



# Saliency-induced reduced-reference quality index for natural scene and screen content images



Xionghuo Min<sup>a</sup>, Ke Gu<sup>b</sup>, Guangtao Zhai<sup>a,\*</sup>, Menghan Hu<sup>a,\*</sup>, Xiaokang Yang<sup>a</sup>

<sup>a</sup>Institute of Image Communication and Network Engineering, Shanghai Jiao Tong University, Shanghai 200240, China

<sup>b</sup>Beijing Key Laboratory of Computational Intelligence and Intelligent System, Faculty of Information Technology, Beijing University of Technology, Beijing 100124, China

## ARTICLE INFO

### Article history:

Received 25 April 2017

Revised 9 October 2017

Accepted 26 October 2017

Available online 31 October 2017

### Keywords:

Image quality assessment

Visual saliency

Natural scene image

Screen content image

Image signature

Reduced-reference

## ABSTRACT

Massive content composed of both natural scene and screen content has been generated with the increasing use of wireless computing and cloud computing, which call for general image quality assessment (IQA) measures working for both natural scene images (NSIs) and screen content images (SCIs). In this paper, we develop a saliency-induced reduced-reference (SIRR) IQA measure for both NSIs and SCIs. Image quality and visual saliency are two widely studied and closely related research topics. Traditionally, visual saliency is often used as a weighting map in the final pooling stage of IQA. Instead, we detect visual saliency as a quality feature since different types and levels of degradation can strongly influence saliency detection. Image quality is described by the similarity between two images' saliency maps. In SIRR, saliency is detected through a binary image descriptor called "image signature", which significantly reduces the reference data. We perform extensive experiments on five large-scale NSI quality assessment databases including LIVE, TID2008, CSIQ, LIVEMD, CID2013, as well as two recently constructed SCI QA databases, i.e., SIQAD and QACS. Experimental results show that SIRR is comparable to state-of-the-art full-reference and reduced-reference IQA measures in NSIs, and it can outperform most competitors in SCIs. The most important is that SIRR is a cross-content-type measure, which works efficiently for both NSIs and SCIs. The MATLAB source code of SIRR will be publicly available with this paper.

© 2017 Elsevier B.V. All rights reserved.

## 1. Introduction

The quick advancements of transmission technologies have boosted various remote applications such as telecommuting and cloud computing, which bring massive computer-generated content called "screen content". The so-called screen content has some distinctive characteristics different from natural scene because of the contained computer generated content, e.g., texts, icons, tables, graphics, etc. Those distinctive characteristics which sometimes violate natural scene statistics (NSS) cause some failures in traditional natural scene image (NSI) based applications. Hence some specialized technologies for screen content image (SCI) have been proposed, such as screen content video compression [1].

The booming of screen content also calls for SCI-specific image quality measures. Limited work has been done concerning SCI quality assessment (QA). In [2], the authors constructed a screen image quality assessment database (SIQAD), which shows that

state-of-the-art image quality assessment (IQA) measures do not work efficiently for SCIs. It is reasonable since current IQA measures are implicitly designed for NSIs and somehow rely on NSS. Wang et al. [3] also constructed a database called quality assessment of compressed SCI (QACS). In [4], the authors proposed a full-reference (FR) saliency-guided quality measure named SQMS for SCI. SQMS exploits gradient magnitude similarity as the quality map, which is then weighted by a specific saliency map. Gu et al. [5,6] learned blind quality evaluation engines for SCI from a huge group of SCIs and corresponding objective quality scores calculated by FR measures.

Although dozens of NSI quality estimators [7–16] and several limited SCI quality measures [2–6] are proposed, they are either implicitly designed for NSIs or specifically developed for SCIs. Only very few quality measures can work for NSIs and SCIs simultaneously. Min et al. [17] proposed a blind blockiness measure which works for JPEG compressed NSIs and SCIs. Xu et al. [18] developed a measure for NSIs and SCIs. In [19], Min et al. constructed a cross-content-type database, and proposed a unified content-type adaptive blind IQA measure for compressed natural, graphic and screen content images. In practical multimedia communication systems, we may encounter both types of images, and sometimes we do not

\* Corresponding authors.

E-mail addresses: [minxionghuo@sjtu.edu.cn](mailto:minxionghuo@sjtu.edu.cn) (X. Min), [guke@bjut.edu.cn](mailto:guke@bjut.edu.cn) (K. Gu), [zhaiguangtao@sjtu.edu.cn](mailto:zhaiguangtao@sjtu.edu.cn) (G. Zhai), [humenghan@sjtu.edu.cn](mailto:humenghan@sjtu.edu.cn) (M. Hu), [xkyang@sjtu.edu.cn](mailto:xkyang@sjtu.edu.cn) (X. Yang).



**Fig. 1.** Influence of quality degradation on image saliency. First row: the reference and degraded images. Second row: corresponding saliency detected by the image signature model.

have any prior knowledge about the image types. Efficient general quality measures ignoring image types are highly needed in such circumstances. In this paper, we extract quality features efficient for both types of images and develop a general reduced-reference (RR) quality measure without any explicit image type classification.

The proposed method is based on visual saliency detection. Visual saliency detection is an important research topic in areas of psychology, image processing and computer vision [20]. Visual attention and quality assessment are two closely related research topics [7,9,10,12,13,21–23]. Quality degradation can influence visual attention [21]. Contrarily, visually salient positions should be more carefully processed since subjects judge the image quality according to the observations of some limited positions, and a typical use of visual attention model is to optimize resource allocation and improve the perceptual quality under the constraints of bandwidth [24–28].

Motivated by the interaction between visual attention and quality assessment, some researchers used visual attention map as a weighting map during the quality pooling stage of IQA [9,10,13,22,23]. Min et al. [13] collected some visual attention data for main-stream IQA databases. Zhang et al. [22,23] studied the use of saliency model in objective quality assessment models. Liu et al. [9] used the saliency map to highlight the visually salient areas. Besides highlighting the salient regions, Saha and Wu [10] used the dissimilarity between saliency maps of the reference and distorted images to highlight the more distorted image content. Besides visual attention maps, some measures utilize other kinds of weighting maps such as phase congruency map [7] and gradient magnitude map [12]. Although without explicit visual attention prediction or visual saliency detection processes, such kinds of weighting maps have also highlighted the visually salient positions, which can be also deemed as one kind of visual saliency.

Instead of as a weighting map, visual saliency can be also used as a quality feature since quality degradation can strongly affect saliency detection. Zhang et al. [8] proposed a FR IQA method named VSI by measuring the similarity between the reference image's and the distorted image's visual saliency. VSI is a FR measure since it utilizes not only saliency, but also the gradient magnitude and chrominance. All extracted feature maps have the same resolution as the reference image. Actually, deriving a gray scale saliency map from a color image is an operation of dimension reduction, which motivates us to develop a saliency-induced reduced-reference (SIRR) IQA measure.

SIRR detects saliency map of the reference image as the reference data, and then measure the similarity between the reference and distorted images' saliency maps. We try to reduce the reference data from two aspects. First, we down-sample the reference image to a coarser scale to detect saliency, whose resolution is only one over sixty-four of the original resolution. We take full advantage of such down-sampling operation to reduce the refer-

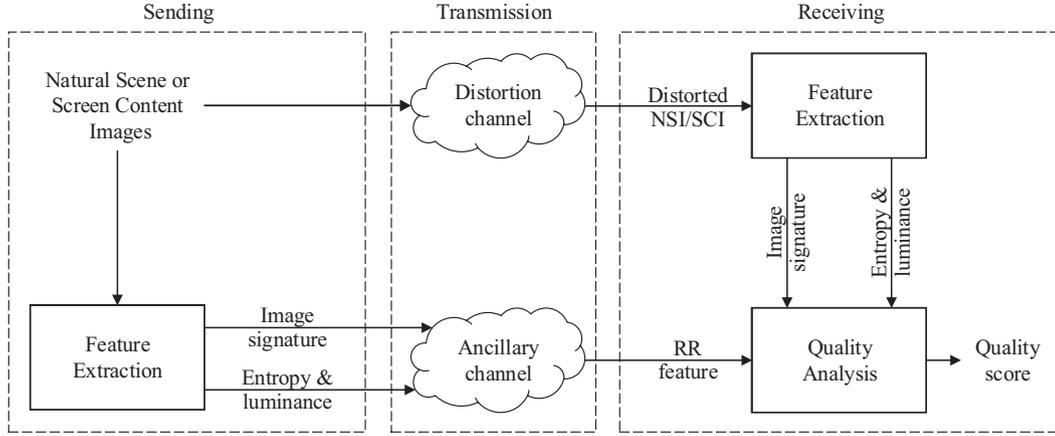
ence data. Second, we exploit a binary image descriptor called “image signature” [29] to detect image saliency. The image saliency is represented by the binary image signature, which also significantly reduces the reference data. The final quality is described by the similarity between two images' saliency maps. In this work, the similarity is evaluated by the classical image fidelity measure SSIM [30].

