Reduced-Reference Quality Assessment Based on the Entropy of DWT Coefficients of Locally Weighted Gradient Magnitudes

SeyedAlireza Golestaneh, *Student Member, IEEE*, and Lina J. Karam, *Fellow, IEEE*

***Abstract*— Perceptual image quality assessment (IQA) attempts to use computational models to estimate the image quality in accordance with subjective evaluations. Reduced-reference IQA (RRIQA) methods make use of partial information or features extracted from the reference image for estimating the quality of distorted images. Finding a balance between the number of RR features and accuracy of the estimated image quality is essential and important in IQA. In this paper, we propose a training-free low-cost RRIQA method that requires a very small number of RR features (six RR features). The proposed RRIQA algorithm is based on the discrete wavelet transform (DWT) of locally weighted gradient magnitudes. We apply human visual system's contrast sensitivity and neighborhood gradient information to weight the gradient magnitudes in a locally adaptive manner. The RR features are computed by measuring the entropy of each DWT subband, for each scale, and pooling the subband entropies along all orientations, resulting in L RR features (one average entropy per scale) for an L-level DWT. Extensive experiments performed on seven large-scale benchmark databases demonstrate that the proposed RRIQA method delivers highly competitive performance as compared with the state-of-the-art RRIQA models as well as full reference ones for both natural and texture images. The MATLAB source code of *REDLOG* and the evaluation results are publicly available online at https://http://lab.engineering.asu.edu/ivulab/software/redlog/.**

***Index Terms*— Reduced reference quality assessment (RRIQA), entropy, contrast sensitivity, gradient magnitude, locally adaptive weighting, discrete wavelet transform.**

I. Introduction

IMAGE and video compression and applications of visual media continue to be in high demand these days. There has been an increasing demand for accurate image and video quality assessment algorithms for different applications, such as image/video compression, communication, printing, display, restoration, segmentation, and fusion. In the past decade, a large number of methods have been designed to evaluate the quality of an image. The objective of image quality assessment (IQA) is to provide computational models to measure the quality of an image as perceived by human subjects. Specifically, existing IQA methods can be categorized into subjective and objective methods [1]. Subjective IQA directly assesses image quality by human subjects. Although this approach is accurate and reliable, it is expensive and time-consuming for real-world applications. In objective IQA, on the other hand, the goal is to provide computational models that can automatically predict the perceived image quality. Objective quality metrics can be divided into full-reference (reference available or FR), no-reference (reference not available or NR), and reduced-reference (RR) methods based on the availability of a reference image.

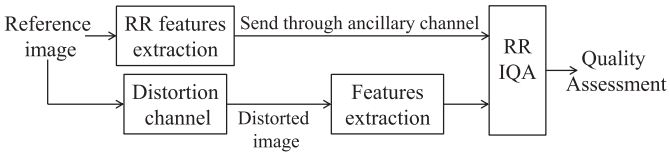

Fig. 1. The general framework for RRIQA systems.

FR methods [2]–[8] usually provide the most precise evaluation results and perform well in predicting the quality scores of human subjects in comparison with NR and RR methods. However, in many practical applications, an IQA system does not have access to reference images. Without the reference image, the IQA task becomes very difficult. NR IQA algorithms can mainly be categorized into two approaches. One approach is to design NR algorithms for a specific type of distortion (such as blocking, blurring, and compression) [9]–[11]. Another approach is training-based and makes use of training methods on databases to estimate quality [12]–[14].

RRIQA methods provide a solution when the reference image is not fully accessible. These methods generally operate by extracting a minimal set of parameters from the reference image (RR features); these parameters are later used with the distorted image to estimate quality. Figure 1 shows the general framework for RRIQA systems. RRIQA systems include a feature extraction process at the sender side for the reference image and a feature extraction at the receiver side for the distorted image. The extracted reference features, also known as side information, usually have a much lower data rate than the image data and are typically transmitted to the receiver through an ancillary channel [1].

In designing RRIQA algorithms it is important to select the features in a way to be sensitive to a variety of image distortions and be relevant to the visual perception of image quality. Another important factor is to maintain a good balance

Manuscript received June 4, 2015; revised February 28, 2016 and May 16, 2016; accepted July 4, 2016. Date of publication August 24, 2016; date of current version September 16, 2016. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Peter Tay.

The authors are with the Image, Video and Usability Laboratory, School of Electrical, Computer and Energy Engineering, Arizona State University, Tempe, AZ 85287-5706 USA (e-mail: sgolest1@asu.edu; karam@asu.edu).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TIP.2016.2601821

between the data rate of RR features and the accuracy of image quality prediction. For designing an effective RR quality metric, we need to consider not only its performance, but also its RR data rate for representing the extracted features. Therefore, how to balance the RR data rate and the performance prediction is important for the RR quality metric development. In this paper we propose a training-free general-purpose RRIQA algorithm. Our proposed algorithm exploits the characteristics of the human visual system, including contrast sensitivity and the perceptual significance of edges, for extractingand enhancing perceptually-significant features from the considered images. For this purpose, gradient magnitudes are computed and are weighted in a locally adaptive manner based on the human visual system's contrast sensitivity and neighborhood gradient information. For an L-level wavelet decomposition of the weighted gradient magnitude image, L RRfeatures are obtained by computing for each wavelet decomposition level, the entropy of DWT coefficients at that level.

This paper is organized as follows. An overview of existing popular RRIQA metrics is given in Section II. Section III presents the proposed RRIQA index. Performance results are presented in Section IV, followed by a conclusion in Section V.

II. Existing RRIQA Methods

RRIQA methods can be classified into three categorizes based on their feature selection methods and their targeted applications [1], [15]. The first category is based on modeling image distortions. RRIQA designed based on this approach are mostly developed for specific distortion types or application environments [16]–[19]. The second category is based on modeling the human visual system [20], [21], where perceptual features motivated from computational models of low-level vision are extracted to provide a reduced description of the image. The third type of approach is based on modeling image statistics and information in different transform domains [22]–[34]. RRIQA methods can be further divided into two main categories [1], [15], training-free and training-based methods. In training-based approaches, training or machine learning methods are used to optimize the IQA performance. Among RRIQA algorithms just a few of them [22]–[28] are general-purpose training-free methods and the rest are either limited in their ability to achieve good performance across different distortion types or involve training on databases or depend on parameters that need to be separately tuned based on prior knowledge of artifacts that are present in the image. Moreover, ztheir performance degrades with the reduction in the data rate (side information) required from the reference.

A. Training-Free Approaches

In [18], Gunawan *et al.* proposed a combined blockiness/blurriness detection method based on a frequency-domain harmonic amplitude analysis that requires a subset of the image information. In [19], Chono *et al.* used distributed source coding for remotely monitoring the image quality by transmitting a Slepian-Wolf encoded bitstream as a feature vector from the original image. In [20] and [21], Cornec *et al.* modeled the human visual system to extract perceptual features motivated from computational models of low-level vision. The concept of quality-aware images was proposed in [22], where partial reference-image information is embedded within the image and can be reliably extracted despite distortions. Based on results in natural image statistics, Wang *et al.* [22] proposed the wavelet-domain natural image statistic metric (WNISM). They modeled the marginal probability distribution of the wavelet coefficients of a natural image using a generalized Gaussian density (GGD) function. The Kullback-Leibler divergence (KLD) [35] is used to depict the distribution difference between the parametrized distribution of the reference and the empirical distribution of the distorted image.

In [23], Xue *et al.* employed the Weibull distribution to describe the statistics of the image gradient magnitude. They proposed an algorithm (βW-SCM) based on the steerable pyramid, where an image is divided into a collection of subbands localized in both scale and orientation. The strongest component map (SCM) is constructed for each scale. This map is composed, at each location, of the coefficient with maximum amplitude among all orientations for the considered scale. So at each location, only the strongest component is kept while other weaker ones are suppressed. Then, the Weibull distribution is employed to describe the statistics of the SCM. The scale and shape parameters of the Weibull distribution are estimated by fitting the Weibull function to the coefficients of the SCM with maximum likelihood estimation. The Weibull scale parameters, one for each pyramid level, represent the RR features. In [24], Zhang *et al.* proposed an edge pattern verification. They utilized the statistics of edges for developing the RRIQA (RR-Edge).

