*, pages 131–145, 2011.

[17] I. P. Gunawanand and M. Ghanbari. "Reduced reference picture quality estimation by using local Harmonic amplitude information". In *Proc. of London Commun. Symp*, pages 137–140, 2003. 1

[18] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli. "Image quality assessment: from error visibility to structural similarity". *IEEE Trans. Image Process.*, 13(4):600–612, 2004. 2, 6

[19] H. R. Sheikh, K. Seshadrinathan, A. K. Moorthy, Z. Wang, A. C. Bovik, and L. K. Cormack. Image and video quality assessment research at LIVE. http://live.ece.utexas.edu/research/quality. 2, 5

[20] Y. Xu, H. Ji, and C. Fermüller. "Viewpoint invariant texture description using fractal analysis". *INT J COMPUT VISION*, 83(1):85–100, 2009. 2, 3

[21] Y. Xu, Y.H. Quan, H.B. Ling, and Hui. Ji. "Dynamic texture classification using dynamic fractal analysis". In *Proc. IEEE Int. Conf. Computer Vision*, pages 1219–1226. IEEE, 2011.

[22] Y.H. Quan, Y. Xu, Y.P. Sun, and Y. Luo. "Lacunarity Analysis on Image Patterns for Texture Classification". In *Proc. of IEEE Conf. Computer Vision and Pattern Recognition*, pages 160–167, 2014. 2, 3

[23] A. P. Pentland. "Fractal-Based description of natural scenes". *IEEE Trans. Pattern Anal. Mach. Intell*, 6(6):661–674, 1984. 2, 3, 5

[24] Mihai Ivanovici, Noël Richard, and Christine Fernandez-Maloigne. "Towards video quality metrics based on colour fractal geometry". *EURASIP J. Image Video Proc.*, 2010:4, 2010. 2

[25] L.H. He, D.X. Wang, X.L. Li, D.C. Tao, X.B. Gao, and F. Gao. "Color fractal structure model for reduced-reference colorful image quality assessment". In *Proc. Neural Information Processing Systems*, pages 401–408, 2012. 2

[26] F.W. Campbell and JG Robson. "Application of Fourier analysis to the visibility of gratings". *The Journal of Physiology*, 197(3):551, 1968. 2, 3

[27] E. Peli, L.E. Arend, G.M. Young, and R.B. Goldstein. "Contrast sensitivity to patch stimuli: effects of spatial bandwidth and temporal presentation". *Spatial Vision*, 7(1):1–14, 1993. 2, 3

[28] D. J. Field. "Relations between the statistics of natural images and the response properties of cortical cells". *J. Opt. Soc. Amer. A*, 4(12):2379–2394, 1987. 2, 3, 7

[29] J. G. Daugman. "Uncertainty relation for resolution in space, spatial frequency, and orientation optimized by two-dimensional visual cortical filters". *J. Opt. Soc. Amer. A*, 2(7):1160–1169, 1985. 3

[30] B. B. Mandelbrot. *The Fractal Geometry of Nature*. San Francisco, CA: Freeman, 1982. 3

[31] N. Sarkar and B. B. Chaudhuri. "An efficient differential box-counting approach to compute fractal dimension of image". *IEEE Trans. Systems, Man, Cybern.,Syst.*, 24(1):115–120, 1994. 3

[32] A. Liu, W. Lin, and M. Narwaria. "Image Quality Assessment Based on Gradient Similarity". *IEEE Trans. Image Process.*, 21(4):1500–1512, 2012. 4

[33] G.H. Chen, C.L. Yang, and S.L. Xie. "Gradient-based structural similarity for image quality assessment". In *Proc. of IEEE Int. Conf. Image Process.*, pages 2929–2932, 2006.

