Reduced-Reference Image Quality Assessment by Structural Similarity Estimation

Abdul Rehman, *Student Member, IEEE*, and Zhou Wang, *Member, IEEE*

Abstract—Reduced-reference image quality assessment (RR-IQA) provides a practical solution for automatic image quality evaluations in various applications where only partial information about the original reference image is accessible. In this paper, we propose an RR-IQA method by estimating the structural similarity index (SSIM), which is a widely used full-reference (FR) image quality measure shown to be a good indicator of perceptual image quality. Specifically, we extract statistical features from a multiscale multiorientation divisive normalization transform and develop a distortion measure by following the philosophy in the construction of SSIM. We find an interesting linear relationship between the FR SSIM measure and our RR estimate when the image distortion type is fixed. A regression-by-discretization method is then applied to normalize our measure across image distortion types. We use six publicly available subject-rated databases to test the proposed RR-SSIM method, which shows strong correlations with both SSIM and subjective quality evaluations. Finally, we introduce the novel idea of partially repairing an image using RR features and use deblurring as an example to demonstrate its application.

Index Terms—Divisive normalization transform, image deblurring, image repairing, natural image statistics, reduced-reference image quality assessment (RR-IQA), structural similarity.

I. INTRODUCTION

OVER the past years, there has been an exponential increase in the demand for image and video services. Nevertheless, the networks in service are not designed to accommodate the current trends of traffic. In practice, the multimedia content delivered over the networks suffers from various kinds of distortions on its way to the destination. It is important for the service providers to be able to identify and quantify the quality degradations in order to maintain the required quality of service. This gives rise to the desire of accurate and efficient perceptual image quality assessment (IQA) algorithms that can estimate the subjective quality of the image content under various kinds of distortions.

Much work has been done in the recent past to develop objective quality assessment measures which can automatically measure the perceived distortion in the visual content.

The most prominent ones include the structure similarity index (SSIM) [1] and its derivatives [2], [3], visual information fidelity [4], visual signal-to-noise ratio [5], and the most apparent distortion [6]. Among these methods, SSIM has often been preferred because of its good tradeoff between accuracy, simplicity, and efficiency [7]. $\sqrt{1 - \text{SSIM}}$ has been shown to be a valid distance metric (that satisfies the identity and symmetry axioms as well as the triangle inequality) and has a number of useful local and quasi-convexity and distance-preserving properties [8]. Besides IQA, SSIM has also found a wide variety of applications, ranging from image coding, restoration, and fusion to watermarking and biometrics [9]–[14]. The success of SSIM motivated us to use it for visual communication applications. The difficulty is that SSIM is a full-reference IQA (FR-IQA) scheme that requires full availability of the reference image in order to estimate the quality of the distorted image. This makes it impractical in visual communication applications, where we have no access to the reference image at the receiver side. No-reference IQA (NR-IQA) is highly desirable because it does not require access to the reference image. In the literature, most NR-IQA algorithms were designed for specific and limited types of distortions [15]–[21]. They may not be good choices in modern communication networks, where the distortions could be a combination of lossy compression, scaling in bit rate and spatial/temporal resolution, network delay and packet loss, and various types of pre- and postprocessing filtering (e.g., error concealment, deblocking filtering, sharpening). On the other hand, general-purpose NR-IQA is still at an immature stage.

The reduced-reference IQA (RR-IQA) method only requires a limited number features extracted from the reference for the IQA task [22]. It provides an interesting compromise between FR and NR approaches in terms of both quality prediction accuracy and the amount of information required to describe the reference. A general framework for the use of RR-IQA in visual communications along with image-repairing capability is shown in Fig. 1. An image x is transmitted to the receiver via a transmission channel, which introduces distortions in the received image y. Meanwhile, RR features X extracted at the transmitter side are sent to the receiver through an ancillary channel. The feature extraction unit at the receiver side calculates the features Y from the received image y in a similar fashion as in the transmitter side. X and Y are compared at the quality assessment unit, which creates a quality score S of the distorted image y. A good RR-IQA approach should achieve a good tradeoff between the rate and accuracy. In general, the larger the rate of the RR features, the more
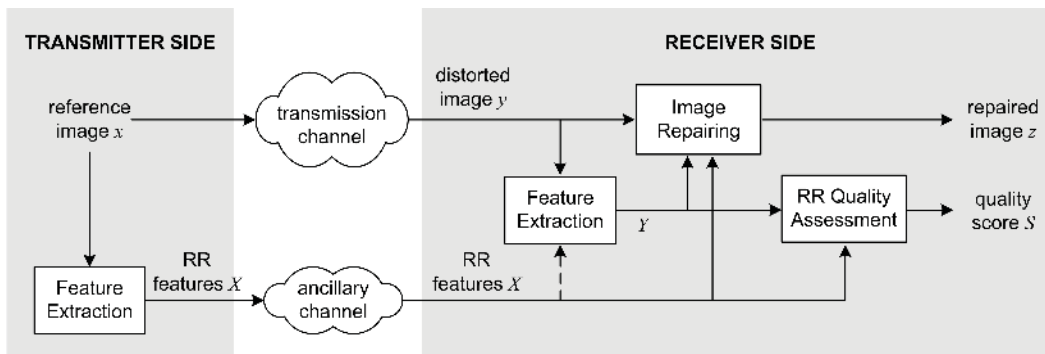
Fig. 1. General framework for the deployment of RR-IQA systems with image repairing capability.

accurate the RR-IQA measure can achieve. In the extreme, when the rate is enough to fully reconstruct the reference, RR-IQA converges to FR-IQA. The performance gap between RR-IQA and FR-IQA may be reduced by selecting RR features that are efficient, perceptually relevant, and sensitive to various kinds of distortions. In addition, since the RR features provide information about what the "correct" image is supposed to look like, they may also be used as side information to repair the received distorted image, as illustrated in Fig. 1.