We perform extensive experiments to test the proposed SIRR in both NSIs and SCIs. Five large-scale NSI QA databases and two recent SCI QA databases are used. Among the NSI databases, LIVE [31], TID2008 [32] and CSIQ [33] are general-purpose IQA databases, whereas LIVEMD [34] focuses on multiply distorted images and CID2013 [35] consists of contrast changed images. As it to the SCI databases, SIQAD [2] is a general-purpose one and QACS [3] concentrates on compressed SCIs. The all seven databases can give an overall description of both NSIs and SCIs. As will be presented in the experiments part, the proposed SIRR is efficient for both types of images. SIRR can be comparable to or outperform state-of-the-art FR and RR IQA measures on all seven IQA databases.

The remainder of this paper is organized as follows. Section 2 describes the proposed saliency-induced reduced-reference quality measure. Experimental results are given in Section 3. We compare the proposed method with state-of-the-art FR and RR quality measures in this section. Section 4 concludes this paper.

## 2. Saliency-induced reduced-reference quality measure

As described in Section 1, visual saliency has been widely used in IQA, but it is generally used as a weighting map during the final pooling. Few work has considered saliency as a quality feature. Most bottom-up saliency models highly rely on the low-level features, which are sensitive to quality degradation. Fig. 1 illustrates the influence of quality degradation on image saliency. From this figure, we can observe that perceptible quality degradation can cause perceptible change of image low-level features, thus lead to the change of detected image saliency. Moreover, heavier quality degradation causes larger saliency change. So we can quantify such saliency change to predict the quality. Zhang et al. [8] proposed VSI based on such phenomena. But VSI is a FR measure since it utilizes not only saliency, but also the gradient magnitude and chrominance information. Motivated by the fact that saliency detection operation is a process of dimension or data reduction, we detect saliency as the reduced-reference information, and propose a saliency-induced reduced-reference quality measure. Fig. 2 illustrates the framework of SIRR. The same feature extraction process is executed for both the reference and the distorted images. Quality of the target image is then estimated by the similarity between two images' saliency maps, which are described through a binary image descriptor named image signature. The entropy and mean



**Fig. 2.** Framework of the proposed SIRR measure. Image signature are detected to represent saliency. Quality is described by the similarity between two images' saliency maps. Entropy and mean luminance are extracted for quality refining.

luminance of both images are also extracted for quality refining. The details are as follows.

### 2.1. Saliency detection: image signature

SIRR evaluates image quality by measuring the saliency change. In this work, we exploit image signature [29] to detect saliency. Given an image  $\mathbf{x}$ , we first down-sample it to a coarser scale

$$\mathbf{x}'(i, j) = \mathbf{a} * \mathbf{x}(ri, rj), \quad (1)$$

where  $\mathbf{x}'$  is the down-sampled image,  $r$  is the down-sampling rate which is set as 8 in this work,  $\mathbf{a}$  is a low-passing filter and  $*$  is the convolution operator,  $i = 1, \dots, \lfloor \frac{l}{r} \rfloor$  and  $j = 1, \dots, \lfloor \frac{l}{r} \rfloor$  are row and column indexes,  $l, j$  represent the total number of rows and columns of image  $\mathbf{x}$ ,  $\lfloor \cdot \rfloor$  is the floor operation.

Then we compute the two-dimensional discrete cosine transform (DCT) of the down-sampled image  $\mathbf{x}'$ , and calculate the sign of all DCT components and discard the amplitude information across the entire frequency spectrum

$$\hat{\mathbf{x}} = \text{sgn}(\text{DCT}(\mathbf{x}')), \quad (2)$$

where  $\hat{\mathbf{x}}$  the image signature of image  $\mathbf{x}'$ , and  $\text{sgn}$  is the sign function described by

$$\text{sgn}(x) = \begin{cases} -1 & \text{if } x < 0 \\ 1 & \text{if } x \geq 0 \end{cases} \quad (3)$$

Note that the image signature  $\hat{\mathbf{x}}$  is a binary map whose resolution is approximate one over sixty-four of the resolution of the original image  $\mathbf{x}$ . As demonstrated in [29], image signature contains some important information related to the foreground of an image. The image signature of the reference image acts as the reference data, which is transmitted to the receiver side through the ancillary channel.

In the receiver side, we first transform image signature  $\hat{\mathbf{x}}$  inversely back into the spatial domain

$$\bar{\mathbf{x}} = \text{IDCT}(\hat{\mathbf{x}}), \quad (4)$$

where  $\text{IDCT}(\cdot)$  computes the inverse discrete cosine transform (IDCT). Then a saliency map  $\mathbf{m}$  can be derived as squared inverse image

$$\mathbf{m} = \bar{\mathbf{x}} \circ \bar{\mathbf{x}}, \quad (5)$$

where  $\circ$  denotes the Hadamard product. If  $\mathbf{m}$  is further smoothed by a gaussian kernel and interpolated to the original resolution, it can act as a saliency map used to predict human fixations [29]. Here we use  $\mathbf{m}$  directly to avoid loss of information during smoothing and interpolation.

### 2.2. Similarity between two saliency maps

In the receiver side, we perform the same image signature-based saliency detection for the distorted image. The saliency maps of the reference and distorted images are denoted as  $\mathbf{m}_r$  and  $\mathbf{m}_d$ , respectively. Fig. 3 illustrates some examples of  $\mathbf{m}_r$  and  $\mathbf{m}_d$ . Note that the saliency maps here are slightly different from the traditional saliency maps used to predict human fixations.  $\mathbf{m}_r$  and  $\mathbf{m}_d$  need to be further smoothed and interpolated to be used to predict fixations. The following work is then to measure the similarity between  $\mathbf{m}_r$  and  $\mathbf{m}_d$ .

During recent years, dozens of FR IQA measures have been proposed [36]. Most of those FR IQA methods are image fidelity measures, which evaluate the fidelity or similarity between two images. In this work, we choose the classical image fidelity measure SSIM [30], and we find that SSIM is also efficient for measuring the similarity between saliency maps. Thus we evaluate the quality by measuring the SSIM between  $\mathbf{m}_r$  and  $\mathbf{m}_d$

$$\mathbf{q} = \text{SSIM}(\mathbf{m}_r, \mathbf{m}_d) = \frac{(2\boldsymbol{\mu}_r\boldsymbol{\mu}_d + c_1)(2\boldsymbol{\sigma}_{rd} + c_2)}{(\boldsymbol{\mu}_r^2 + \boldsymbol{\mu}_d^2 + c_1)(\boldsymbol{\sigma}_r^2 + \boldsymbol{\sigma}_d^2 + c_2)}, \quad (6)$$

where  $c_1, c_2$  are two stabilizing constants,  $\boldsymbol{\mu}_r, \boldsymbol{\mu}_d, \boldsymbol{\sigma}_r, \boldsymbol{\sigma}_d$  and  $\boldsymbol{\sigma}_{rd}$  represent the corresponding saliency maps' local mean, variance and covariance, which can be described as

$$\boldsymbol{\mu}_r(i, j) = \sum_{k,l} \mathbf{a}(k, l) \mathbf{m}_r(i+k, j+l), \quad (7)$$

$$\boldsymbol{\sigma}_r(i, j) = \sqrt{\sum_{k,l} \mathbf{a}(k, l) [\mathbf{m}_r(i+k, j+l) - \boldsymbol{\mu}_r(i, j)]^2}, \quad (8)$$