[34] D.O. Kim, H.S. Han, and R.H. Park. "Gradient information-based image quality metric". *IEEE Trans. Power Electron.*, 56(2):930–936, 2010. 4

[35] 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*, 10(10):30–45, 2009. 5

[36] P. Ye, Y. Li, and D. Doermann. "Convolutional Neural Networks for No-Reference Image Quality Assessment". In *Proc. of IEEE Conf. Computer Vision and Pattern Recognition*, pages 1733–1740, 2014. 4

[37] E. C. Larson and D. M. Chandler. Categorical Image Quality (CSIQ) Database. http://vision.okstate.edu/csiq. 5

[38] A. Ninassi, P. Le Callet, and F. Autrusseau. Subjective quality assessment-IVC database. http://www2.irccyn.ec-nantes.fr/ivcdb. 5

[39] Y. Horita, K. Shibata, Y. Kawayoke, and Z. M. Parves Sazzad. MICT image quality evaluation database. http://mict.eng.u-toyama.ac.jp/mict/index2.html. 5

[40] U. Engelke, H. J. Zepernick, and M. Kusuma. Wireless imaging quality database, 2010. http://www.bth.se/tek/rcg.nsf/pages/wiq-db. 5

[41] D. M. Chandler and S. S. Hemami. A57 database. http://foulard.ece.cornell.edu/dmc27/vsnr/vsnr.html. 5

[42] Z. Wang and Q. Li. "Information content weighting for perceptual image quality assessment". *IEEE Trans. Image Process.*, 20(5):1185–1198, 2011. 6

[43] VQEG. Final report from the video quality experts group on the validation of objective models of video quality assessment, 2000. http://www.vqeg.org. 6

[44] H. R. Sheikh, M. F. Sabir, and A. C. Bovik. "A statistical evaluation of recent full reference image quality assessment algorithms". *IEEE Trans. Image Process.*, 15(11):3440–3451, 2006. 6

[45] Z. Wang. "SSIM Index for Image Quality Assessment". https://ece.uwaterloo.ca/%7Ez70wang/research/ssim/. 6

Yong Xu received the B.S., M.S. and Ph.D. degrees in Mathematics from Nanjing University, Nanjing, China, in 1993, 1996 and 1999, respectively. He was a Post-Doctoral Research Fellow of computer science with South China University of Technology, Guangzhou, China, from 1999 to 2001, where he became a Faculty Member and where he is currently a Professor with the School of Computer Science & Engineering. His current research interests include image analysis, image and video recognition, image quality assessment, object recognition. Dr. Xu is a member of the IEEE Computer Society and the ACM.

Delei Liu received the Ph.D. degree in Computer Science from South China University of Technology, Guangzhou, China, in 2014. He is currently an assistant professor with the Software School, Shanxi Agricultural University, Taigu, China. His current research interests include computer vision, and perceptual image/video quality assessment.

Yuhui Quan received the Ph.D. degree in Computer Science from South China University of Technology in 2013. He is currently the postdoc research fellow in Mathematics at National University of Singapore. His research interests include computational vision, sparse representation, and machine learning.

Patrick Le Callet received both an M.Sc. and a PhD degree in image processing from Ecole polytechnique de l'Université de Nantes. He was also a student at the Ecole Normale Superieure de Cachan where he sat the "Aggrégatio "(credentialing exam) in electronics of the French National Education. He worked as an Assistant Professor from 1997 to 1999 and as a full time lecturer from 1999 to 2003 at the Department of Electrical Engineering of Technical Institute of the University of Nantes (IUT). Since 2003 he teaches at Ecole polytechnique de l'Université de Nantes (Engineering School) in the Electrical Engineering and the Computer Science departments where is now a Full Professor. Since 2006, he is the head of the Image and Video Communication lab at CNRS IRCCyN, a group of more than 35 researchers. He is mostly engaged in research dealing with the application of human vision modeling in image and video processing. His current centers of interest are 3D image and video quality assessment, watermarking techniques and visual attention modeling and applications. He is co-author of more than 200 publications and communications and co-inventor of 13 international patents on these topics. He served or is currently serving as associate editor for several journal such as IEEE transactions on Image Processing, IEEE transactions on Circuit System and Video Technology, SPRINGER EURASIP Journal on Image and Video Processing, or SPIE Electronic Imaging.