Based on the underlying design philosophy, existing RR-IQA algorithms may be loosely classified into three categories. The first type of methods are primarily built upon models of the image *source*. Since the reference image is not available in the deterministic sense, these models are often statistical that capture *a priori* the low-level statistical properties of natural images. The model parameters provide a highly efficient way to summarize the image information, and thus these methods often lead to RR-IQA algorithms with low RR data rate. In [23] and [24], the marginal distribution of wavelet subband coefficients is modeled using a generalized Gaussian density (GDD) function, and GGD model parameters are used as RR features are employed to quantify the variations of marginal distributions in the distorted image. The model was further improved in [25] by employing a nonlinear divisive normalization transform (DNT) after the linear wavelet decomposition, which resulted in enhanced quality prediction performance, especially when images with different distortion types are mixed together. The second category of RR-IQA methods are oriented to capture image *distortions*. These methods provide useful and straightforward solutions when we have sufficient knowledge about the distortion process that the images underwent, e.g., standard image or video compression [26]–[29]. The limitation of such approaches is in their generalization capability. Generally, it is inappropriate to apply these methods beyond the distortions they are designed to capture. The third category of RR-IQA algorithms are based on models of the image *receiver* [i.e., the hierarchical visualisation system (HVS)] [30], [31], where computational models from physiological and/or psychophysical vision studies may be employed. These methods have demonstrated good performance for JPEG and JPEG2000 compression [30], [31]. Among the three classes of RR-IQA approaches, the first and third ones, i.e., methods

based on modeling the image source and the receiver, have more potential to be extended for general-purpose applications because the statistical and perceptual features being employed are not restricted to any specific distortion process. There are also interesting conceptual connections between these two types of approaches, because it is a general belief in biological vision science that the HVS is highly tuned for efficient statistical encoding of the natural visual environment [32], [33].

This paper focuses on a general-purpose RR-IQA based on *natural image statistics* modeling. In addition, motivated by the success of the FR SSIM index, we develop our method as an attempt to estimate SSIM rather than directly predicting subjective quality. The benefits of this approach are twofold. First, the successful design principle in the construction of SSIM can be naturally incorporated into the development of the RR algorithm. Second, when the algorithm design involves a supervised learning stage, it is much easier to obtain training data, because SSIM can be readily computed, as opposed to the expensive and time-consuming subjective evaluations. In [34], an interesting RR video quality measure based on SSIM estimation was proposed for quantifying visual degradations caused by channel transmission errors. It is based on local spatial statistical features and uses distributed source coding techniques to reduce the required bandwidth to transmit RR features. Our method differs from this approach in three ways. First, our method is based on natural image statistical modeling and makes use of the perceptually and statistically motivated DNT transform. Second, instead of decomposing the problem of SSIM estimation into many local problems and estimating each component in SSIM expression separately, our method uses global statistics to estimate global SSIM value. This allows for a much more efficient description of the image content, and thus significantly lowers the number of RR features. Third, our approach aims for a general-purpose RR-IQA that can be applied to assess images with a wide variety of distortion types.

The value of RR-IQA measures is beyond quality evaluations. As illustrated in Fig. 1, they may also be employed to partially "repair" the distorted image. In this paper, we attempt to repair an image by matching the subband statistical properties of the distorted image with those of the reference, and use deblurring as an example to demonstrate the idea.

The interesting feature of this method is that it requires no knowledge about the blur kernel. Instead, the same repairing procedure is successful in correcting images of not only homogeneous blur (e.g., out-of-focus blur) but also directional blur (e.g., motion blur).

II. RR-SSIM ESTIMATION

The proposed RR-SSIM estimation algorithm starts with a feature extraction process of the reference image based on a multiscale multiorientation DNT. Divisive normalization was found to be an effective mechanism to account for many neuronal behaviors in biological perceptual systems [35]–[37]. It also provides a useful model to describe the psychophysical visual masking effect [38], [39]. DNT is typically applied after a multiscale linear transform (loosely referred to as wavelet transform) that decomposes the image into transform coefficients representing localized structures in space, frequency (scale), and orientation. The DNT-domain representation of the image is then calculated by dividing each coefficient by a local energy measure based on its neighboring coefficients. It was found that the histogram of DNT coefficients within a wavelet subband can often be well fitted with a zero-mean Gaussian density function [40], [41], which is a one-parameter function that allows efficient summarization of the statistics of the reference image.

In [25], the effect of image distortions on the statistics of DNT coefficients was studied. It was found that different types of distortions modify the statistics of the reference image in different ways, and the levels of statistical differences may be used to quantify image distortions. In order to estimate FR SSIM, we desire the variations of the statistics of the DNT coefficients with respect to different types and levels of distortions to be coherent with the corresponding effects on FR SSIM.

The Gaussian scale mixture (GSM) model provides a convenient framework to define a DNT [40]. A vector Y of length N is regarded as a GSM if it can be represented as the product of two independent components: i.e., $Y \doteq zU$, where z is a scalar random variable called the mixing multiplier, and U is a zero-mean Gaussian-distributed random vector with covariance C_U. In image processing applications, GSM may be used to model a cluster of wavelet coefficients that are neighbors in space, scale, and orientation. If we assume that z takes a fixed value for each cluster but varies across the image, then putting all z values together constitutes a variance field. DNT can then be accomplished by $v = Y/z$, which produces a random vector that is Gaussian. This had been observed in empirical studies in [40], where z is replaced by a local estimation $\hat{z}$ using a maximum-likelihood estimator [40]

$$\hat{z} = \arg\max_z \{\log p(Y|z)\} = \sqrt{Y^T C_U^{-1} \frac{Y}{N}}. \tag{1}$$

The Gaussianization produced by the DNT process largely reduces the complication in describing the distribution of the subband coefficient x

$$p_m(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{x^2}{2\sigma^2}\right) \tag{2}$$

where only a single parameter σ needs to be recorded for each subband.