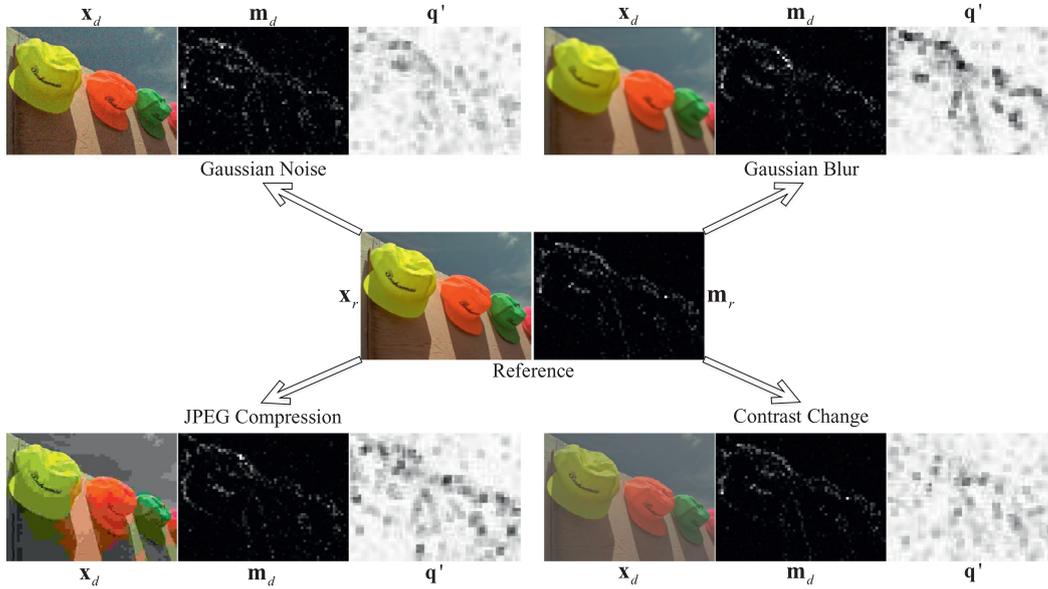
$$\boldsymbol{\sigma}_{rd}(i, j) = \sum_{k,l} \mathbf{a}(k, l) [\mathbf{m}_r(i+k, j+l) - \boldsymbol{\mu}_r(i, j)] \cdot [\mathbf{m}_d(i+k, j+l) - \boldsymbol{\mu}_d(i, j)], \quad (9)$$

where  $\mathbf{a}(k, l)$  indicates a local Gaussian window. Although SIRR adopts SSIM as the similarity function, it can be replaced by other similarity functions. We will give some analyses in Section 3.

In Eq. (6),  $\mathbf{q}$  is the spatial quality map, which has the same resolution as the down-sampled image  $\mathbf{x}'$ . Finally, the overall quality of the degraded image can be derived through average pooling

$$Q = \frac{1}{Z} \sum_{i,j} \mathbf{q}(i, j), \quad (10)$$

where  $Z$  is a normalization factor.



**Fig. 3.** The reference and distorted images  $\mathbf{x}_r$ ,  $\mathbf{x}_d$ , the corresponding saliency maps  $\mathbf{m}_r$ ,  $\mathbf{m}_d$ , and quality maps  $\mathbf{q}'$ . Subscripts  $r$ ,  $d$  denote the reference and distorted images, respectively. Four types of distortions are illustrated.

### 2.3. Quality refining: contrast change

The method described in Sections 2.1 and 2.2 works fine in most situations since most types of distortions can cause saliency change. But it is not so efficient in some unusual circumstances. As illustrated in Fig. 3, saliency is not quite sensitive to contrast change, whereas contrast change has a great influence on subjective image quality [37]. Thus we need to do some specific refining for the contrast changed images.

Since the entropy and the luminance of an image are two important features which are sensitive to contrast change, we extract the entropy difference  $D_H$  and the luminance difference  $D_L$  for the refining

$$D_H = H_r - H_d, \quad (11)$$

$$D_L = L_r - L_d, \quad (12)$$

where  $H_r$ ,  $H_d$  and  $L_r$ ,  $L_d$  are the entropy and mean luminance of the reference and distorted images, respectively, and  $H$ ,  $L$  can be defined as

$$H = - \sum_i p_i \log(p_i), \quad (13)$$

$$L = \frac{1}{Z} \sum_{i,j} \mathbf{x}(i, j), \quad (14)$$

where  $\mathbf{p} = [p_1, \dots, p_i, \dots, p_{256}]$  denotes the histogram probability of the luminance of the down-sampled image, and  $Z$  is a normalization factor.  $H_r$  and  $L_r$  are also transmitted as RR features.

The refining is done by modifying the quality map  $\mathbf{q}$  directly

$$\mathbf{q}' = \mathbf{q}^{\hat{f}}, \quad (15)$$

where  $\hat{\cdot}$  indicates element-wise power, and  $f$  is a contrast change sensitive factor defined as

$$f = \begin{cases} k_1 |D_H| + k_2 |D_L| & \text{if } Q > \tau_1, |D_H| > \tau_2, \\ 1 & \text{otherwise} \end{cases}, \quad (16)$$

where  $D_H$ ,  $D_L$ ,  $Q$  are defined in Eqs. (11), (12) and (10), respectively,  $k_1$ ,  $k_2$  are two linear coefficients,  $\tau_1$ ,  $\tau_2$  are two thresholds used

to filter out the contrast changed images,  $k_1$ ,  $k_2$ ,  $\tau_1$ ,  $\tau_2$  are constants and empirically set. We will test SIRR's sensitivity to these parameters in Section 3. The modifying factor  $f$  acts as a punishment on those contrast changed images whose entropy and mean luminance are far deviated from the reference image. After modification, the final quality index of the image can be derived as

$$Q' = \frac{1}{Z} \sum_{i,j} \mathbf{q}'(i, j). \quad (17)$$

Some examples of  $\mathbf{q}'$  are shown in Fig. 3.

## 3. Validation of SIRR

The proposed SIRR measure is validated on both NSIs and SCIs. The details are as follows.

### 3.1. Experimental settings

#### 3.1.1. Test databases

As described in Section 1, the following two categories of IQA databases are chosen as test beds:

- NSI QA databases: LIVE [31], TID2008 [32], CSIQ [33] which are general-purpose IQA databases, together with a multiply distorted IQA database LIVEMD [34] and a contrast changed IQA database CID2013 [35]. The included images depict various natural scenes including indoor and outdoor views, humans and animals, etc. All five databases are frequently used in different IQA studies.
- SCI QA databases: A general-purpose SCI QA database SIQAD [2] and a compressed SCI QA database QACS [3]. Both databases include SCIs which cover a variety of computer operation scenes, e.g., web pages, documents, user interfaces, etc.

Table 1 gives an overview of all test databases. Fig. 4 illustrates some example NSIs and SCIs from two representative databases.

#### 3.1.2. Comparing algorithms

We compare the proposed SIRR with state-of-the-art FR and RR IQA algorithms. For FR measures, we consider the mainstream peak signal-to-noise ratio (PSNR), SSIM [30], VSNR [38], MSSSIM [39],

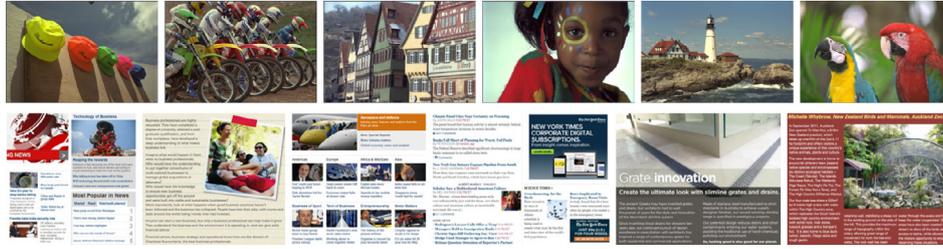


Fig. 4. Example NSIs and SCIs from two representative databases: LIVE (1st row) and SIQAD (2nd row).