In [25] and [26], Ma *et al.* proposed a RRIQA method based on GGD modeling of the coefficient distributions of the reorganized DCT (RDCT) subbands. After applying the RDCT, they categorized the block-based DCT coefficients into a three-level coefficient tree, resulting in ten RDCT subbands and then modeled each RDCT subband by the GGD function. The city-block distance (CBT) and mutual information (MI) values were employed to depict the intra and inter RDCT relationships, respectively. The frequency ratio descriptor (FRD) calculated in the RDCT domain is used to depict the frequency distribution of the considered images. Finally the RRIQA is obtained by linearly combining the intra RDCT subband GGD modeling, inter RDCT subband MI values, and the FRD value. In [27], Wu *et al.* proposed a RRIQA index based on evaluating the visual content fidelities of the primary visual information and the residual uncertainty. They used an autoregressive (AR) model to predict the visualcontent and to decompose the input image into two portions, the orderly portion and the disorderly portion, where the orderly portion possessed the primary visual information of the input scene and the disorderly portion consists of the residual uncertainty. Then the fidelities of the two types of information are separately evaluated for quality assessment.

In [28], Soundararajan *et al.* developed a training-free RRIQA framework (RRED) based on an information theoretic framework. The image quality is measured via the difference

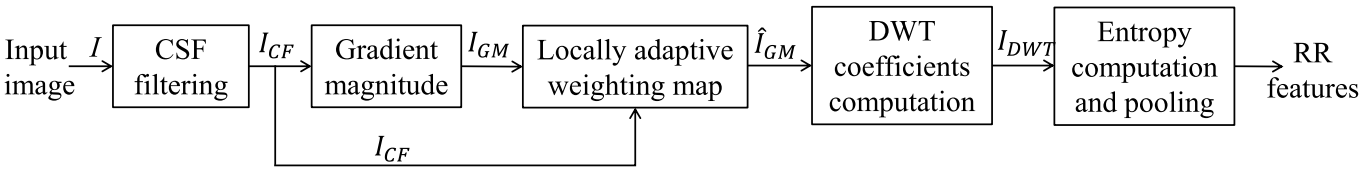

Fig. 2. Block diagram illustrating the computation of the RR features for the proposed RRIQA index, where I_{CF} represents the output of the CSF filtering of the input image I. I_{GM} and $\hat{I}_{GM}$ represent, respectively, the gradient magnitude map and the proposed locally weighted gradient magnitude map. I_{DWT} represents the wavelet coefficients of $\hat{I}_{GM}$.

between the entropies of wavelet coefficients of reference and distorted images.

B. Training/Tuning-Based Approaches

In [16], Engel *et al.* proposed a RRIQA metric for use in wireless imaging by considering different structural information observed in the distortion model of wireless links. In [29], Gao *et al.* proposed a RRIQA framework based on incorporating multiscale geometry analysis (MGA), contrast sensitivity function (CSF), and the Weber's law of just noticeable difference (JND). In this scheme, images are represented by normalized histograms, which correspond to visually sensitive coefficients. The quality of a distorted image is measured by comparing the normalized histogram of the distorted image to that of the reference image. Although this algorithm performs well, it needs to tune the JND threshold free parameter for different distortions.

In [31], Lin *et al.* proposed a method based on the average directional information (ADI), which is obtained from complex wavelet coefficients. The directional information is represented by the inter-coefficient product (ICP) obtained from the complex wavelet domain. In [32], Liu *et al.* proposed a RRIQA algorithm (SPCRM), which measures the difference of the regularity of the phase congruency (PC) between the reference image and the distorted image. The features are extracted from the PC via fractal analysis. Then the image features are pooled as the quality score using the ℓ_1 norm. The results depend on a tunable parameter, the image block size, to estimate the fractal dimension of the image surfaceand also to determine the feature length. Two versions were proposed in [32], one based on the intensity image (SPCRM-INT) and the other based on the partial gradient image (SPCRM-SCHARR).

Modeling the marginal probability distribution of the wavelet coefficients of natural images using the GGD is further extended in [33], where an additional divisive normalization transform (DNT) [36] step is introduced before computing the KLD. In [33], Li *et al.* proposed an algorithm that makes use of Gaussian scale mixtures (GSM) to model the natural image statistics. The KLD between the DNT coefficients' distributions of the reference and distorted images is used to represent the image perceptual quality. We refer to this method as RR-DNT. In [34], Rehman *et al.* proposed a general-purpose RRIQA (RR-SSIM) based on natural image statistics modeling, which is motivated by the success of the FR SSIM index. RR-SSIM combined the GSM-based statistics in a multi-scale and multi-orientation DNT-domain. A regression-by-discretization method is then applied to fit the metric to the FR SSIM index during a training stage and to normalize the measure across image distortion types.

III. Proposed RRIQA Index

In the proposed RRIQA method, given a reference or distorted image I as an input, RR features are computed as shown in Figure 2. After computing the RR features for both the reference and distorted images, the quality of the distorted image is estimated by computing a weighted sum of squared difference between the reference and distorted image features. For computing the RR features (Figure 2), the input image (reference or distorted) is first processed by applying a Contrast Sensitivity Function (CSF) based filter resulting in the filtered image I_{CF}. The gradient magnitude map, I_{GM}, of the resulting filtered image, I_{CF}, is then computed and locally weighted An L-level DWT of the weighted gradient magnitude image is performed. Details about the CSF-based filtering, local adaptive weighting of gradient magnitudes, DWT and entropy computations are given below. Figure 3 illustrates the steps of the proposed RRIQA method on a reference and distorted image. The RR features are obtained by computing the entropy of each DWT subband and, for each scale, averaging the subband entropies along all orientations, resulting in L RR features for an L-level DWT.

A. CSF Based Filtering

The CSF measures the sensitivity of the HVS to the various frequencies of visual stimuli. Here we apply an adjusted CSF model given by:

$$H_{CSF}(f,\theta) = H_1(f,\theta)H_2(f,\theta), \tag{1}$$

where f denotes the radial spatial frequency in cycles per degree of visual angle (c/deg), and $\theta \in [-\pi, \pi]$ denotes the angular frequency. In (1), $H_1(f,\theta)$ is the frequency response of a circularly symmetric Gaussian filter and $H_2(f,\theta)$ is the frequency response of a CSF model originally proposed by Mannos and Sakrison [37] and adjusted by Daly [38]. The frequency response of the Gaussian filter, $H_1(f,\theta)$, is given by:

$$H_1(f,\theta) = \exp(-2\pi^2\sigma^2 f^2), \tag{2}$$

where σ is a parameter that controls the cutoff of the filter. In our implementation, we an 8×8 filter size and a small $\sigma = 0.5$, which results in a lowpass filter with a relatively very high cutoff frequency in order to filter out very high frequency

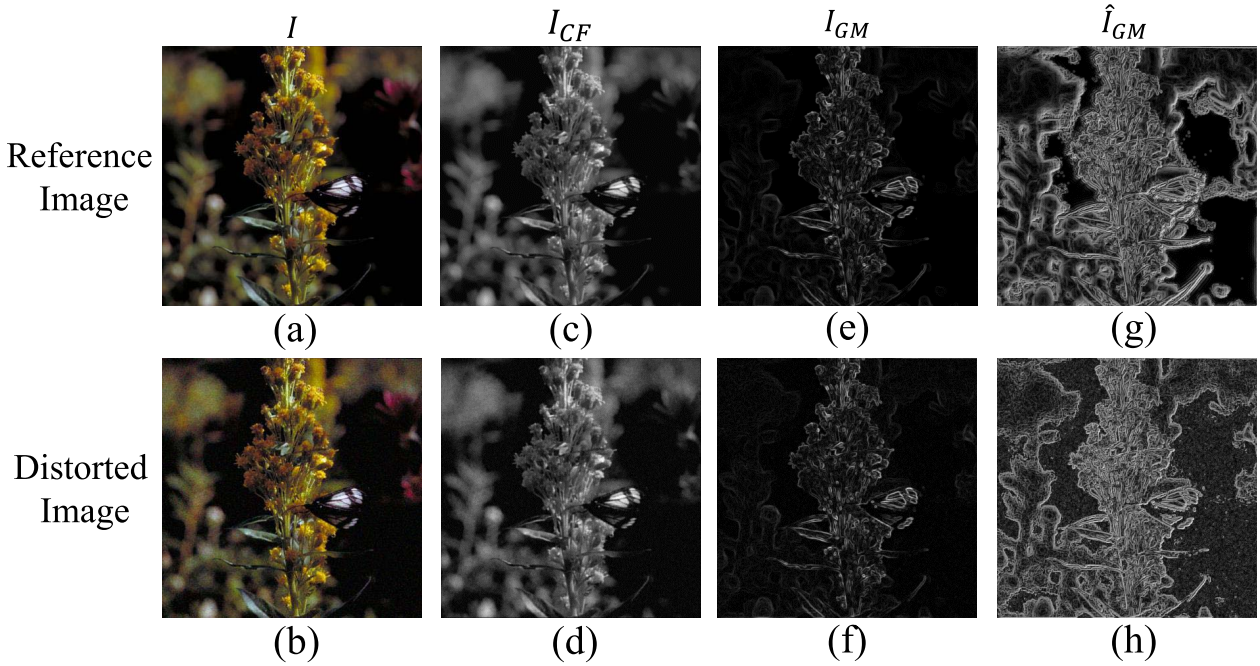

Fig. 3. Illustration of the steps of the proposed RRIQA method. (a) and (b) reference and distorted input images; (c) and (d) CSF and Gaussian filtered images; (e) and (f) gradient magnitude images; (g) and (h) locally weighted gradient images computed via (4).