In addition to σ, the Kullback–Leibler divergence (KLD) [42] between model Gaussian distribution, $p_m(x)$, and the true probability distribution of the DNT-domain coefficients, $p(x)$, denoted by $d(p_m||p)$ is extracted as the second feature for each subband

$$d(p_m||p) = \int p_m(x) \log \frac{p_m(x)}{p(x)} dx. \tag{3}$$

This improves model accuracy when the probability distribution is not exactly Gaussian.

The subband distortion of the distorted image can be evaluated by the KLD between the probability distribution of the original image, $p(x)$, and that of the distorted image, $q(x)$

$$d(p||q) = \int p(x) \log \frac{p(x)}{q(x)} dx. \tag{4}$$

Direct computation of this quantity requires full access to $p(x)$, which would require a large number of RR features to be described. Fortunately, the Gaussian model of the DNT coefficients (2) provides a good approximation. Therefore, we can estimate $p(x)$ by

$$\hat{d}(p||q) = \int p_m(x) \log \frac{p(x)}{q(x)} dx \tag{5}$$

$$= d(p_m||q) - d(p_m||p) \tag{6}$$

where $d(p_m||q)$ is the KLD between the model Gaussian distribution and the distribution computed from the distorted image. Although different types of distortions affect the statistics of the reference image in different manners, they are all summarized in (6) to a single distortion measure. An added nice feature of this measure is that it equals zero when the two distributions $p(x)$ and $q(x)$ are identical.

At the receiver side, the KLD between the subband coefficient probability distributions of the original and distorted images is calculated as in (6). By assuming independence between subbands, the subband-level distortion measure of (6) can be combined to provide an overall distortion assessment of the whole image by

$$D = \log\left(1 + \frac{1}{D_0} \sum_{k=1}^{K} \left|\hat{d}^k(p^k||q^k)\right|\right) \tag{7}$$

where K is the total number of subbands, p^k and q^k are the probability distributions of the kth subband of the reference and distorted images, respectively, $\hat{d}^k$ represents the KLD between p^k and q^k, and D_0 is a constant to control the scale of the distortion measure.

The limitation of the measure in (7) is that it does not take into account the relationship (or structures) between the distortions across different subbands. Such distortion structure is a critical issue behind the philosophy of the SSIM approach [43], which attempts to distinguish structural and nonstructural distortions. To understand this better, let us look at the FR SSIM algorithm [1], which is based on measuring the similarities of luminance, contrast, and structure between
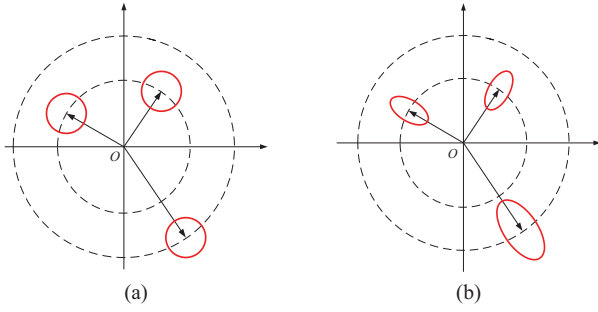
Fig. 2. Equal-distortion contours with respect to the central reference vectors. (a) MSE measure. (b) SSIM measure.

Fig. 3. Relationship between D_n and SSIM for blur, JPEG compression, JPEG2000 compression, and noise contamination distortions.

local image patches $\mathbf{x}$ and $\mathbf{y}$ extracted from a reference and a distorted images

$$l(\mathbf{x}, \mathbf{y}) = \frac{2\mu_x \mu_y + C_1}{\mu_x^2 + \mu_y^2 + C_1} \tag{8}$$

$$c(\mathbf{x}, \mathbf{y}) = \frac{2\sigma_x \sigma_y + C_2}{\sigma_x^2 + \sigma_y^2 + C_2} \tag{9}$$

$$s(\mathbf{x}, \mathbf{y}) = \frac{2\sigma_{xy} + C_3}{\sigma_x \sigma_y + C_3} \tag{10}$$

where μ, σ, and σ represent the mean, standard derivation, and covariance of the image patches, respectively, and C_1, C_2, and C_3 are positive constants used to avoid instability when the denominators are close to zero. Subsequently, the local SSIM index is defined as the product of the three components, which gives

$$\text{SSIM}(\mathbf{x}, \mathbf{y}) = [l(\mathbf{x}, \mathbf{y})]^\alpha [c(\mathbf{x}, \mathbf{y})]^\beta [s(\mathbf{x}, \mathbf{y})]^\gamma. \tag{11}$$

The SSIM index of the whole image is obtained by averaging (or weighted averaging) the local SSIM indices obtained using a sliding window that runs across the image.

Fig. 2 gives a graphical explanation in the vector space of image components, where the image components can be pixels, wavelet coefficients, or extracted features from the reference image. For the purpose of illustration, 2-D diagrams are shown here. However, the actual dimensions may be equal to the number of pixels or features being compared. The three vectors represent three reference images and the contours around them represent the images with the same distortion level using (a) MSE and (b) SSIM as the distortion/quality measures, respectively. The critical difference is in the shapes of the contours. Unlike MSE (where all three contours have the same size and shape), SSIM is adaptive according to the reference image. In particular, if the "direction" of distortion is consistent with the underlying reference (aligned with the direction of the reference vector), the distortion is nonstructural and is much less objectionable than structural distortions (the distortions perpendicular to the reference vector direction). The formulation of SSIM in (11) provides a flexible framework to adjust the relative importance between structural (last term) and nonstructural (first two terms) distortions.