Table 1  
Overview of the test databases.

Content type	Name	No. of ref.	No. of dist.	Distortion type
NSI	LIVE [31]	29	779	General
	TID2008 [32]	25	1700	General
	CSIQ [33]	30	866	General
	LIVEMD [34]	15	450	Multiple distortions
	CID2013 [35]	15	400	Contrast change
SCI	SIQAD [2]	20	980	General
	QACS [3]	24	492	Compression

GSIM [40] and SQMS [4]. As it for RR measures, we compare SIRR with RRED [41], FTQM [42], REDLOG [43] and RQMSH [44]. Among all competitors, SQMS and RQMSH are two recent quality measures specifically designed for SCIs. Whereas the other measures are implicitly designed for NSIs.

For RRED, we choose  $RRED_{16}^{M_{16}/16}$  as its score, which transmits  $L/576$  scalars as the reference data ( $L$  denotes the total number of pixels in the image). Whereas for FTQM, we adopt the third score ( $Q_{Phase}^{(2)}$ ) of FTQM as the objective quality. 403 scalars act as reference data, which requires approximately 1/490 of the reference information for a  $512 \times 384$  image [42].

3.1.3. Evaluation criteria

We follow the suggestion given by the video quality experts group (VQEG) [45]. The predicted objective scores are first nonlin-

early mapped using a five-parameter logistic function

$$q(s) = \beta_1 \left( \frac{1}{2} - \frac{1}{1 + \exp(\beta_2(s - \beta_3))} \right) + \beta_4 s + \beta_5, \quad (18)$$

where  $\beta_i$  ( $i = 1, 2, \dots, 5$ ) are parameters obtained from a nonlinear regression process,  $s$  and  $q(s)$  are predicted and mapped scores, respectively. Then the following three evaluation criteria are applied to assess the performance of all compared IQA algorithms.

- Spearman rank-order correlation coefficient (SRCC). It computes the monotonicity while ignoring the relative distance between the data

$$SRCC = 1 - \frac{6 \sum_{i=1}^N d_i^2}{N(N^2 - 1)}, \quad (19)$$

where  $d_i$  denotes the difference between the  $i$ th image's ranks in subjective and objective evaluations, and  $N$  is the number of images in the testing database.

- Pearson linear correlation coefficient (PLCC), which measures prediction linearity

$$PLCC = \frac{\sum_i (q_i - \bar{q}) \cdot (o_i - \bar{o})}{\sqrt{\sum_i (q_i - \bar{q})^2 \cdot (o_i - \bar{o})^2}}, \quad (20)$$

where  $o_i$  and  $q_i$  are the  $i$ -th image's subjective rating and the converted objective score after nonlinear mapping;  $\bar{o}$  and  $\bar{q}$  are mean values of  $o_i$  and  $q_i$ .

Table 2  
Performance comparison on NSI and SCI QA databases. We highlight the best-performing model in each row.

Type	Database	Metric	FR measures						RR measures				
			PSNR	SSIM	VSNR	MSSSIM	GSIM	SQMS	RRED	FTQM	REDLOG	RQMSH	SIRR
NSI	LIVE	SRCC	0.8756	0.9104	0.9279	0.9513	<b>0.9561</b>	0.9363	0.9169	0.9454	0.9456	0.7534	0.9489
		PLCC	0.8723	0.9042	0.9236	0.9489	<b>0.9512</b>	0.9301	0.9192	0.9426	0.9373	0.7562	0.9466
		RMSE	13.360	11.669	10.476	8.6181	<b>8.4323</b>	10.034	11.181	9.1218	9.5217	17.878	8.8078
	TID2008	SRCC	0.5531	0.6251	0.7045	<b>0.8542</b>	0.8504	0.7307	0.7952	0.7745	0.6864	0.5452	0.8073
		PLCC	0.5734	0.6413	0.6818	<b>0.8451</b>	0.8422	0.7596	0.7950	0.8063	0.6859	0.5479	0.8160
		RMSE	1.0994	1.0297	0.9813	<b>0.7173</b>	0.7234	0.8727	0.8140	0.7937	0.9765	1.1226	0.7758
	CSIQ	SRCC	0.8058	0.8369	0.8109	0.9133	0.9108	0.8643	0.8730	0.7762	0.8576	0.6462	<b>0.9144</b>
		PLCC	0.8000	0.8154	0.8005	0.8880	0.8964	0.8466	0.8662	0.8804	0.8565	0.6170	<b>0.9297</b>
		RMSE	0.1575	0.1520	0.1573	0.1207	0.1163	0.1397	0.1312	0.1245	0.1355	0.2066	<b>0.0967</b>
	LIVEMD	SRCC	0.6771	0.6459	0.7719	0.8363	0.8454	0.7565	<b>0.8765</b>	0.8553	0.8274	0.4116	0.8469
		PLCC	0.7419	0.7343	0.8117	0.8747	0.8808	0.8200	<b>0.9025</b>	0.8777	0.8658	0.5115	0.8772
		RMSE	12.681	12.839	11.045	9.1650	8.9562	10.825	<b>8.1463</b>	9.0628	9.4635	16.250	9.0796
	CID2013	SRCC	0.6649	0.8132	0.4306	<b>0.8580</b>	0.8372	0.8048	0.7153	0.8232	0.5977	0.6797	0.8229
		PLCC	0.6503	0.8125	0.4624	<b>0.8686</b>	0.8353	0.8215	0.7249	0.8316	0.5800	0.6997	0.8254
		RMSE	0.4734	0.3633	0.5525	<b>0.3088</b>	0.3426	0.3553	0.4293	0.3461	0.5077	0.4452	0.3518
SCI	SIQAD	SRCC	0.5605	0.7566	0.5703	0.6112	0.5483	<b>0.8803</b>	0.4926	0.1801	0.7315	0.7534	0.7286
		PLCC	0.5869	0.7561	0.5966	0.6195	0.5686	<b>0.8872</b>	0.4775	0.1388	0.7468	0.7555	0.7540
		RMSE	11.590	9.3676	11.487	11.236	11.775	<b>6.6039</b>	12.577	11.176	9.5192	9.3784	9.4031
	QACS	SRCC	0.8656	0.8683	0.7172	0.8922	0.8947	0.9096	0.9190	0.7881	0.8262	0.7786	<b>0.9227</b>
		PLCC	0.8679	0.8696	0.7050	0.8867	0.8921	0.9059	0.9166	0.7905	0.8217	0.7728	<b>0.9214</b>
		RMSE	1.1019	1.0953	1.5733	1.0257	1.0025	0.9396	0.8871	1.3588	1.2642	1.4079	<b>0.8623</b>
	Average	SRCC	0.7147	0.7796	0.7048	0.8452	0.8347	0.8404	0.7984	0.7347	0.7818	0.7849	<b>0.8560</b>
		PLCC	0.7275	0.7905	0.7117	0.8474	0.8381	0.8530	0.8003	0.7526	0.6526	0.6658	<b>0.8672</b>

**Table 3**  
Statistical significance matrix. Symbols '1'/'0' indicate that the row model is statically better/worse than the column model, respectively. Symbol '-' indicates that the row and column models are statically indistinguishable. Each seven symbols indicate the results on seven databases: LIVE, TID2008, CSIQ, LIVEMD, CID2013, SIQAD, and QACS, respectively.