components that are perceptually insignificant without significantly affecting the perceptually relevant edge components. $H_2(f,\theta)$ is given by [37], [38]:

$$H_2(f,\theta) = \begin{cases} 2.6(0.0192 + \lambda f_\theta)\exp[-\lambda f_\theta], & f \geq f_{\text{peak}} \\ 0.981, & \text{otherwise,} \end{cases} \tag{3}$$

where $\lambda = 0.114$ as in [3], [39], and [40], and $f_\theta = \frac{f}{0.15cos(4\theta)+0.85}$ accounts for the oblique effect [40]; in other words, f_θ represents an orientation-based modification of f which leads to a decrease in the contrast sensitivity along diagonal orientations.

The adjusted CSF model (1) is applied by filtering the input image, where the filtering can be performed in the frequency domain via $I_{CF} = F^{-1}[\tilde{H}_{CSF}(u,v)F[I]]$, where $F[\cdot]$ and $F^{-1}[\cdot]$ denote the DFT and inverse DFT, respectively. The quantity $\tilde{H}_{CSF}(u,v)$ denotes a DFT-based version of $H_{CSF}(f,\theta)$, where u, v are the DFT indices. The resulting filtered image is denoted as I_{CF}. Examples of I_{CF} are shown in Figures 3(c) and (d) for a reference image and corresponding distorted version, respectively.

B. Locally Adaptive Weighting of Gradient Magnitudes

Perceptually motivated features at different scales and orientations can be modeled by Gaussian derivative functions in terms of retino-cortical information [41], and can be represented using gradient magnitudes (GM). The image gradient is a popular feature in IQA [8], [14], [23], [32], [42] since it can effectively capture local image structures, to which the HVS is highly sensitive. At this stage, we compute the gradient magnitude I_{GM} of the image I_{CF} resulting from the previous CSF-based filtering stage. I_{GM} is computed as the root mean square of image directional gradients along two orthogonal directions (see Figure 3 (e) and (f)). I_{GM} measures the strength of local luminance change. Although I_{GM} conveys important visual information, it is not sufficient by itself to estimate the image quality accurately. As shown in the literature [14], [43]–[45], a locally adaptive normalization process is performed by the human visual system, and such a normalization process has been shown to result in a stable statistical image representation [36]. Different normalization methods were adopted in the literature including adaptive gain control [43], divisive normalization models [44], [45], and local energy methods [14].

In the proposed method, our normalization model makes use of I_{CF} and I_{GM} in order to enhance the local image structure while removing the contrast variations in a locally adaptive manner across the whole image. This adaptive weighting is performed based on the local background luminance and the local structure in order to account for the effect of luminance masking and contrast/texture masking and how these influence the visibility of distortions [10], [11]. Toward this end, we compute a locally weighted gradient image that uses the information in I_{CF} and I_{GM} jointly to take into account the effect of local background luminance (given by information in I_{CF}) as well as local image structure (given by information in I_{GM}) on distortion visibility. In our proposed method, the locally weighted gradient image is computed as follows:

$$\hat{I}_{GM}(i,j) = \frac{I_{GM}(i,j)}{\alpha(i,j) + C}, \tag{4}$$

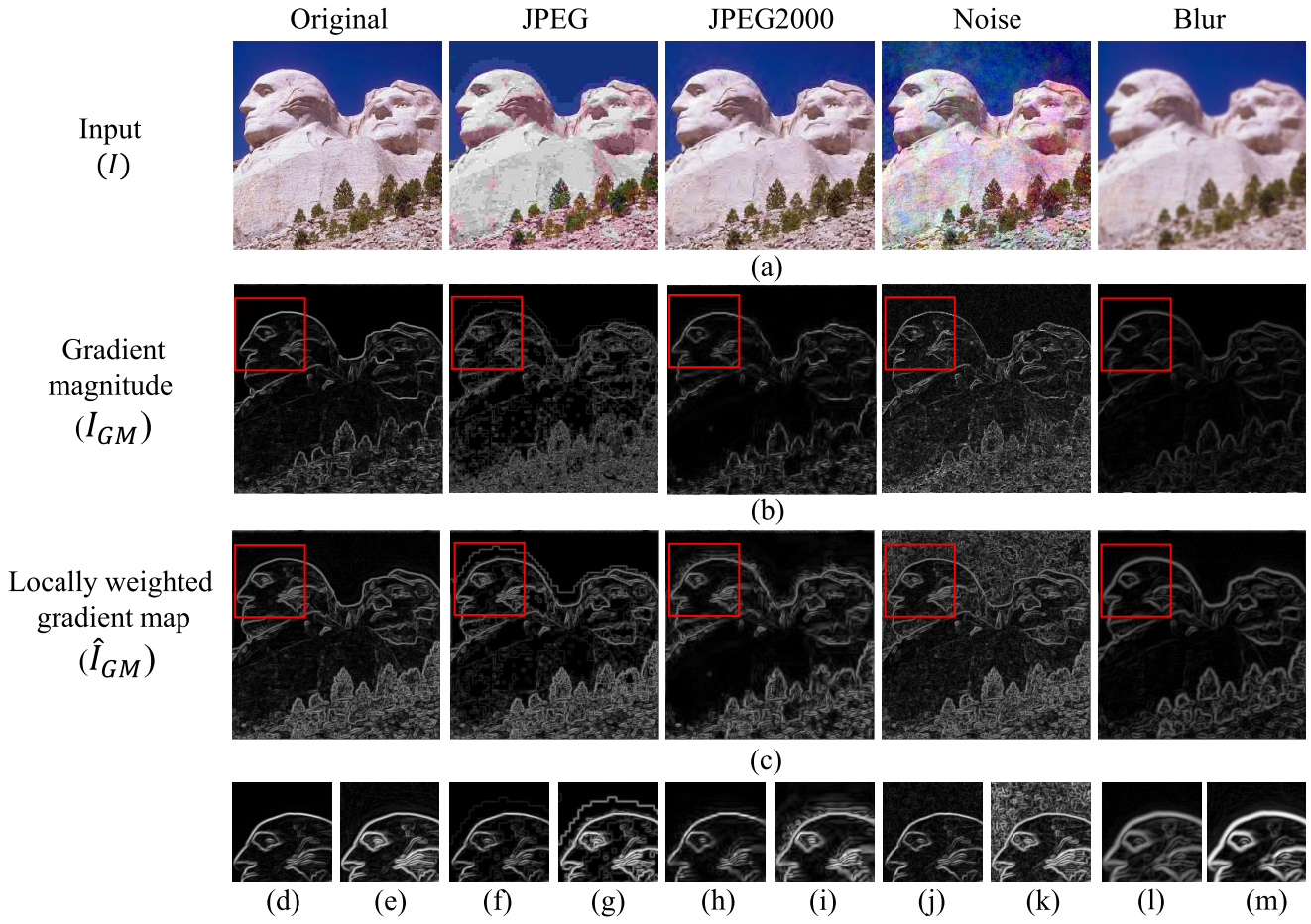

Fig. 4. Comparison between I_{GM} and $\hat{I}_{GM}$ maps. (a) First row represents the input images (I). (b) Second row represents the corresponding gradient magnitude images (I_{GM}) of the first row images. (c) Third row represents the corresponding locally weighted gradient magnitude map images ($\hat{I}_{GM}$) of the first row images. Close-up images of I_{GM} and $\hat{I}_{GM}$ are provided in the forth row, where (d), (f), (h), (j), and (l) belong to the corresponding I_{GM} images in the second row, and (e), (g), (i), (k), and (m) belong to the corresponding $\hat{I}_{GM}$ images in the third row.

where

$$\alpha(i,j) = \sqrt{\sum\sum_{(i',j')\in\Omega_{i,j}} I_{NF}(i',j')w(i',j')}, \tag{5a}$$

$$lI_{NF}(i,j) = \frac{I^2_{GM}(i,j) + I^2_{CF}(i,j)}{2}, \tag{5b}$$

$$w(i,j) = \frac{K(i,j)}{\sum_{i,j} K(i,j)}. \tag{5c}$$

In (5a) $\Omega_{i,j}$ is a local window centered at (i,j), and $w(i',j')$ are positive symmetric weights satisfying $\sum_{i',j'} w(i',j') = 1$ via (5c), where $K(i,j)$ denotes a Gaussian kernel coefficient at (i,j) with window size 8×8 and $\sigma = 0.5$. In (4), a constant is added when $\alpha(i,j)$ has a small value. Examples of locally weighted gradient images are shown in Figures 3 (g) and (h). Figure 4 illustrates the gradient magnitude map without the proposed local weighting fails in effectively capturing some types and levels of distortions. Moreover, from Figure 4, it can be clearly seen that, in contrast to the gradient magnitude map which does not change much under different types and levels of distortions (Figure 4 (b) and close-ups shown in Figures 4 (d), (f), (h), (j), (l)), our proposed locally weighted gradient magnitude map (Figure 4(c) and close-ups shown in Figures 4 (e), (g), (i), (k), (m)) can reflect more clearly the local changes in the image structures caused by different distortion types and levels.