Here we borrow the design philosophy of FR SSIM, but apply it to a completely different domain of image representation. In particular, 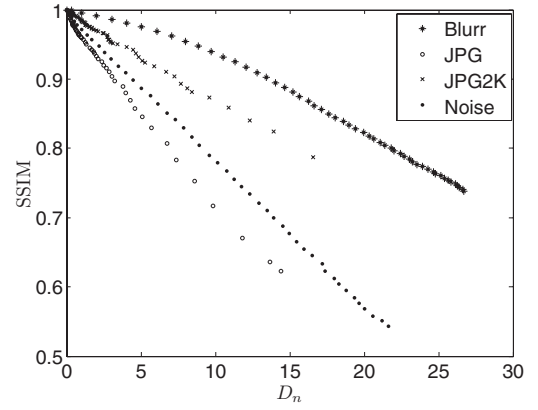we attempt to distinguish structural and nonstructural changes of the cluster of statistical features extracted from the DNT coefficients from different subbands. This is intuitively sensible because the distortion that is consistent with the underlying signal in the feature vector space needs to be treated differently as compared to nonstructural distortions. For example, in the case where the distorted image is a globally contrast-scaled (contrast reduction or enhancement) version of the reference image, then the standard deviations of all subbands should scale by the same factor, which is considered consistent nonstructural distortion and is less objectionable than the case where the subband standard deviations change in different ways.

Let $\boldsymbol{\sigma}_r$ and $\boldsymbol{\sigma}_d$ represent the vectors containing the standard deviation σ values of the DNT coefficients from each subband in the reference and distorted images, respectively. We define a new RR distortion measure as

$$D_n = g(\boldsymbol{\sigma}_r, \boldsymbol{\sigma}_d) \log \left(1 + \frac{1}{D_0} \sum_{k=1}^K \left| \hat{d}^k (p^k || q^k) \right| \right). \tag{12}$$

Compared with (7), the key difference here is the added function $g(\boldsymbol{\sigma}_r, \boldsymbol{\sigma}_d)$ in the front. This function should serve the purpose of differentiating nonstructural from structural distortion directions in the feature vector space of subband σ values, so as to scale the distortion measure D in a way that penalizes more on structural than nonstructural distortions. Motivated by the successful normalized correlation formulation in SSIM [43], we define

$$g(\boldsymbol{\sigma}_r, \boldsymbol{\sigma}_d) = \frac{\|\boldsymbol{\sigma}_r\|^2 + \|\boldsymbol{\sigma}_d\|^2 + C}{2(\boldsymbol{\sigma}_r \cdot \boldsymbol{\sigma}_d) + C} \tag{13}$$

where a positive constant C is included to avoid instability when the dot product $\boldsymbol{\sigma}_r \cdot \boldsymbol{\sigma}_d$ is close to 0. This function is lower-bounded by 1 when $\boldsymbol{\sigma}_r$ and $\boldsymbol{\sigma}_d$ are fully correlated, or in other words, when their directions in the feature vector space are completely aligned (corresponding to nonstructural distortions). With the decrease of correlation, $g(\boldsymbol{\sigma}_r, \boldsymbol{\sigma}_d)$ increases, and thus imposes more penalty to structural distortions.

Fig. 3 plots the D_n values computed using distorted images from the LIVE database [44] for four common distortion types at different distortion levels, and compares them with the corresponding FR SSIM values. Interestingly, for each

fixed distortion type, D_n exhibits a nearly perfect linear relationship with SSIM. We regard this as a consequence of the similarity between their design principle, even though the principle is applied to completely different domains of signal representation. The clean linear relationship helps in reducing the SSIM estimation problem to the estimation of the slope factor. Once the slope is determined, we can then use the following straight-line relationship to estimate SSIM:

$$\hat{S} = 1 - \alpha D_n. \qquad (14)$$

The slope factor α in (14) varies across distortion types and needs to be learned from examples. Specifically, we adopt a regression-by-discretization approach [45], which is a regression scheme that employs a classifier on a copy of the data that has the class attribute discretized, and the predicted value is the expected value of the mean class value for each discretized interval. The training images were obtained from six image databases described in Section III. The classification is performed using random forests [46], which are built using $|\sigma_r - \sigma_d|$ and $|k_r - k_d|$ values in each subband as the attributes, where k_r and k_d are the kurtosis values of the DNT coefficients computed from the reference and distorted images, respectively. It has been observed with the help of ground-truth data that the values of α tend to lie in various closely packed clusters. Each cluster may contain images belonging to one distortion type. It provides a natural order to the distortion types and therefore does not require an undesirable distortion classification stage which limits the generalization capability of the proposed method. Therefore, the proposed method has the potential to extrapolate to extended distortion types that may not be included in the training samples.

The specification of our implementation is as follows. To extract RR features, the reference image is first decomposed into 12 subbands using a three-scale four-orientation steerable pyramid decomposition [47], which is a type of redundant wavelet transform that avoids aliasing in subbands. DNT is then performed using 13 neighboring coefficients, including 9 spatial neighbors from the same subband, 1 from parent subband, and 3 from the same spatial location in the other orientation bands at the same scale. The value of the constant C in (13) is set to 0.1, which is found to be an insensitive parameter in terms of the performance of the proposed IQA measure. Three features, σ_r, k_r, and $d(p_m||p)$, are extracted for each subband, resulting in a total of 36 scalar RR features for a reference image.