	PSNR	SSIM	VSNR	MSSSIM	GSIM	SQMS	RRED	FTQM	REDLOG	RQMSH	SIRR
PSNR	-----	00--00-	00-01-1	00000--	00000-0	0000000	0000010	0000011	0000-01	1-11-01	0000000
SSIM	11--11-	-----	0--0111	000001-	0000-10	0000-00	-000110	0000-11	00001-1	11111-1	0000--0
VSNR	11-10-0	1--1000	-----	00000-0	00000-0	-00-000	1000010	0000010	0-00000	1111000	0000000
MSSSIM	11111--	111110-	11111-1	-----	-----1--	111-100	1110110	-1--111	111-101	1111101	-10-100
GSIM	11111-1	1111-01	11111-1	-----0--	-----	111-0-0	1110110	111--11	111-101	1111101	-10-100
SQMS	1111111	1111-11	-11-111	000-011	000--1-	-----	100-11-	000--11	-1--111	1111111	000--10
RRED	1111101	-11001	0111101	0001001	0001001	011-00-	-----	0--1011	01-1101	1111-01	000100-
FTQM	1111100	1111-00	1111101	-0--000	000--00	111--00	1--0100	-----	-11-10-	111110-	--0--00
REDLOG	1111-10	11110-0	1-11111	000-010	000-010	-0--000	10-0010	-00-01-	-----	11110-1	000-0-0
RQMSH	0-00-10	00000-0	0000111	0000010	0000010	0000000	0000-10	000001-	11110-1	-----	00000-0
SIRR	1111111	1111--1	1111111	-01-011	-01--11	111--01	111011-	--1--11	111-1-1	11111-1	-----

**Table 4**  
Reference data and computational cost comparison. The units of reference data are p: pixels; s: scalars; b: bits. Time: average running time for one image (tested on LIVE database).

	PSNR	SSIM	VSNR	MSSSIM	GSIM	SQMS	RRED	FTQM	REDLOG	RQMSH	SIRR
Type	FR	FR	FR	FR	FR	FR	RR	RR	RR	RR	RR
Reference data	$L p$	$\frac{L}{576} s$	403 s	6 s	1 s	$\frac{L}{64} + 16 b$					
Time (s/image)	0.0010	0.0741	0.3612	0.1297	0.0359	0.0583	0.9141	0.2084	0.6995	0.1105	0.0131

**Table 5**  
SRCC performance of using different saliency models and similarity functions.

Model	LIVE	TID2008	CSIQ	LIVEMD	CID2013	SIQAD	QACS
SIRR <sub>Itti</sub>	0.8738	0.6018	0.8571	0.8311	0.8178	0.3144	0.7752
SIRR <sub>GBVS</sub>	0.8525	0.5716	0.8256	0.8125	0.8175	0.0505	0.7560
SIRR <sub>SR</sub>	0.8433	0.6116	0.8728	0.5963	0.8261	0.0807	0.6914
SIRR <sub>PFT</sub>	0.8393	0.6349	0.8312	0.7819	0.8192	0.0284	0.7472
SIRR <sub>Li</sub>	0.9418	0.8080	0.9181	0.8266	0.8246	0.5544	0.9217
SIRR <sub>r</sub>	0.9500	0.8147	0.9168	0.8468	0.8248	0.7304	0.9224
SIRR <sub>s</sub>	0.9495	0.8192	0.9183	0.8439	0.8300	0.7167	0.9214
SIRR	0.9489	0.8073	0.9144	0.8469	0.8229	0.7286	0.9227

- Root-mean-square error (RMSE), which is an accuracy measure and calculates the difference between  $o_i$  and  $q_i$

$$RMSE = \sqrt{\frac{1}{N} (q_i - o_i)^2}. \quad (21)$$

### 3.2. Experimental results

Table 2 lists the performance comparison results on all seven benchmark databases.

#### 3.2.1. On NSI QA databases

The proposed SIRR generally outperforms FR measures like PSNR, SSIM, VSNR and SQMS on all five NSI QA databases. It is also comparable to the better measures like MSSSIM and GSIM, whose performance is quite close to SIRR on a lot of NSI QA databases. Whereas for the RR measures, SIRR generally performs better than RRED, REDLOG, RQMSH, and the performance of SIRR and FTQM is quite close. Note that SQMS and RQMSH are not comparable to most competitors on NSI QA databases, since they are specifically designed for SCIs.

#### 3.2.2. On SCI QA databases

The superiority of the proposed method is more obvious in SCIs, where most state-of-the-art IQA measures work not so well and suffer from certain degrees of performance drop. On QACS, SIRR performs better than all competitors. On SIQAD, SQMS shows the best performance, which is expectable since SQMS is optimized on this database. SIRR is close to SSIM, REDLOG, RQMSH, and they

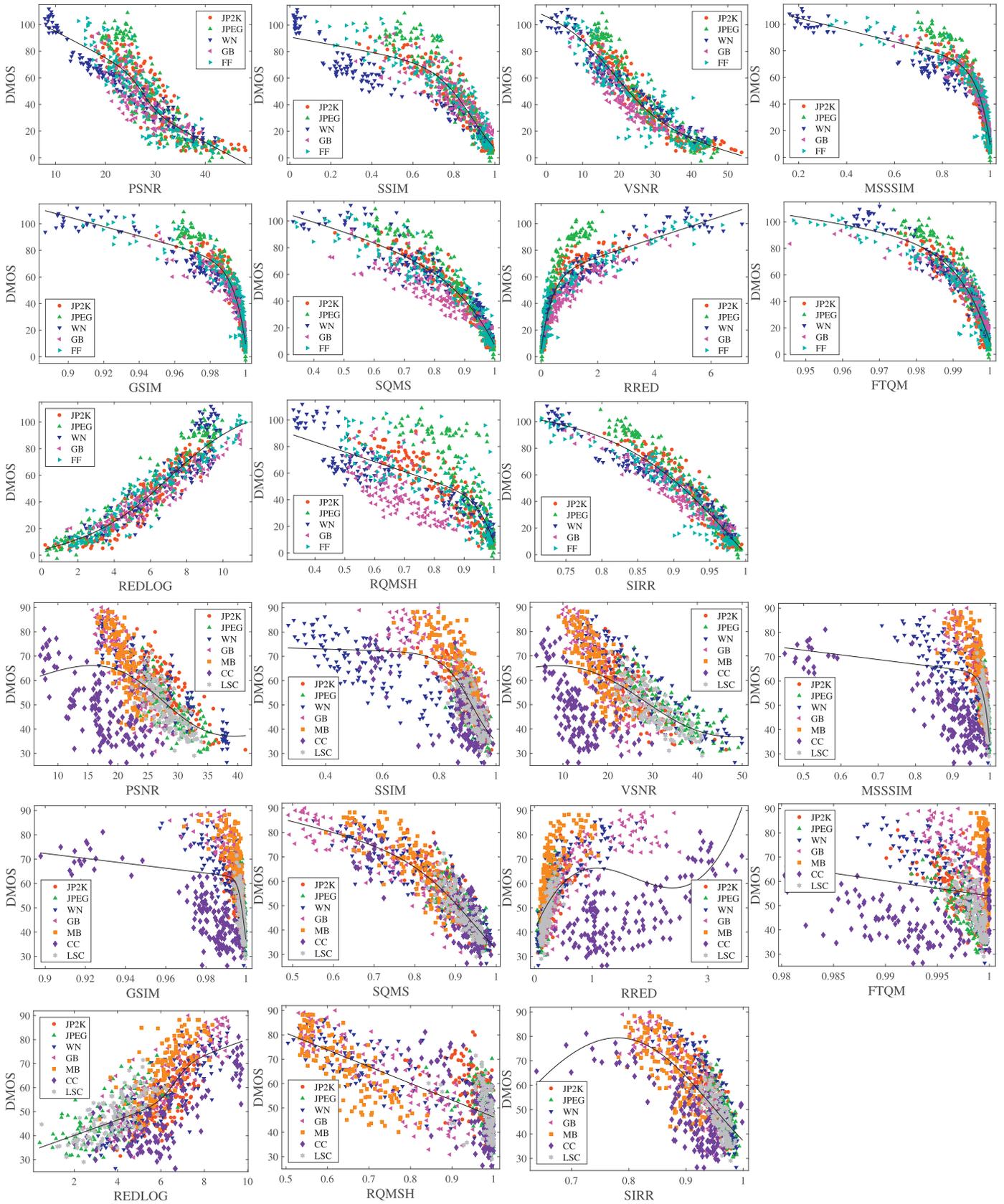
**Table 6**  
SRCC performance of different parameter settings.