C. DWT Coefficients Computation

Visual images are subjected to local spatial frequency decompositions in the visual cortex [5], [46]. The structural information in natural images may loosely be described as smoothness, texture, and edge information composed by local spatial frequencies that constructively and destructively interfere over scales to produce the spatial structure in natural scenes [12]. Although the proposed locally weighted gradient map captures the local image structure and structural distortions, it is important for reduced-reference quality assessment to efficiently represent the captured structure using as few discriminating features as possible. This can be achieved using a multi-scale representation as different structural components can be more effectively captured and represented at certain scales.

In the literature of neural physiology [44], it has been shown that wavelet transforms provide a convenient framework for localized representation of signals simultaneously in space

and frequency. They have been widely used to model the processing in the early stages of biological visual systems and have also become the preferred form of representations for many image processing and computer vision algorithms [22], [47], [48]. Therefore, in our reduced-reference IQA model, a multiscale DWT decomposition is used to decompose the proposed locally weighted gradient magnitude map into components at different scales (fine to coarse) in order to extract relevant features that are sensitive to various image distortions.

Both the reference and distorted images are decomposed into different subbands using a steerable pyramid wavelet decomposition. $I_{DWT}^{s,o}$ is used to denote the wavelet coefficients of image $\hat{I}_{GM}$ at the o^{th} orientation and the s^{th} scale.

D. Entropy Computation and Pooling

The proposed training-free **R**RIQA method is based on the **E**ntropy of **D**WT coefficients of **Lo**cally weighted **G**radient (*REDLOG*).

Differential entropy is a fundamental measure of the uncertainty associated with a probability distribution [35]. The uncertainty of shape and structure knowledge of an image can be computed by its entropy [49]. Images with varying edge orientations and complex shapes have a relatively high entropy. Alternatively, images with straight edges will produce energy at a single orientation or very peaked distributions so they have a low entropy [49], [50]. Computing the entropy of the wavelet coefficients of $\hat{I}_{GM}$ at each scale and orientation is used in this work to represent efficiently the information about the image structure using as few features as possible for reduced-reference visual quality assessment. The pooled entropies across orientations at each scale are used as RR features and to estimate the visual quality.

Let $H_{s,o}$ denote the computed entropy value of the DWT coefficients, after removing the mean value and converting to unsigned 8-bit integers, for the s^{th} scale and o^{th} orientation ($I_{DWT}^{s,o}$), and let $I_{DWT}^{s,o,k}$, $k = 1, \ldots, M_{s,o}$, denote all the values of $I_{DWT}^{s,o}$. The entropy of $I_{DWT}^{s,o}$ is given by:

$$H_{s,o} = -\sum_{k=1}^{M_{s,o}} p(I_{DWT}^{s,o,k}) \log p(I_{DWT}^{s,o,k}), \tag{6}$$

where p is the probability density function associated with $I_{DWT}^{s,o}$. After computing the entropies for all the scales and orientations, the entropies $H_{s,o}$ are pooled at each scale s, $s = 1, ..., N_s$ along the orientations as follows:

$$H_s = \sum_{o=1}^{N_o} \log(1 + H_{s,o}), \tag{7}$$

where H_s represents the pooled value of entropies over all the orientations of the s^{th} scale.

Let H_s^r and H_s^d represent pooled entropies of the reference and distorted images at the s^{th} scale, respectively. The value of H_s^r at each scale constitutes a RR feature and would be sent through the ancillary channel for the reference image. The total number of RR features is equal to the number of scales N_s. In our implementation, we use a 6-level decomposition, with 4 orientations at each scale, which results in only 6 RR features.

E. REDLOG Index Computation

The proposed reduced-reference quality index, which we denote as *REDLOG*, is computed as follows:

$$REDLOG = \log\left(C \sum_{s=1}^{N_s} (H_s^r - H_s^d)^2 (M^d + 1) + 1\right), \tag{8}$$

where H_s^r and H_s^d denote, respectively, the pooled DWTentropies at the s^{th} scale for the reference and distorted images, N_s denotes the number of decomposition levels (scales), and C is a scaling constant. In our implementation, C is set to 50. In (8), M^d is given by:

$$M^d = \sum_{s=1}^{N_s} \sum_{o=1}^{N_o} M_{s,o}^d, \tag{9}$$

where N_s and N_o denote the number of scales and orientations, respectively, and $M_{s,o}^d$ is the mean absolute value of DWT coefficients at the s^{th} scale and o^{th} orientation for the distorted image. As the quality of an image degrades, the *REDLOG* index value increases.

The proposed *REDLOG* index has a low data rate, as only a small set of RR features are extracted from the reference image (6 features). The amount of information required as RR features can be reduced further by changing the number of scales.

As shown in Section IV, our proposed *REDLOG* index not only works well for natural images, but also works well for texture images with different distortions.

IV. Results

In this section, the performance of the proposed *REDLOG* index is analyzed in terms of its ability to predict subjective ratings of image quality for both natural and texture images. We evaluate the performance in terms of prediction accuracy, prediction monotonicity, and prediction consistency, and provide comparisons with existing state-of-the-art RRIQA methods.

We also evaluate the performance of *REDLOG* for different numbers of scales and orientations and therefore different numbers of RR features. Moreover, to make statistically meaningful conclusions on the models' performance, we further conduct a series of hypothesis tests based on the prediction residuals (errors in predictions) after nonlinear regression.

A. Databases and Performance Measures

To quantify the performance, we applied *REDLOG* to seven different image quality databases including natural image (LIVE, CSIQ, TID2008, TID2013, Toyama, and IVC) and texture (QualTex) databases.

The LIVE database [51] contains 779 distorted images with five types of distortions (JPEG2000 compression, JPEG compression, white noise contamination, Gaussian blur,

TABLE I

PLCC AND SROCC OF DIFFERENT RRIQA ALGORITHMS FOR DIFFERENT TYPES OF DISTORTIONS IN THE LIVE DATABASE. BOLD ENTRIES ARE THE BEST AND SECOND-BEST PERFORMERS

PLCC							
	# Features	JPEG	JPEG2000	AWGN	Gaussian Blur	Fast Fading	Overall
RRED [28]	$342 \leq \frac{Image\ Size}{576}$	**0.9788**	**0.9600**	**0.9845**	0.9318	0.7838	0.9066
DNT marginal [33]	48	0.9329	0.8461	0.8769	0.9265	0.9178	0.8949
RR-SSIM [34]	36	0.9543	0.9435	**0.9772**	0.9154	0.9178	**0.9194**
SPCRM-SCHARR [32]	32	0.9512	0.9293	0.9311	**0.9493**	0.8924	0.8379
Wu *et al.* [27]	30	0.8952	0.9325	0.9574	**0.9554**	**0.9445**	0.7252
WNISM [22]	18	0.9291	0.9261	0.8897	0.8873	0.9230	0.7365
Ma *et al.* [25]	18	0.9311	0.8452	0.8914	0.9304	0.9191	0.8831
βW-SCM [23]	6	0.9225	0.9531	0.9755	0.9454	0.9243	0.8173
REDLOG (proposed)	6	**0.9612**	**0.9582**	0.9204	0.9483	**0.9597**	**0.9372**