III. Validation of RR-IQA Algorithm

Six databases were used to test the proposed algorithm and compare its performance with other IQA algorithms. The databases include.

1) The LIVE database [44] contains seven datasets of 982 subject-rated images, including 779 distorted images with five types of distortions at different distortion levels. The distortion types include: a) JPEG2000 compression (2 sets); b) JPEG compression (2 sets); c) white noise contamination (1 set); d) Gaussian blur (1 set); and e) fast fading channel distortion of JPEG2000

compressed bitstream (1 set). The subjective test was carried out with each dataset individually, and a cross-comparison set that mixes images from all distortion types is then used to align the subject scores across datasets. The alignment process is rather crude, but the aligned subjective scores (all data) are still useful references for testing general-purpose IQA algorithms, for which cross-distortion comparisons are highly desirable.

2) The Cornell-A57 database [48] contains 54 distorted images with six types of distortions: a) quantization of the LH subbands of a five-level discrete wavelet transform, where the subbands were quantized via uniform scalar quantization with step sizes chosen such that the RMS contrast of the distortions was equal; b) additive Gaussian white noise; c) baseline JPEG compression; d) JPEG2000 compression without visual frequency weighting; e) JPEG2000 compression with the dynamic contrast-based quantization algorithm, which applies greater quantization to the fine spatial scales relative to the coarse scales in an attempt to preserve global precedence; and f) blurring by using a Gaussian filter.

3) The IVC database [49], [50] includes 185 distorted images with four types of distortions, which are: a) JPEG compression; b) JPEG2000 compression: c) local adaptive resolution (LAR) coding: and d) blurring.

4) The Toyama-MICT database [51] contains 196 images, including 168 distorted images generated by JPEG and JPEG2000 compression.

5) The Tampere Image database 2008 (TID2008) [52], [53] includes 1700 distorted images with 17 distortion types at four distortion levels. The types of distortions are: a) additive Gaussian noise; b) additive noise in color components more intense than additive noise in the luminance component; c) Spatially correlated noise; d) masked noise; e) high-frequency noise; f) impulse noise; g) quantization noise; h) Gaussian blur; i) image denoising; j) JPEG compression; k) JPEG2000 Compression; l) JPEG transmission errors; m) JPEG2000 transmission Errors, n) Non-eccentricity pattern noise; o) local block-wise distortions of different intensity; p) mean shift (intensity shift); and q) contrast change.

6) The Categorical Image Quality (CSIQ) database [54] contains 866 distorted images of six types of distortions at 4 and 5 distortion levels. The distortion types include JPEG compression, JPEG2000 compression, global contrast decrements, additive pink Gaussian noise, and Gaussian blurring.

To validate the proposed RR-SSIM algorithm, we first test how well it predicts FR SSIM. Fig. 4 shows the scatter plots obtained using all six databases, where each point in the plots represents one test image, and the vertical and horizontal axes are FR-SSIM and RR-SSIM, respectively. If the prediction is perfect, then the point should lie on the diagonal line. To provide a quantitative measure, Table I shows the mean absolute error (MAE) and Pearson linear correlation coefficient (PLCC) between FR SSIM and our RR-SSIM estimate. It can be
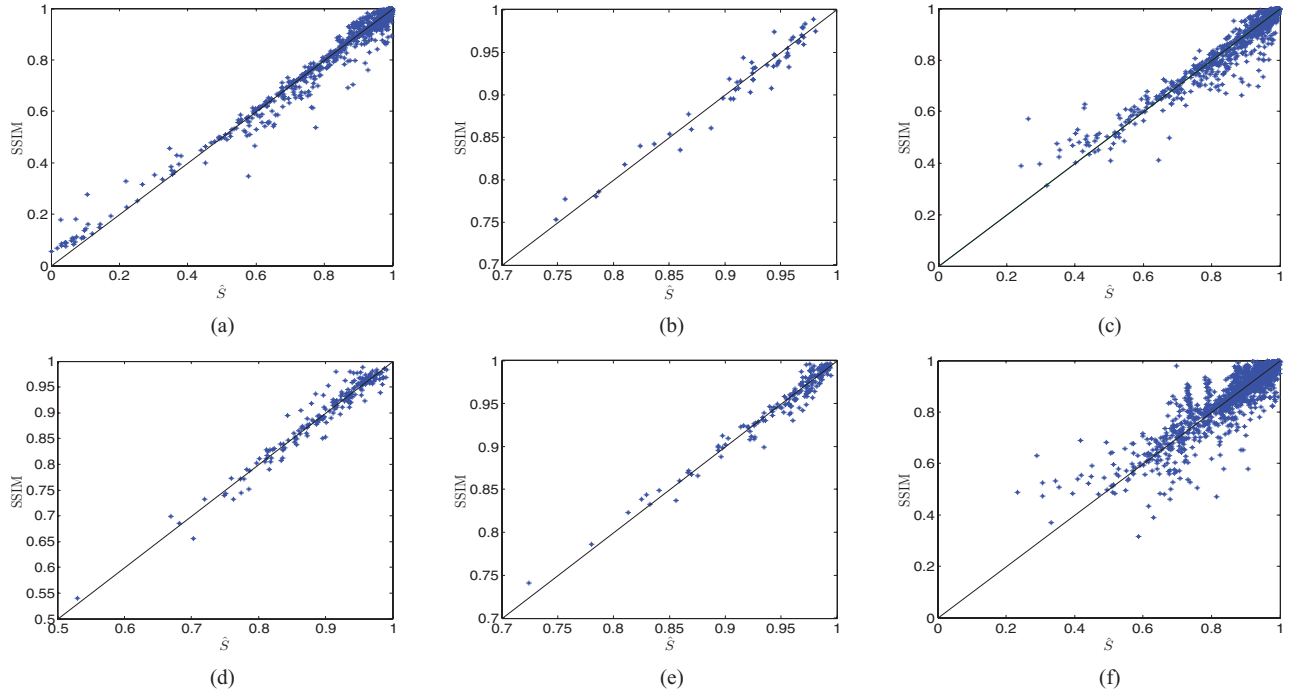
Fig. 4. Scatter plots of SSIM versus RR-SSIM estimation $\hat{S}$ for six test databases. (a) LIVE Image Database. (b) Cornell A57 Database. (c) CSIQ Database. (d) IVC Database. (e) Toyama-MICT Database. (f) TID 2008 Database.