	$k_1$	2	5	8	11	14
$k_1$						
TID2008	0.7982	0.8117	0.8073	0.8159	0.8149	
CSIQ	0.8840	0.9172	0.9144	0.9096	0.9019	
CID2013	0.8239	0.8239	0.8229	0.8239	0.8239	
SIQAD	0.7291	0.7310	0.7286	0.7259	0.7235	
$k_2$						
TID2008	0.02	0.05	0.08	0.11	0.14	
CSIQ	0.8158	0.8158	0.8073	0.8158	0.8158	
CSIQ	0.9166	0.9166	0.9144	0.9164	0.9164	
CID2013	0.8239	0.8239	0.8229	0.8239	0.8239	
SIQAD	0.7297	0.7295	0.7286	0.7271	0.7254	
$\tau_1$						
TID2008	0.95	0.96	0.97	0.98	0.99	
TID2008	0.8148	0.8156	0.8073	0.8167	0.7982	
CSIQ	0.9192	0.9192	0.9144	0.9097	0.8679	
CID2013	0.8239	0.8239	0.8229	0.8239	0.8239	
SIQAD	0.6995	0.7244	0.7286	0.7262	0.6648	
$\tau_2$						
TID2008	0.1	0.3	0.5	0.7	0.9	
TID2008	0.8172	0.8175	0.8073	0.8167	0.8097	
CSIQ	0.9256	0.9246	0.9144	0.9058	0.8884	
CID2013	0.8236	0.8239	0.8229	0.8239	0.8239	
SIQAD	0.7249	0.7249	0.7286	0.7249	0.7258	

outperform the rest measures. Note that we have not done any particular optimization for SCIs, but SIRR still performs pretty well on SIQAD and QACS. It indicates that the extracted features are efficient for both NSIs and SCIs. Table 2 also lists the average performance of all seven databases. As shown, SIRR outperforms all comparing algorithms from an average perspective.

#### 3.2.3. Significance tests

To testify if the quality prediction abilities between any two models are statistically different, we conduct a series of statistical significance tests. Following the strategy adopted in [46], we exam the quality prediction ability by comparing the variances of residuals between the nonlinear mapped scores  $q_i$  and subjective ratings  $o_i$ . Lower variance indicates better performance. Thus we conduct  $F$ -tests, whose test statistic is the ratio of two models' residual variances. The null hypothesis is that the residuals of two quality models are statistically indistinguishable (with 95% confidence) and they are from the same distribution. We compare every two models. Table 3 lists the significance test results on all seven



**Fig. 5.** Scatter plots of all compared methods on LIVE (top three rows) and SIQAD (bottom three rows) database. The (black) lines are curves fitted with the five-parameter logistic function. JP2K: JPEG2000, JPEG: JPEG, WN: white noise, GB: Gaussian blur, FF: fast fading, MB: motion blur, CC: contrast change, LSC: layer-segmentation based compression.

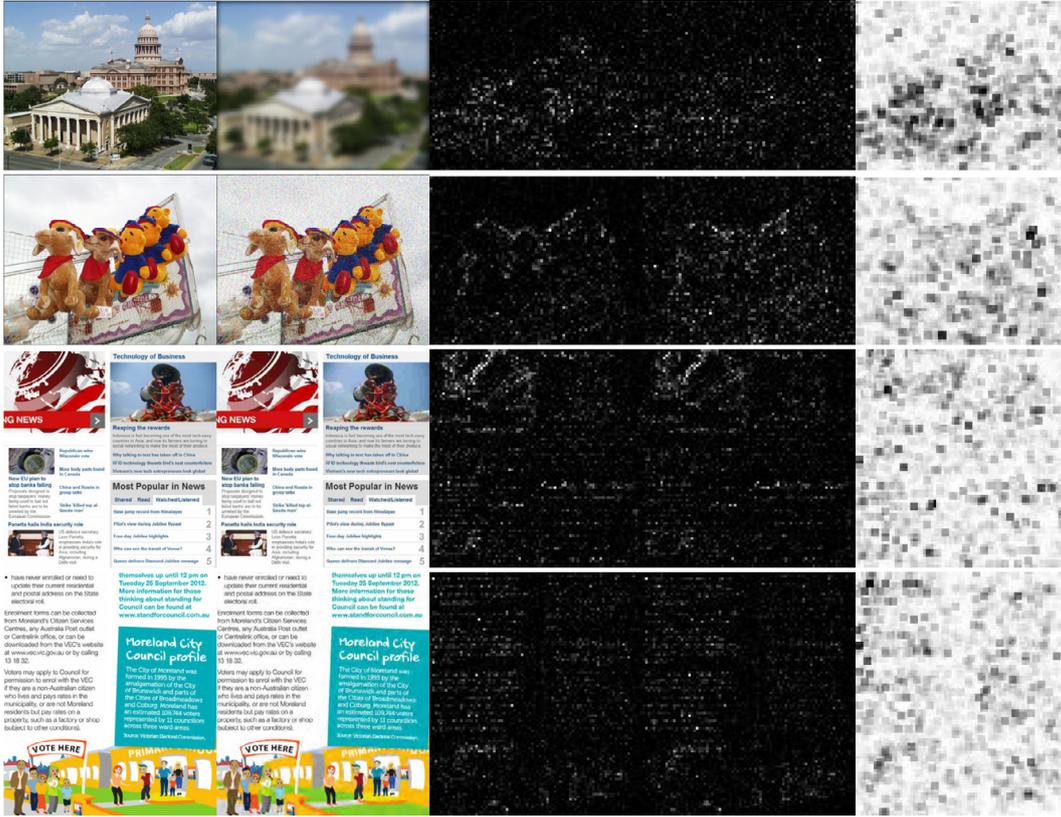


Fig. 6. Illustration of relative saliency and quality maps in SIRR. 1st, 2nd column: the reference, distorted images. 3rd, 4th column: saliency maps of the reference, distorted images. 5th column: quality map. 1st row: NSI, Gaussian blur. 2nd row: NSI, white noise. 3rd row: SCI, JPEG. 4th row: SCI, JPEG2000.

databases. From this table, we can get similar performance comparison results as described in the previous two paragraphs.

### 3.3. Reference data

The amount of reference data is also an important aspect of a RR IQA measure. A better RR measure should be able to achieve better performance using less reference data. For a gray-scale image with a total of  $L$  pixels, FR measures need to transmit all  $L$  pixels, and each pixel generally costs 8 bits.  $RRED_{16}^{M_{16}/16}$  is chosen as the quality score of RRED. It needs to transmit  $L/576$  scalars as the reference data. FTQM's reference data is composed of 403 scalars since we adopt the third score  $Q_{\text{Phase}}^{(2)}$ . That means approximately  $L/490$  scalars for a  $512 \times 384$  image [42]. REDLOG and RQMSH require 6 and 1 scalars as reference data, respectively. For the proposed SIRR, the reference data consists of three parts: a  $L/64$  binary image signature map, the mean luminance and entropy of the reference image, which cost  $L/64 + 16$  bits in total.  $RRED_{16}^{M_{16}/16}$  and  $Q_{\text{Phase}}^{(2)}$  are selected such that the amount of reference data is close to SIRR.