SROCC							
	# Features	JPEG	JPEG2000	AWGN	Gaussian Blur	Fast Fading	Overall
RRED [28]	$342 \leq \frac{Image\ Size}{576}$	**0.9725**	**0.9536**	**0.9763**	0.9221	0.7549	0.8964
DNT marginal [33]	48	0.9238	0.8371	0.8619	0.9214	0.9066	0.8882
RR-SSIM [34]	36	**0.9532**	0.9278	0.9642	0.8692	0.9178	0.9129
SPCRM-SCHARR [32]	32	0.9444	0.9273	0.9273	**0.9401**	0.8861	**0.9131**
Wu *et al.* [27]	30	0.8851	0.9503	0.9463	**0.9612**	**0.9413**	0.7325
WNISM [22]	18	0.9542	0.9335	0.8701	0.9145	0.9227	0.7651
Ma *et al.* [25]	18	0.9242	0.8383	0.8805	0.9302	0.9102	0.8792
βW-SCM [23]	6	0.9121	0.9521	**0.9715**	0.9371	0.9258	0.8391
REDLOG (proposed)	6	0.9500	**0.9523**	0.9302	0.9349	**0.9638**	**0.9455**

and fast fading channel distortion of JPEG2000 compressed bitstream) at different distortion levels. The Categorical Image Quality (CSIQ) database [7] contains 866 distorted images of six types of distortions at four and five distortion levels. The distortion types include JPEG compression, JPEG2000 compression, global contrast decrements, white noise, additive pink Gaussian noise, and Gaussian blurring. The Tampere Image database 2008 (TID2008) [52] includes 1700 distorted images with 17 distortion types at four distortion levels. The TID2013 database [53] contains 3000 distorted images, including 25 reference images, 24 types of distortions for each reference image, and five levels of distortion. The IVC database [54] includes 185 distorted images with four types of distortions. The Toyama-MICT database [55] contains 196 images, including 168 distorted images generated by JPEG and JPEG2000 compression. The QualTex database [56] contains 340 distorted images extracted from 10 reference texture images, with six types of distortions (JPEG2000 compression, JPEG compression, white noise contamination, Gaussian blur, subpixel-shift, and synthesis with parameter quantization) for each reference image. These images cover the range of different texture types such as regular, non-regular and stochastic textures as well as varying ability to mask different types of noise and distortions.

For performance evaluation, we employ three commonly used performance metrics. We measure the prediction monotonicity of *REDLOG* via the Spearman rank-order correlation coefficient (SROCC). This metric operates only on the rank of the data points and ignores the relative distance between data points. We also apply a regression analysis to provide a nonlinear mapping between the objective scores and either the subjective mean opinion scores (MOS) or difference of mean opinion scores (DMOS). We measure the Pearson linear correlation coefficient (PLCC) between MOS (DMOS) and the objective scores after nonlinear regression. The root mean squared error (RMSE) between MOS (DMOS) and the objective scores after nonlinear regression is also measured. For the nonlinearity relation measurement, we used the following logistic nonlinearity mapping function as suggested by [4]:

$$f(x) = \frac{\beta_1 - \beta_2}{1 + e^{(\frac{x-\beta_3}{\beta_4})}} + \beta_2, \tag{10}$$

where x denotes the predicted quality for the image and $f(x)$ denotes the quality score after nonlinear fitting, and $\beta_i, i = 1, 2, .., 4,$ are regression model parameters.

B. Performance Comparison

As stated earlier, seven databases were used to evaluate the proposed *REDLOG* index and compare its performance with existing RRIQA methods.

1) Natural Images: Most of the existing RRIQA were designed to predict the quality of natural images. In Table I, we compare the results obtained using the proposed *REDLOG* and various state-of-the-art RRIQA algorithms for different distortion types in the LIVE database. From Table I, it can

TABLE II

COMPARISON OF *REDLOG* VS. VARIOUS IQA ALGORITHMS ON DIFFERENT DATABASES. BOLD ENTRIES ARE THE BEST AND SECOND-BEST PERFORMERS

IQA measure	# Features	LIVE database PLCC	LIVE database SROCC	LIVE database RMSE	CSIQ database PLCC	CSIQ database SROCC	CSIQ database RMSE	IVC database PLCC	IVC database SROCC	IVC database RMSE
PSNR	FR	0.8700	0.8760	13.368	0.7512	0.8058	0.1732	0.6719	0.6884	0.9023
MS-SSIM [57]	FR	**0.9430**	**0.9445**	**9.0956**	**0.8998**	**0.9138**	**0.1145**	**0.8934**	**0.8847**	**0.5474**
RRED [28]	$342 \leq \frac{Image\ Size}{576}$	0.9066	0.8964	11.7944	0.8003	0.8083	0.1574	0.7698	0.8384	1.2184
DNT marginal [33]	48	0.8949	0.8882	11.7862	0.7009	0.7027	0.1872	0.6316	0.6099	0.9446
RR-SSIM [34]	36	0.9194	0.9129	11.3026	0.8426	0.8527	0.1413	0.8177	0.8154	0.7014
SPCRM-SCHARR [32]	32	0.8379	0.9131	27.3086	0.7892	**0.8889**	0.2625	0.7919	0.8186	1.2184
WNISM [22]	18	0.7365	0.7651	18.4814	0.7124	0.7431	0.1842	0.5311	0.4114	1.0322
βW-SCM [23]	6	0.8173	0.8391	15.7351	0.6827	0.6455	0.1918	0.5267	0.4512	1.0357
REDLOG (proposed)	6	**0.9372**	**0.9455**	**9.5224**	**0.8560**	0.8576	**0.1357**	**0.8559**	**0.8544**	**0.6301**

IQA measure	# Features	Toyoma database PLCC	Toyoma database SROCC	Toyoma database RMSE	TID2008 database PLCC	TID2008 database SROCC	TID2008 database RMSE	TID2013 database PLCC	TID2013 database SROCC	TID2013 database RMSE
PSNR	FR	0.6329	0.6131	0.9688	0.5232	0.5530	1.1435	0.7018	0.6394	0.8832
MS-SSIM [57]	FR	**0.8935**	**0.8864**	**0.5621**	**0.8425**	**0.8528**	**0.7299**	**0.8299**	**0.7872**	**0.6917**
RRED [28]	$342 \leq \frac{Image\ Size}{576}$	0.7564	0.8199	1.2514	0.5637	0.6272	1.3419	0.6321	0.6625	1.2397
DNT marginal [33]	48	0.6733	0.6521	0.9253	0.5964	0.5722	1.0772	0.6526	0.6532	0.9831
RR-SSIM [34]	36	0.8051	0.8003	0.7423	0.7231	0.7210	0.9270	0.7324	0.6621	0.8231
SPCRM-SCHARR [32]	32	0.7145	0.8067	1.2514	0.6812	**0.7614**	1.3419	0.7099	0.7190	1.2397
WNISM [22]	18	0.6542	0.6322	0.9464	0.5891	0.5119	1.0843	0.6247	0.5202	0.9680
βW-SCM [23]	6	0.7008	0.7036	0.8927	0.5536	0.5414	1.1176	0.6291	0.6083	0.9637
REDLOG (proposed)	6	**0.8829**	**0.8876**	**0.5877**	**0.7326**	0.6864	**0.9134**	**0.7400**	**0.6829**	**0.8339**

TABLE III

EFFECT OF CHANGING THE NUMBER OF SCALES AND ORIENTATIONS IN *REDLOG* IN TERMS OF PLCC AND SROCC FOR THE LIVE AND TID2013 DATABASES

	LIVE PLCC	LIVE SROCC	TID2013 PLCC	TID2013 SROCC	# Features
$\text{REDLOG}^{1,2}$	0.5633	0.5766	0.5481	0.5358	1
$\text{REDLOG}^{2,2}$	0.6761	0.6841	0.5702	0.5801	2
$\text{REDLOG}^{3,2}$	0.8826	0.8901	0.6799	0.5991	3
$\text{REDLOG}^{4,2}$	0.9008	0.9021	0.7023	0.6602	4
$\text{REDLOG}^{5,2}$	0.9031	0.9081	0.7322	0.6701	5
$\text{REDLOG}^{6,2}$	0.9064	0.9072	0.7324	0.6739	6
$\text{REDLOG}^{1,4}$	0.6929	0.6951	0.5621	0.5141	1
$\text{REDLOG}^{2,4}$	0.7524	0.7522	0.6031	0.5922	2
$\text{REDLOG}^{3,4}$	0.8815	0.8917	0.6917	0.6125	3
$\text{REDLOG}^{4,4}$	0.9263	0.9325	0.7325	0.6521	4
$\text{REDLOG}^{5,4}$	0.9317	0.9391	0.7392	0.6818	5
$\text{REDLOG}^{6,4}$	0.9372	0.9455	0.7400	0.6829	6