TABLE I
MAE AND PLCC COMPARISONS BETWEEN SSIM AND
RR SSIM ESTIMATION FOR SIX DATABASES

Database	MAE	PLCC
LIVE [44]	0.0317	0.9432
Cornell A57 [48]	0.0266	0.9299
IVC [49], [50]	0.0244	0.9211
Toyama-MICT [51]	0.0119	0.9405
TID2008 [52], [53]	0.0303	0.9004
CSIQ [54]	0.0339	0.9243

observed that, for all databases, the points are scattered close to the diagonal lines in Fig. 4 and the correlation coefficients are above 0.9, indicating good prediction accuracy of the proposed method. The breakdown prediction performance for individual distortion types in different databases are provided in Table II.

The ultimate goal of RR-IQA algorithms is to predict subjective quality evaluation of images. Therefore, the more important test is to evaluate how well they predict subjective scores. For this purpose, we use five evaluation metrics to assess the performance of IQA measures.

1) PLCC after a nonlinear mapping between the subjective and objective scores. For the ith image in an image database of size N, given its subjective score o_i [mean opinion score (MOS) or difference of MOS (DMOS) between reference and distorted images] and its raw objective score r_i, we first apply a nonlinear function to r_i given by [55]

$$q(r) = a_1 \left\{ \frac{1}{2} - \frac{1}{1 + \exp[a_2(r - a_3)]} \right\} + a_4 r + a_5 \quad (15)$$

where a_1–a_5 are model parameters found numerically using a nonlinear regression process in MATLAB

optimization toolbox to maximize the correlations between subjective and objective scores. The PLCC value can then be computed as

$$\text{PLCC} = \frac{\sum_i (q_i - \bar{q}) * (o_i - \bar{o})}{\sqrt{\sum_i (q_i - \bar{q})^2 * \sum_i (o_i - \bar{o})^2}}. \quad (16)$$

2) MAE is calculated using the converted objective scores after the nonlinear mapping described above

$$\text{MAE} = \frac{1}{N} \sum |q_i - o_i|. \quad (17)$$

3) Root mean-squared (RMS) error is computed similarly as

$$\text{RMS} = \sqrt{\frac{1}{N} \sum (q_i - o_i)^2}. \quad (18)$$

4) Spearman's rank correlation coefficient (SRCC) is defined as

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

where d_i is the difference between the ith image's ranks in subjective and objective evaluations. SRCC is a non-parametric rank-based correlation metric, independent of any monotonic nonlinear mapping between subjective and objective scores.

5) Kendall's rank correlation coefficient (KRCC) is another nonparametric rank correlation metric given by

$$\text{KRCC} = \frac{N_c - N_d}{\frac{1}{2} N(N - 1)} \quad (20)$$

where N_c and N_d are the number of concordant and discordant pairs in the dataset, respectively.

TABLE II
DISTORTION TYPE BREAKDOWN FOR MAE AND PLCC COMPARISONS
BETWEEN SSIM AND RR-SSIM ESTIMATION

Distortion type	Database	MAE	PLCC
Additive Gaussian noise	LIVE	0.0340	0.9903
	TID2008	0.0185	0.9522
	CSIQ	0.0274	0.9771
Noise in color comp.	TID2008	0.0080	0.8978
Spatially corr. noise	TID2008	0.0331	0.9580
Masked noise	TID2008	0.0057	0.5982
High frequency noise	TID2008	0.0227	0.9621
Additive pink noise	CSIQ	0.0212	0.9712
Impulse noise	TID2008	0.0222	0.9667
Quantization noise	TID2008	0.0316	0.7584
Gaussian blur	LIVE	0.0412	0.8973
	IVC	0.0342	0.9288
	TID2008	0.0416	0.8892
	CSIQ	0.0260	0.9783
Image denoising	TID2008	0.0444	0.8721
JPEG compression	LIVE (Set 1)	0.0214	0.9867
	LIVE (Set 2)	0.0235	0.9840
	IVC	0.0141	0.9476
	Toyama-MICT	0.0144	0.9007
	TID2008	0.0253	0.9325
	CSIQ	0.0490	0.8895
JPEG2000 compression	LIVE (Set 1)	0.0197	0.9820
	LIVE (Set 2)	0.0229	0.9792
	IVC	0.0296	0.9321
	Toyama-MICT	0.0093	0.9472
	TID2008	0.0482	0.9009
	CSIQ	0.0452	0.9223
LAR compression	IVC	0.0227	0.9426
JPEG trans. error	TID200	0.0420	0.8990
JPEG2000 trans. error	LIVE	0.0434	0.9138
	TID2008	0.0601	0.9074
Non-ecc. patt. noise	TID2008	0.0149	0.8863
Local blockwise dist.	TID2008	0.0117	0.8837
Mean shift	TID2008	0.0367	0.8205
Contrast change	TID2008	0.0485	0.7085
	CSIQ	0.0372	0.9486

Among the above metrics, PLCC, MAE, and RMS are adopted to evaluate prediction accuracy [56], and SRCC and KRCC are employed to assess prediction monotonicity [56]. A better objective IQA measure should have higher PLCC, SRCC, and KRCC, with lower MAE and RMS values. All these evaluation metrics are adopted from previous IQA studies [55]–[57]. Only the distorted images in the six databases were employed in our tests (i.e., reference images are excluded). This avoids several difficulties in computing the evaluation metrics. Specifically, the reference images have infinite peak signal-to-noise-ratio (PSNR) values, making it hard to perform nonlinear regression and compute PLCC, MAE, and MSE values. In addition, since all reference images are assumed to have perfect quality, there are no natural relative ranks between them, resulting in ambiguities when computing SRCC and KRCC metrics.