Table 4 gives an overview of all compared algorithms' reference data. Note that the authors' implementations of RRED, FTQM, REDLOG and RQMSH adopt the double-precision data type for the reference scalars. It is the default data type for floating-point numbers in MATLAB, and it takes 64 bits for one scalar, which is quite bit-costly. The reference data can be reduced by adopting some other low-bit data types, but the models may suffer from some performance drop because of the precision loss. SIRR does not have such problems since the majority of reference data is binary. Considering the low rate of reference data and the competitive performance, SIRR is quite promising.

### 3.4. Ablation experiment

#### 3.4.1. Influence of the saliency model

SIRR adopts the image signature model [29] to detect saliency. Other saliency models can act the same role. We replace the image signature model with several other saliency models, including Itti1998 [47], GBVS [48], SR [49], and PFT [50]. All other settings are kept the same except for the saliency model. We denote the corresponding RR measures as  $SIRR_{\text{Itti}}$ ,  $SIRR_{\text{GBVS}}$ ,  $SIRR_{\text{SR}}$ , and  $SIRR_{\text{PFT}}$ , respectively, whose performance is listed in Table 5. We can observe that other saliency models are effective too, but they may suffer from some degree of performance drop, which is not surprising since some settings of SIRR are optimized for the image signature model. The performance of using other saliency models can be improved through fine tuning, for example, using other similarity functions [8]. One advantage of using the image signature model is that the reference data is much less, since it uses a binary image descriptor, while other models generate gray scale saliency maps.

#### 3.4.2. Influence of the similarity function

We test three similarity functions, i.e.,  $\frac{2\mu_r\mu_d}{\mu_r^2 + \mu_d^2}$ ,  $\frac{2\sigma_r\sigma_d}{\sigma_r^2 + \sigma_d^2}$ ,  $\frac{2s_r s_d}{s_r^2 + s_d^2}$ , where  $\mu_r$ ,  $\mu_d$ ,  $\sigma_r$ ,  $\sigma_d$  and  $\sigma_{rd}$  are the local mean, variance and covariance of the reference and distorted images' saliency maps  $s_r$  and  $s_d$ , respectively. The corresponding RR measures are denoted as  $SIRR_{\mu}$ ,  $SIRR_{\sigma}$ , and  $SIRR_s$ , respectively, whose performance is also listed in Table 5. We can observe that the performance is very close, which means that the form of the similarity function is not the key factor of SIRR.

### 3.5. Parameter sensitivity

We test the sensitivity of SIRR with respect to four parameters— $k_1$ ,  $k_2$ ,  $\tau_1$ ,  $\tau_2$ , which are introduced during quality refining. Since the quality refining is introduced specifically for distortions such as mean shift and contrast change, we only test parameter sensitivity on four databases including such distortions, i.e., TID2008, CSIQ, CID2013, and SIQAD. When varying one parameter, other parameters are fixed at the default values. The parameter range and step can be found in Table 6, from which we can observe that the SRCC performance remains stable within a significantly wide range.

### 3.6. Computational cost

Computational cost is also one important aspect of an algorithm. We test the computational cost of all compared methods on LIVE database. Table 4 lists the average running time for each image. Here the running time includes all feature extraction and comparison procedures at both the sender and receiver sides. All algorithms are tested on MATLAB R2013a platform which is operated on a computer with Intel I5-3470 CPU @3.20GHz and 4 GB RAM. The proposed SIRR is the fastest among all compared methods except PSNR. Compared with state-of-the-art RR methods, SIRR does not involve too many complicate transform domain processes. Moreover, the most computationally costly processes are executed in the down-sampled image whose resolution is only one over sixty-four of the original resolution. It significantly reduces the computational cost of SIRR, which makes it possible for real-time applications.

### 3.7. Visualization

We select two representative databases, LIVE for NSIs and SIQAD for SCIs, and illustrate the scatter plots of all compared methods in Fig. 5. As analyzed, the quality prediction abilities of SIRR, MSSSIM, GSIM, FTQM, REDLOG are quite close on LIVE, and they outperform the rest methods. Whereas on SIQAD, SQMS is the best. SIRR, SSIM, REDLOG and RQMSH are close, and they outperform the rest significantly. As illustrated in the scatter plots, SIRR is quite balanced for different types of distortions. What is more important is that it is efficient for both NSI and SCI.

To give an intuitive illustration of SIRR, Fig. 6 shows some relative saliency maps and quality maps. We select two NSIs and two SCIs degraded by four most common distortions, i.e., Gaussian blur, white noise, JPEG, JPEG2000. Saliency maps of both the reference and distorted images, and the quality maps are shown. As illustrated, quality degradation can significantly change the detected saliency, which is the basis of the proposed method.

## 4. Conclusion

In this paper, we propose a saliency-induced reduced-reference (SIRR) IQA measure for the two most common but quite different types of images encountered in realistic multimedia communication systems, i.e., NSI and SCI. We develop SIRR based on the observations that quality degradation can significantly affect saliency detection, and that saliency detection is in fact an operation of dimension and data reduction. SIRR evaluates quality by measuring the similarity between two images' saliency maps, which are described through a binary image descriptor, i.e., image signature. Reference data is significantly reduced through down-sampling and the binary image descriptor. We validate SIRR on five large-scale NSI QA databases and two existing SCI QA databases. Validation results show that SIRR is comparable to well-performed FR IQA measures like MSSSIM and GSIM in NSI. What is more important is that SIRR outperforms most of the state-of-the-art FR and RR IQA

measures in SCI. The reference data rate is relatively low and it is computationally costless. It is worth noting that SIRR does not involve any image type classification process, but it is still efficient for both NSI and SCI under the circumstance that there is not any prior knowledge about the image types.

## Acknowledgments

This work was supported in part by National Natural Science Foundation of China under grants 61422112, 61371146, 61521062, and 61527804.