TABLE IV

COMPARISON OF *REDLOG* VS. VARIOUS IQA ALGORITHMS ON THE QUALTEX DATABASE. BOLD ENTRIES ARE THE BEST AND SECOND-BEST PERFORMERS

QualTex database				
IQA measure	# Features	PLCC	SROCC	RMSE
PSNR	FR	0.1985	0.5153	1.17356
MS-SSIM [57]	FR	0.7525	0.6972	0.5600
VIF [5]	FR	0.7478	0.6921	0.5726
VSNR [6]	FR	0.6704	0.6252	0.6680
CWSSIM [58]	FR	0.7881	0.7694	0.5232
MAD [7]	FR	0.7539	0.7346	0.7817
STSSIM [59]	FR	0.7601	0.6785	0.7732
RRED [28]	$3072 < \frac{Image\ Size}{64}$	0.6713	0.6363	0.8819
RR-SSIM [34]	36	0.7423	0.7059	0.7922
SPCRM-SCHARR [32]	32	0.6791	0.8387	1.1898
WNISM [22]	18	0.7230	0.7216	0.8220
βW-SCM [23]	6	**0.8516**	**0.8405**	**0.5413**
REDLOG (proposed)	6	**0.9022**	**0.8564**	**0.5131**

be seen that the proposed *REDLOG* index achieves the best or (near) second best performance for JPEG, JPEG2000, and Fast Fading distortions, while requiring the lowest number of RR features. For the AWGN and Gaussian blur distortions, although the proposed *REDLOG* does not yield the best results, it still achieves a consistently high performance with a very small number of RR features.

Table II shows the obtained performance evaluation results of our proposed algorithm on the LIVE, CSIQ, IVC, Toyama, TID2008, and TID2013 databases in comparison with state-of-the-art RRIQA algorithms as well as FR IQA methods such as PSNR and MS-SSIM [57]. As shown in Table II, the proposed *REDLOG* index is not only competitive with training-free RR and FR IQA algorithms, but also has a comparable or even better performance as compared to trained/tuned-based RRIQA methods. The results show that our index (*REDLOG*) yields high correlation with the subjective quality ratings while requiring the smallest number of RR features (only 6) and being independent from the size of the image. In comparison, the performance of RRED [28] and SPCRM-SCHARR [32] depends on the number of features. As the number of features decreases to a smaller number, their performance degrades. Table II provides the results for SPCRM-SCHARR [32] with 32 features and RRED [28] with $\frac{Image\ Size}{576}$ features (between 342 to 682 features depending on the image size in the database). The proposed *REDLOG* achieves the best results in terms of PLCC, SROCC, and RMSE values for the LIVE, CSIQ, TOYAMA, and TID2013 among all the mentioned RRIQAs. For the TID2008 and IVC databases, *REDLOG* achieves the best PLCC and RMSE, and the second best rank for SROCC among all the RRIQAs. In Table II, although

TABLE V

STATISTICAL SIGNIFICANCE TESTS OF THE COMPETING IQA MODELS ON THE (a) LIVE, (b) CSIQ AND (c) TID2013 DATABASES. A VALUE OF '1' INDICATES THAT THE MODEL IN THE ROW IS SIGNIFICANTLY BETTER THAN THE MODEL IN THE COLUMN, A VALUE '0' INDICATES THAT THE MODEL IN THE ROW IS WORSE THAN THE ONE IN THE COLUMN, AND A VALUE OF '-' INDICATES THAT THE TWO MODELS HAVE NO STATISTICAL DIFFERENCE IN PERFORMANCE

LIVE	PSNR	MS-SSIM [58]	WNISM [22]	RR-SSIM [34]	RRED [28]	SPCRM-SCHARR [32]	βW-SCM [23]	REDLOG
PSNR	-	0	0	0	0	0	1	0
MS-SSIM [57]	1	-	0	0	0	1	1	0
WNISM [22]	1	1	-	-	0	0	0	0
RR-SSIM [34]	1	1	-	-	-	0	1	0
RRED [28]	1	1	1	-	-	0	1	1
SPCRM-SCHARR [32]	1	0	1	1	1	-	-	-
βW-SCM [23]	0	0	1	0	0	-	-	0
REDLOG	1	1	1	1	0	-	1	-

(a)

CSIQ	PSNR	MS-SSIM [58]	WNISM [22]	RR-SSIM [34]	RRED [28]	SPCRM-SCHARR [32]	βW-SCM [23]	REDLOG
PSNR	-	0	0	0	0	0	-	0
MS-SSIM [57]	1	-	1	1	0	0	1	0
WNISM [22]	1	0	-	0	0	0	0	0
RR-SSIM [34]	1	0	1	-	-	0	0	0
RRED [28]	1	1	1	-	-	0	0	0
SPCRM-SCHARR [32]	1	1	1	1	1	-	1	-
βW-SCM [23]	-	0	1	1	1	0	-	0
REDLOG	1	1	1	1	1	-	1	-

(b)

TID2013	PSNR	MS-SSIM [58]	WNISM [22]	RR-SSIM [34]	RRED [28]	SPCRM-SCHARR [32]	βW-SCM [23]	REDLOG
PSNR	-	0	0	0	0	0	-	0
MS-SSIM [57]	1	-	1	0	0	1	-	0
WNISM [22]	1	0	-	1	0	0	-	0
RR-SSIM [34]	1	1	0	-	0	1	0	0
RRED [28]	1	1	1	1	-	1	0	0
SPCRM-SCHARR [32]	1	0	1	0	0	-	1	1
βW-SCM [23]	-	-	-	1	1	0	-	0
REDLOG	1	1	1	1	1	0	1	-

(c)

SPCRM-SCHARR [32] yields a prediction performance that is comparable to *REDLOG*, SPCRM-SCHAR [32] requires a significantly higher number of RR features (32 RR features) as compared to the proposed *REDLOG* index (only 6 RR features). Moreover, the proposed *REDLOG* index performs nearly as well as the popular MS-SSIM [57].

Comparing our proposed *REDLOG* index to the RRIQA methods [23], [32] that make use of the image gradient magnitude, it can be seen from Table II that, although the method of [23] needs the same small number of features as our algorithm, the method of [23] results in a significant drop in performance and does not perform consistently well across different datasets. In addition, the proposed *REDLOG* index not only results in a better performance as compared to the method of [32] but it also requires a significantly lower number of RR features.

In Table III we evaluate our algorithm in terms of the number of features that need to be sent, where $REDLOG^{s,o}$ denotes evaluating *REDLOG* with s scales, resulting in s RR features, and o orientations. We apply *REDLOG* on the LIVE and TID2013 databases and its performance is evaluated by varying the number of scales and orientations. As shown in Table III, for a fixed number of orientations, increasing the number of scales results on average in an increase in the performance of the proposed *REDLOG*. Moreover, decreasing the number of orientations from 4 to 2, results in a significant drop in accuracy of the estimated quality.

2) Texture Images: Most of the IQA algorithms were developed for natural images and they do not typically perform well on texture images. Here we show that our proposed *REDLOG* index not only performs well on natural images, but it also works well for estimating the quality of texture images. The proposed locally weighted gradient map is able to capture the perceived structural distortions while reducing the effect of those that are masked by the original image/visual texture structure. Furthermore, the proposed multi-scale feature representation can efficiently capture changes in the image/visual texture structure using very few parameters. Table IV shows the performance of our index together with existing FR and RRIQA methods on the QualTex database. As shown in Table IV, *REDLOG* exhibits the best performance among all existing FR and RRIQA methods.

C. Statistical Analysis

To make statistically meaningful conclusions on the models' performance, we further conducted a series of hypothesis tests based on the prediction residuals (errors in predictions) of each model after nonlinear regression. Tables V (a)-(c) present significance test results on the LIVE, CSIQ and TID2013 databases, respectively. The Jarque-Bera (JB) statistic [60] was used to verify that the methods' prediction residuals follow the Gaussian distribution. The left-tailed F-test is applied to the residuals of every two methods to be compared; in other words, the left-tail F-test is used to determine whether

the population variance of the first input (indicated by the row) is less than that of the second input (indicated by the column). If the residuals are not Gaussian, the test for significance is considerably more difficult and often inconclusive. A value of H=1 indicates that the F-test rejects the null hypothesis at the default 5% significance level, which means that the first model (indicated by the row) has better IQA performance than the second model (indicated by the column) with a confidence greater than 95%. A value of H=0 means that the first model is not significantly better than the second one, in which case thesecond model can be better than or similar to the first model. If H=0 is obtained no matter which one of the two models is taken as the first one, then the two models have no significant difference in performance, and the symbol '-' is shown in Table V to indicate that the two models are equivalent in performance; otherwise, a value of 0 is shown in Table V indicating that the first model is worse than the second model. As shown in Tables II and V, *REDLOG* performs the best on the considered databases interms of prediction accuracy and the number of needed RR features. It should be noted that, while SPCRM-SCHARR [32] is shown in Table V to result in a prediction performance that is comparable to *REDLOG*, SPCRM-SCHAR [32] requires a significantly higher number of RR features (32 RR features) as compared to the proposed *REDLOG* index (only 6 RR features).