The test results are given in Tables III and IV. To provide background comparisons, we have also included in the tables four other objective IQA algorithms, among which two are FR-IQA measures, i.e., PSNR and SSIM, and three are RR-IQA measures, i.e., wavelet marginal-based method [23], [24] and

DNT marginal-based method [25]. Other RR-IQA methods are not included in the comparison because they are not designed and tested for general-purpose applications. Although it is unfair to compare RR-IQA with FR-IQA measures, the PSNR and SSIM results supply useful references on the current status of RR approaches. To provide an overall evaluation of the IQA algorithms, we also calculate the direct and weighted average of PLCC, SRCC, and KRCC values across all six databases (where the weight assigned to a database is determined by the number of test images in a database). The average results are given in Table IV. It can be seen that, in general, the proposed RR-SSIM method performs slightly inferior to SSIM (which is as expected) but significantly outperforms PSNR and the other RR-IQA methods under comparison.

Statistical significant analysis has been carried out based on variance-based hypothesis testing, which follows the approach introduced in [55] and subsequently adopted by many later papers in the literature. Specifically, the residual difference between the DMOS and the predicted quality given by each objective IQA algorithm is assumed to be Gaussian-distributed and F-statistic is employed to compare the variances of two sets of sample points. With such a test, we can make a statistically sound judgment of the superiority or inferiority of one IQA algorithm over another. A statistical significance matrix is calculated and given in Table V. Each entry in the table consists of six characters which correspond to the six publicly available databases in the order of {LIVE, A57, CSIQ, IVC, Toyama, TID2008}. The symbol "-" denotes that the two IQA methods are statistically indistinguishable, "1" denotes that the IQA method of the row is statistically better than that of the column, and "0" denotes that the IQA method of the column is better than that of the row. It can be observed that FR-SSIM performs the best among the IQA algorithms under comparison and the performance of the proposed RR-SSIM algorithm is quite close to that of SSIM and is superior to all other IQA methods being compared.

The assumption of Gaussianity is verified with the help of kurtosis values obtained from the prediction residuals. As in [55], the residual values are considered to be Gaussian-distributed if the kurtosis value lies between 2 and 4. The results of Gaussianity tests are given in Table VI, where "1" means the distribution is considered Gaussian and "0" otherwise. It can be observed that the assumption is met in most cases with only a few exceptions.

To examine how the proposed RR-SSIM method performs for different distortion types, we compare it with five other recently proposed RR-IQA algorithms using individual distortion types as well as the "All data" case of the LIVE database. The results are given in Table VII, where the best results for each distortion type are highlighted in bold. It can be observed that the proposed method exhibits highly competitive performance in most cases.

Finally, we compare the computational complexity of the proposed RR-SSIM method with five other RR-IQA algorithms. The results are reported in Table VIII, where we present the average time taken per image, over all the images in the LIVE database, using a computer with Intel i7 processor at 2.67 GHz (the only exception is the method by

TABLE III
PERFORMANCE COMPARISONS OF IQA MEASURES USING SIX DATABASES

IQA measure	Type	LIVE database (779 Images) [44]					Cornell A57 database (54 Images) [48]				
		PLCC	MAE	RMS	SRCC	KRCC	PLCC	MAE	RMS	SRCC	KRCC
PSNR	FR	0.8721	10.5248	13.3683	0.8755	0.6863	0.6346	0.1606	0.1899	0.6188	0.4309
SSIM [1]	FR	0.9448	6.9324	8.9455	0.9479	0.7962	0.8017	0.1209	0.14688	0.8066	0.6058
Wavelet marginal [24]	RR	0.8226	10.5248	13.3683	0.8755	0.6863	0.5125	0.1971	0.2317	0.31398	0.2210
DNT marginal [25]	RR	0.8949	9.7321	11.7862	0.8882	0.7126	0.5474	0.1659	0.2057	0.5058	0.3638
RR-SSIM	RR	0.9194	9.1889	11.3026	0.9129	0.7349	0.7044	0.1433	0.1744	0.7301	0.5345

IQA measure	Type	IVC database (185 Images) [49], [50]					Toyama-MICT database (168 Images) [51]				
		PLCC	MAE	RMS	SRCC	KRCC	PLCC	MAE	RMS	SRCC	KRCC
PSNR	FR	0.6719	0.7190	0.9023	0.6884	0.5217	0.6329	0.7817	0.9688	0.6131	0.4442
SSIM [1]	FR	0.9119	0.3776	0.4999	0.9018	0.7223	0.8886	0.4385	0.5738	0.8793	0.6939
Wavelet marginal [24]	RR	0.5311	0.8550	1.0322	0.4114	0.2907	0.6542	0.7742	0.9464	0.6322	0.4570
DNT marginal [25]	RR	0.6316	0.7842	0.9446	0.6099	0.4364	0.6733	0.7507	0.9253	0.6521	0.4764
RR-SSIM	RR	0.8177	0.5619	0.7014	0.8154	0.6164	0.8051	0.5648	0.7423	0.8003	0.6090