## References

- [1] J. Xu, R. Joshi, R.A. Cohen, Overview of the emerging HEVC screen content coding extension, *IEEE Trans. Circuits Syst. Video Technol.* 26 (1) (2016) 50–62.
- [2] H. Yang, Y. Fang, W. Lin, Perceptual quality assessment of screen content images, *IEEE Trans. Image Process.* 24 (11) (2015) 4408–4421.
- [3] S. Wang, K. Gu, X. Zhang, W. Lin, L. Zhang, S. Ma, W. Gao, Subjective and objective quality assessment of compressed screen content images, *IEEE J. Emerging Sel. Top. Circuits Syst.* 6 (4) (2016) 532–543.
- [4] K. Gu, S. Wang, H. Yang, W. Lin, G. Zhai, X. Yang, W. Zhang, Saliency-guided quality assessment of screen content images, *IEEE Trans. Multimedia* 18 (6) (2016) 1098–1110.
- [5] K. Gu, G. Zhai, W. Lin, X. Yang, W. Zhang, Learning a blind quality evaluation engine of screen content images, *Neurocomputing* 196 (2016) 140–149.
- [6] K. Gu, J. Zhou, J. Qiao, G. Zhai, W. Lin, A. Bovik, No-reference quality assessment of screen content pictures, *IEEE Trans. Image Process.* 26 (8) (2017) 4005–4018.
- [7] 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.
- [8] 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.
- [9] Y. Liu, J. Yang, Q. Meng, Z. Lv, Z. Song, Z. Gao, Stereoscopic image quality assessment method based on binocular combination saliency model, *Signal Process.* 125 (2016) 237–248.
- [10] A. Saha, Q.J. Wu, Full-reference image quality assessment by combining global and local distortion measures, *Signal Process.* 128 (2016) 186–197.
- [11] F. Gao, J. Yu, Biologically inspired image quality assessment, *Signal Process.* 124 (2016) 210–219.
- [12] K. Gu, L. Li, H. Lu, X. Min, W. Lin, A fast reliable image quality predictor by fusing micro-and macro-structures, *IEEE Trans. Ind. Electron.* 64 (5) (2017) 3903–3912.
- [13] X. Min, G. Zhai, Z. Gao, K. Gu, Visual attention data for image quality assessment databases, in: *Proceedings of the IEEE International Symposium on Circuits and Systems*, 2014, pp. 894–897.
- [14] G. Zhai, X. Wu, X. Yang, W. Lin, W. Zhang, A psychovisual quality metric in free-energy principle, *IEEE Trans. Image Process.* 21 (1) (2012) 41–52.
- [15] G. Zhai, J. Cai, W. Lin, X. Yang, W. Zhang, M. Etoh, Cross-dimensional perceptual quality assessment for low bit-rate videos, *IEEE Trans. Multimedia* 10 (7) (2008) 1316–1324.
- [16] K. Gu, V. Jakhethiya, J.-F. Qiao, X. Li, W. Lin, D. Thalmann, Model-based referenceless quality metric of 3d synthesized images using local image description, *IEEE Trans. Image Process.* (2017).
- [17] X. Min, G. Zhai, K. Gu, Y. Fang, X. Yang, X. Wu, J. Zhou, X. Liu, Blind quality assessment of compressed images via pseudo structural similarity, in: *Proceedings of the IEEE International Conference on Multimedia and Expo*, 2016, pp. 1–6.
- [18] J. Xu, P. Ye, Q. Li, H. Du, Y. Liu, D. Doermann, Blind image quality assessment based on high order statistics aggregation, *IEEE Trans. Image Process.* 25 (9) (2016) 4444–4457.
- [19] X. Min, K. Ma, K. Gu, G. Zhai, Z. Wang, W. Lin, Unified blind quality assessment of compressed natural, graphic, and screen content images, *IEEE Trans. Image Process.* 26 (11) (2017) 5462–5474.
- [20] A. Borji, L. Itti, State-of-the-art in visual attention modeling, *IEEE Trans. Pattern Anal. Mach. Intell.* 35 (1) (2013) 185–207.
- [21] X. Min, G. Zhai, Z. Gao, C. Hu, Influence of compression artifacts on visual attention, in: *Proceedings of the IEEE International Conference on Multimedia and Expo*, 2014, pp. 1–6.
- [22] W. Zhang, A. Borji, Z. Wang, P. Le Callet, H. Liu, The application of visual saliency models in objective image quality assessment: a statistical evaluation, *IEEE Trans. Neural Netw. Learn. Syst.* 27 (6) (2016) 1266–1278.
- [23] W. Zhang, H. Liu, Study of saliency in objective video quality assessment, *IEEE Trans. Image Process.* 26 (3) (2017) 1275–1288.
- [24] X. Min, G. Zhai, K. Gu, X. Yang, Fixation prediction through multimodal analysis, *ACM Trans. Multimedia Comput. Commun. Appl.* 13 (1) (2016) 6:1–6:23.
- [25] X. Min, G. Zhai, C. Hu, K. Gu, Fixation prediction through multimodal analysis, in: *Proc. IEEE Vis. Commun. Image Process.*, 2015, pp. 1–4.
- [26] X. Min, G. Zhai, K. Gu, Visual attention on human face, in: *Proc. IEEE Vis. Commun. Image Process.*, 2015, pp. 1–4.
- [27] X. Min, G. Zhai, K. Gu, J. Liu, S. Wang, X. Zhang, X. Yang, Visual attention analysis and prediction on human faces, *Inf. Sci.* 420 (2017) 417–430.

- [28] Q. Wei, G. Zhai, C. Hu, X. Min, Visual attention analysis and prediction on human faces with mole, in: Proc. IEEE Vis. Commun. Image Process., 2016, pp. 1–4.
- [29] X. Hou, J. Harel, C. Koch, Image signature: highlighting sparse salient regions, IEEE Trans. Pattern Anal. Mach. Intell. 34 (1) (2012) 194–201.
- [30] Z. Wang, A.C. Bovik, H.R. Sheikh, E.P. Simoncelli, Image quality assessment: from error visibility to structural similarity, IEEE Trans. Image Process. 13 (4) (2004) 600–612.
- [31] H.R. Sheikh, Z. Wang, L. Cormack, A.C. Bovik, LIVE image quality assessment database release 2, [Online]. Available: <http://live.ece.utexas.edu/research/quality>.
- [32] N. Ponomarenko, V. Lukin, A. Zelensky, K. Egiazarian, M. Carli, F. Battisti, TID2008—a database for evaluation of full-reference visual quality assessment metrics, Adv. Mod. Radioelectron. 10 (4) (2009) 30–45.
- [33] E.C. Larson, D.M. Chandler, Most apparent distortion: full-reference image quality assessment and the role of strategy, J. Electron. Imaging 19 (1) (2010) 011006.
- [34] D. Jayaraman, A. Mittal, A.K. Moorthy, A.C. Bovik, Objective quality assessment of multiply distorted images, in: Proc. IEEE Asilomar Conf. Sig., Syst. Comput., 2012, pp. 1693–1697.
- [35] K. Gu, G. Zhai, X. Yang, W. Zhang, M. Liu, Subjective and objective quality assessment for images with contrast change, in: Proc. IEEE Int. Conf. Image Process., 2013, pp. 383–387.
- [36] W. Lin, C.-C. J. Kuo, Perceptual visual quality metrics: a survey, J. Vis. Commun. Image Represent. 22 (4) (2011) 297–312.
- [37] K. Gu, G. Zhai, X. Yang, W. Zhang, C.W. Chen, Automatic contrast enhancement technology with saliency preservation, IEEE Trans. Circuits Syst. Video Technol. 25 (9) (2015) 1480–1494.
- [38] D.M. Chandler, S.S. Hemami, VSNR: a wavelet-based visual signal-to-noise ratio for natural images, IEEE Trans. Image Process. 16 (9) (2007) 2284–2298.
- [39] Z. Wang, E.P. Simoncelli, A.C. Bovik, Multiscale structural similarity for image quality assessment, in: Proc. IEEE Asilomar Conf. Sig., Syst., Comput., 2, 2003, pp. 1398–1402.
- [40] A. Liu, W. Lin, M. Narwaria, Image quality assessment based on gradient similarity, IEEE Trans. Image Process. 21 (4) (2012) 1500–1512.
- [41] R. Soundararajan, A.C. Bovik, RRED indices: reduced reference entropic differencing for image quality assessment, IEEE Trans. Image Process. 21 (2) (2012) 517–526.
- [42] M. Narwaria, W. Lin, I.V. McLoughlin, S. Emmanuel, L.-T. Chia, Fourier transform-based scalable image quality measure, IEEE Trans. Image Process. 21 (8) (2012) 3364–3377.
- [43] S. Golestaneh, L.J. Karam, Reduced-reference quality assessment based on the entropy of DWT coefficients of locally weighted gradient magnitudes, IEEE Trans. Image Process. 25 (11) (2016) 5293–5303.
- [44] S. Wang, K. Gu, X. Zhang, W. Lin, S. Ma, W. Gao, Reduced-reference quality assessment of screen content images, IEEE Trans. Circuits Syst. Video Technol. PP (99) (2016) 1.
- [45] VQEG, Final report from the Video Quality Experts Group on the validation of objective models of video quality assessment, 2003, [Online]. <http://www.vqeg.org/>.
- [46] H.R. Sheikh, M.F. Sabir, A.C. Bovik, A statistical evaluation of recent full reference image quality assessment algorithms, IEEE Trans. Image Process. 15 (11) (2006) 3440–3451.
- [47] L. Itti, C. Koch, E. Niebur, A model of saliency-based visual attention for rapid scene analysis, IEEE Trans. Pattern Anal. Mach. Intell. 20 (11) (1998) 1254–1259.
- [48] J. Harel, C. Koch, P. Perona, Graph-based visual saliency, in: Adv. Neural Inf. Process. Syst., 2006, pp. 545–552.
- [49] X. Hou, L. Zhang, Saliency detection: a spectral residual approach, in: Proc. IEEE Conf. Comput. Vis. Pattern Recognit., 2007, pp. 1–8.
- [50] C. Guo, Q. Ma, L. Zhang, Spatio-temporal saliency detection using phase spectrum of quaternion fourier transform, in: Proc. IEEE Conf. Comput. Vis. Pattern Recognit., 2008, pp. 1–8.