V. Conclusion

Most of the existing RRIQA methods need to either train or tune their algorithms to assess the quality of the images accurately. There are just a few training-free RRIQA algorithms that perform well without training and tuning. However, these algorithms need to send a relatively large amount of information as RR features. Finding a balance between the number of RR features and the predicted image quality is at the core of the design of RRIQA methods. In this paper, we proposed a RR-training-free-IQA method. Our proposed method, *REDLOG*, not only needs a very small number of RR features (6 RR features) independent from the size of the image, but also results in a high quality prediction accuracy over a variety of image databases, including both natural and texture images.

The MATLAB source code of *REDLOG* and the evaluation results are publicly available online at ivulab.asu.edu.

References

[1] Z. Wang and A. C. Bovik, "Modern image quality assessment," in *Synthesis Lectures on Image, Video, and Multimedia Processing*, vol. 2. San Rafael, CA, USA: Morgan & Claypool, Jan. 2006, no. 1, pp. 1–156.

[2] 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.*, vol. 13, no. 4, pp. 600–612, Apr. 2004.

[3] 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.*, vol. 9, no. 4, pp. 636–650, Apr. 2000.

[4] 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.*, vol. 15, no. 11, pp. 3440–3451, Nov. 2006.

[5] H. R. Sheikh and A. C. Bovik, "Image information and visual quality," *IEEE Trans. Image Process.*, vol. 15, no. 2, pp. 430–444, Feb. 2006.

[6] D. M. Chandler and S. S. Hemami, "VSNR: A wavelet-based visual signal-to-noise ratio for natural images," *IEEE Trans. Image Process.*, vol. 16, no. 9, pp. 2284–2298, Sep. 2007.

[7] E. C. Larson and D. M. Chandler, "Most apparent distortion: Full-reference image quality assessment and the role of strategy," *J. Electron. Imag.*, vol. 19, no. 1, pp. 6–27, 2010.

[8] L. Zhang, L. Zhang, X. Mou, and D. Zhang, "FSIM: A feature similarity index for image quality assessment," *IEEE Trans. Image Process.*, vol. 20, no. 8, pp. 2378–2386, Aug. 2011.

[9] H. R. Sheikh, A. C. Bovik, and L. Cormack, "No-reference quality assessment using natural scene statistics: JPEG2000," *IEEE Trans. Image Process.*, vol. 14, no. 1, pp. 1918–1927, Nov. 2005.

[10] R. Ferzli and L. J. Karam, "A no-reference objective image sharpness metric based on the notion of just noticeable blur (JNB)," *IEEE Trans. Image Process.*, vol. 18, no. 4, pp. 717–728, Apr. 2009.

[11] T. Zhu and L. Karam, "A no-reference objective image quality metric based on perceptually weighted local noise," *EURASIP J. Image Video Process.*, vol. 2014, no. 1, pp. 1–8, Dec. 2014.

[12] M. A. Saad, A. C. Bovik, and C. Charrier, "Blind image quality assessment: A natural scene statistics approach in the DCT domain," *IEEE Trans. Image Process.*, vol. 21, no. 8, pp. 3339–3352, Aug. 2012.

[13] Y. Zhang, A. K. Moorthy, D. M. Chandler, and A. C. Bovik, "C-DIIVINE: No-reference image quality assessment based on local magnitude and phase statistics of natural scenes," *Signal Process., Image Commun.*, vol. 29, no. 7, pp. 725–747, Aug. 2014.

[14] W. Xue, X. Mou, L. Zhang, A. C. Bovik, and X. Feng, "Blind image quality assessment using joint statistics of gradient magnitude and Laplacian features," *IEEE Trans. Image Process.*, vol. 23, no. 11, pp. 4850–4862, Nov. 2014.

[15] D. M. Chandler, M. M. Alam, and T. D. Phan, "Seven challenges for image quality research," *Proc. SPIE*, vol. 9014,

[16] U. Engelke, M. Kusuma, H.-J. Zepernick, and M. Caldera, "Reduced-reference metric design for objective perceptual quality assessment in wireless imaging," *Signal Process., Image Commun.*, vol. 24, no. 7, pp. 525–547, Aug. 2009.

[17] T. M. Kusuma and H.-J. Zepernick, "A reduced-reference perceptual quality metric for in-service image quality assessment," in *Proc. Joint 1st Workshop Mobile Future Symp. Trends Commun.*, vol. 11. Oct. 2003, pp. 71–74.

[18] I. P. Gunawan and M. Ghanbari, "Reduced-reference picture quality estimation by using local harmonic amplitude information," in *Proc. London Commun. Symp.*, 2003, pp. 353–358.

[19] K. Chono, Y.-C. Lin, D. Varodayan, Y. Miyamoto, and B. Girod, "Reduced-reference image quality assessment using distributed source coding," in *Proc. IEEE Int. Conf. Multimedia Expo*, Apr./Jun. 2008, pp. 609–612.

[20] M. Carnec, P. Le Callet, and D. Barba, "An image quality assessment method based on perception of structural information," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, vol. 2. Sep. 2003, pp. III-185–III-188.

[21] M. Carnec, P. Le Callet, and D. Barba, "Objective quality assessment of color images based on a generic perceptual reduced reference," *Signal Process., Image Commun.*, vol. 23, no. 4, pp. 239–256, Apr. 2008.

[22] Z. Wang, G. Wu, H. R. Sheikh, E. P. Simoncelli, E.-H. Yang, and A. C. Bovik, "Quality-aware images," *IEEE Trans. Image Process.*, vol. 15, no. 6, pp. 1680–1689, Jun. 2006.

[23] W. Xue and X. Mou, "Reduced reference image quality assessment based on Weibull statistics," in *Proc. 2nd Int. Workshop Quality Multimedia Exper. (QoMEX)*, Jun. 2010, pp. 1–6.

[24] M. Zhang, W. Xue, and X. Mou, "Reduced reference image quality assessment based on statistics of edge," *Proc. SPIE*, vol. 7876, pp. 787611-1–787611-7, Jan. 2011.

[25] L. Ma, S. Li, F. Zhang, and K. N. Ngan, "Reduced-reference image quality assessment using reorganized DCT-based image representation," *IEEE Trans. Multimedia*, vol. 13, no. 4, pp. 824–829, Aug. 2011.

[26] L. Ma, S. Li, and K. N. Ngan, "Reduced-reference image quality assessment in reorganized DCT domain," *Signal Process., Image Commun.*, vol. 28, no. 8, pp. 884–902, Sep. 2013.

[27] J. Wu, W. Lin, G. Shi, and A. Liu, "Reduced-reference image quality assessment with visual information fidelity," *IEEE Trans. Multimedia*, vol. 15, no. 7, pp. 1700–1705, Nov. 2013.

[28] R. Soundararajan and A. C. Bovik, "RRED indices: Reduced reference entropic differencing for image quality assessment," *IEEE Trans. Image Process.*, vol. 21, no. 2, pp. 517–526, Feb. 2012.

[29] X. Gao, W. Lu, D. Tao, and X. Li, "Image quality assessment based on multiscale geometric analysis," *IEEE Trans. Image Process.*, vol. 18, no. 7, pp. 1409–1423, Jul. 2009.

[30] D. Tao, X. Li, W. Lu, and X. Gao, "Reduced-reference IQA in contourlet domain," *IEEE Trans. Syst., Man, Cybern. B, Cybern.*, vol. 39, no. 6, pp. 1623–1627, Dec. 2009.
[31] L. Zhichao, T. Jinxu, and Z. Zhufeng, "Reduced-reference image quality assessment based on average directional information," in *Proc. IEEE Int. Conf. Signal Process. (ICSP)*, vol. 2. Oct. 2012, pp. 787–791.
[32] D. Liu, Y. Xu, Y. Quan, and P. Le Callet, "Reduced reference image quality assessment using regularity of phase congruency," *Signal Process., Image Commun.*, vol. 29, no. 8, pp. 844–855, Sep. 2014.
[33] Q. Li and Z. Wang, "Reduced-reference image quality assessment using divisive normalization-based image representation," *IEEE J. Sel. Topics Signal Process.*, vol. 3, no. 2, pp. 202–211, Apr. 2009.
[34] A. Rehman and Z. Wang, "Reduced-reference image quality assessment by structural similarity estimation," *IEEE Trans. Image Process.*, vol. 21, no. 8, pp. 3378–3389, Aug. 2012.
[35] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. New York, NY, USA: Wiley, 2012.
[36] S. Lyu and E. P. Simoncelli, "Statistically and perceptually motivated nonlinear image representation," *Proc. SPIE*, vol. 6492, pp. 649207-1–649207-15, Feb. 2007.
[37] J. Mannos and D. J. Sakrison, "The effects of a visual fidelity criterion of the encoding of images," *IEEE Trans. Inf. Theory*, vol. 20, no. 4, pp. 525–536, Jul. 1974.
[38] S. Daly, "Subroutine for the generation of a two dimensional human visual contrast sensitivity function," Eastman Kodak, Rochester, NY, USA, Tech. Rep. 233203Y, 1987.
[39] T. Mitsa and K. L. Varkur, "Evaluation of contrast sensitivity functions for the formulation of quality measures incorporated in halftoning algorithms," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, vol. 5. Apr. 1993, pp. 301–304.
[40] S. J. Daly, "Visible differences predictor: An algorithm for the assessment of image fidelity," *Proc. SPIE*, vol. 1666, pp. 2–15, Aug. 1992.
[41] A. J. Bell and T. J. Sejnowski, "The 'independent components' of natural scenes are edge filters," *Vis. Res.*, vol. 37, no. 23, pp. 3327–3338, Dec. 1997.
[42] A. Liu, W. Lin, and M. Narwaria, "Image quality assessment based on gradient similarity," *IEEE Trans. Image Process.*, vol. 21, no. 4, pp. 1500–1512, Apr. 2012.
[43] D. J. Heeger, "Normalization of cell responses in cat striate cortex," *J. Neurosci.*, vol. 9, no. 2, pp. 181–197, 1992.
[44] E. P. Simoncelli and B. A. Olshausen, "Natural image statistics and neural representation," *Annu. Rev. Neurosci.*, vol. 24, no. 1, pp. 1193–1216, 2001.
[45] S. Lyu and E. P. Simoncelli, "Nonlinear image representation using divisive normalization," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2008, pp. 1–8.
[46] R. Blake and R. Sekuler, *Perception*. New York, NY, USA: McGraw-Hill, 2006.
[47] J. Portilla and E. P. Simoncelli, "A parametric texture model based on joint statistics of complex wavelet coefficients," *Int. J. Comput. Vis.*, vol. 40, no. 1, pp. 49–70, Oct. 2000.
[48] E. P. Simoncelli and D. J. Heeger, "A model of neuronal responses in visual area MT," *Vis. Res.*, vol. 38, no. 5, pp. 743–761, 1998.
[49] L. W. Renninger, P. Verghese, and J. Coughlan, "Where to look next? Eye movements reduce local uncertainty," *J. Vis.*, vol. 7, no. 3, p. 6, 2007.
[50] B. A. Olshausen and D. J. Field, "Natural image statistics and efficient coding," *Network, Comput. Neural Syst.*, vol. 7, no. 2, pp. 333–339, 1996.
[51] H. R. Sheikh, Z. Wang, L. Cormack, and A. C. Bovik. (2005). *LIVE Image Quality Assessment Database Release 2*. [Online]. Available: http://live.ece.utexas.edu/research/quality/subjective.htm
[52] 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," *Adv. Modern Radioelectron.*, vol. 10, no. 4, pp. 30–45, 2009.
[53] N. Ponomarenko *et al.*, "Color image database TID2013: Peculiarities and preliminary results," in *Proc. 4th Eur. Workshop Vis. Inf. Process. (EUVIP)*, Jun. 2013, pp. 106–111.
[54] P. Le Callet and F. Autrusseau. (2005). *Subjective Quality Assessment IRCCyN/IVC Database*. [Online]. Available: http://www.irccyn.ec-nantes.fr/ivcdb/
[55] Y. Horita, K. Shibata, Y. Kawayoke, and Z. P. Sazzad. (2011). *Toyama-MICT Image Quality Evaluation Database*. [Online]. Available: http://mict.eng.u-toyama.ac.jp/mictdb.html
[56] M. S. Gide and L. J. Karam, "On the assessment of the quality of textures in visual media," in *Proc. 44th Annu. Conf. Inf. Sci. Syst. (CISS)*, Mar. 2010, pp. 1–5.
[57] Z. Wang, E. P. Simoncelli, and A. C. Bovik, "Multiscale structural similarity for image quality assessment," in *Proc. 37th Asilomar Conf. Signals, Syst. Comput.*, Nov. 2003, pp. 1398–1402.
[58] X. Zhao, M. G. Reyes, T. N. Pappas, and D. L. Neuhoff, "Structural texture similarity metrics for retrieval applications," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Oct. 2008, pp. 1196–1199.
[59] J. Zujovic, T. N. Pappas, and D. L. Neuhoff, "Structural texture similarity metrics for image analysis and retrieval," *IEEE Trans. Image Process.*, vol. 22, no. 7, pp. 2545–2558, Jul. 2013.
[60] C. M. Jarque and A. K. Bera, "Efficient tests for normality, homoscedasticity and serial independence of regression residuals," *Econ. Lett.*, vol. 6, no. 3, pp. 255–259, 1980.

S. Alireza Golestaneh received the B.Sc. degree in electrical engineering from Shiraz University, Shiraz, Iran, in 2010, and the M.Sc. degree in electrical and computer engineering from the Computational Perception and Image Quality Laboratory, Oklahoma State University, Stillwater, OK, USA, in 2013.

He is currently pursuing the Ph.D. degree in electrical engineering with Arizona State University. His current research interests include image and video quality assessment, statistical modeling of images, and perceptual based image and video processing, and visual perception and computer vision. He was a recipient of the Graduate Fellowship from Arizona State University from 2014 to 2015.

He is also a Teaching Assistant with the School of Electrical, Computer and Energy Engineering. He is a member of the Image, Video, and Usability Laboratory with Arizona State University.

Lina J. Karam (F'90) received the B.E. degree in computer and communications engineering from the American University of Beirut, Beirut, Lebanon, in 1989, and the M.S. and Ph.D. degrees in electrical engineering from the Georgia Institute of Technology, Atlanta, GA, USA, in 1992 and 1995, respectively.

She is currently a Full Professor with the School of Electrical, Computer and Energy Engineering, Arizona State University, Tempe, AZ, USA, where she directs the Image, Video, and Usability Research Laboratory. Her industrial experience includes image and video compression development with AT&T Bell Labs, Murray Hill, NJ, USA, multidimensional data processing and visualization with Schlumberger, biomedical image processing with Muscale, and collaboration on computer vision, image/video processing, compression, and transmission projects with industries including Intel, Google, Qualcomm, NTT, Motorola, General Dynamics, and NASA. She has authored over 100 technical publications. She is a co-inventor of a number of patents.

Dr. Karama is a member of the IEEE Signal Processing Magazine Senior Editorial Board, the IEEE Signal Processing Society's Image, Video & Multidimensional Signal Processing Technical Committee, and the IEEE Circuits and Systems Society's DSP Technical Committee. She is a member of the Signal Processing, Circuits and Systems, and Communications societies of the IEEE. She received the U.S. National Science Foundation CAREER Award, the NASA Technical Innovation Award, the 2012 Intel Outstanding Researcher Award, the 2012 IEEE Phoenix Section Outstanding Faculty Award, and the 2014 IEEE SPS Best Paper Award. She has served on several journal editorial boards, several conference organization committees, and several IEEE technical committees. She served as the Lead Guest Editor of the Proceedings of the IEEE, Special Issue on Perceptual-Based Media Processing. She served as the Technical Program Chair of the 2009 IEEE International Conference on Image Processing, as the General Chair of the 2011 IEEE International DSP/SPE Workshops, and as the Lead Guest Editor of the IEEE JOURNAL ON SELECTED TOPICS IN SIGNAL PROCESSING, Special Issue on Visual Quality Assessment. She has co-founded two international workshops VPQM and QoMEX. She is currently serving as the General Chair of the 2016 IEEE International Conference on Image Processing. She is also serving on the IEEE SPS Board of Governors and the IEEE Publications Board Strategic Planning Committee.