IQA measure	Type	TID 2008 database (1700 Images) [52], [53]					CSIQ database (866 Images) [54]				
		PLCC	MAE	RMS	SRCC	KRCC	PLCC	MAE	RMS	SRCC	KRCC
PSNR	FR	0.5232	0.8683	1.1435	0.5530	0.4027	0.7512	0.1366	0.1732	0.8058	0.6083
SSIM [1]	FR	0.7731	0.6546	0.8510	0.7749	0.5767	0.8612	0.0991	0.1334	0.8756	0.6906
Wavelet marginal [24]	RR	0.5891	0.8666	1.0843	0.5119	0.3589	0.7124	0.1492	0.1842	0.7431	0.5457
DNT marginal [25]	RR	0.5964	0.8287	1.0772	0.5722	0.4188	0.7009	0.1535	0.1872	0.7027	0.5176
RR-SSIM	RR	0.7231	0.7190	0.9270	0.7210	0.5236	0.8426	0.1092	0.1413	0.8527	0.6540

TABLE IV
AVERAGE PERFORMANCE OF IQA MEASURES OVER SIX DATABASES

IQA measure	Type	Direct average			Database-size weighted average		
		PLCC	SRCC	KRCC	PLCC	SRCC	KRCC
PSNR	FR	0.6811	0.6924	0.5157	0.6622	0.6887	0.5172
SSIM [1]	FR	0.8636	0.8643	0.6809	0.8416	0.8455	0.6615
Wavelet marginal [24]	RR	0.6371	0.5813	0.4266	0.6651	0.6383	0.4691
DNT marginal [25]	RR	0.6741	0.6552	0.4876	0.6870	0.6724	0.5053
RR-SSIM	RR	0.8021	0.8054	0.6121	0.7995	0.7996	0.6061

TABLE V
STATISTICAL SIGNIFICANCE MATRIX BASED ON IQA − DMOS RESIDUALS

Model	PSNR	SSIM	Wavelet marginal [24]	DNT marginal [25]	RR-SSIM
PSNR	- - - - - -	0 - 0 0 0 0	1 - - - - 0	0 - 1 - - 0	0 - 0 0 0 0
SSIM	1 - 1 1 1 1	- - - - - -	1 1 1 1 1 1	1 1 1 1 1 1	1 - - 1 - -
Wavelet Marginal [24]	0 - - - - 1	0 0 0 0 0 0	- - - - - -	0 - - - - -	0 0 0 0 0 0
DNT Marginal [25]	1 - 0 - - 1	0 0 0 0 0 0	1 - - - - -	- - - - - -	0 - 0 0 0 0
RR-SSIM	1 - 1 1 1 1	0 - - 0 - -	1 1 1 1 1 1	1 - 1 1 1 1	- - - - - -

Ma *et al.* [60], which was tested on a slightly faster computer). This measurement provides a rough estimate of the relative computational complexity between different RR-IQA algorithms, as no code optimization has been done. It can be seen that the proposed method takes only slightly more time than most of the other methods under comparison, mainly due to the computation of the DNT. The additional computational cost is compensated by the improved quality prediction performance, as shown in Table VII.

IV. IMAGE REPAIRING USING RR FEATURES

Since the RR features reflect certain properties about the reference image and these properties may be altered in the distorted image, they may be employed to partially "repair" the distorted image. Here we provide an example that uses DNT-domain RR features to correct blurred images without any knowledge about the blur kernel.

Since blur reduces energy at mid- and high-frequencies, the subband standard deviation σ_d of DNT coefficients in the distorted image is smaller than that of the reference image σ_r. A straightforward way to enforce a "corrected" image to have the same statistical properties as the reference image is to scale up all DNT coefficients in the subband of the distorted image by a fixed scale factor, followed by an inverse DNT to create a reconstructed image. In practice, however, inverting a DNT transform is a nontrivial issue that requires specific conditions of the coefficients and may involve computationally expensive algorithms [61].

Here we propose a different approach that attempts to match DNT-domain statistics but avoids direct inversion of DNT. The idea is to use the DNT-domain statistics to estimate the scale factors and then apply them in the wavelet domain rather than DNT domain. As a result, only inverse wavelet transform is necessary, and the remaining question becomes whether the

TABLE VI

GAUSSIANITY OF IQA−DMOS RESIDUALS

	LIVE	A57	CSIQ	IVC	Toyama	TID 2008
PSNR	1	1	1	1	1	1
SSIM	1	1	0	0	1	1
Wavelet Marginal [24]	1	1	1	1	1	1
DNT Marginal [25]	1	1	1	1	1	1
RR-SSIM	1	1	1	0	1	1

TABLE VII

PERFORMANCE COMPARISON OF RR-IQA ALGORITHMS USING LIVE DATABASE

Distortion	JP2(1)	JP2(2)	JPG(1)	JPG(2)	Noise	Blur	FF	All data
	PLCC							
Wavelet marginal [24]	0.9339	0.9488	0.8278	0.9566	0.8769	0.8395	0.9230	0.8284
DNT marginal [25]	0.9470	0.9625	0.8228	0.9627	0.9598	**0.9523**	**0.9438**	0.8949
βW-SCM [58]	0.9514	0.9569	0.8673	0.9568	0.9755	0.9454	0.9243	0.8353
Zhang *et al.* [59]	0.9087	0.9511	0.9094	**0.9777**	0.8623	0.9234	0.9392	0.8744
Ma *et al.* [60]	0.8065	0.8819	0.8180	0.9663	0.8769	0.9092	0.9178	0.8841
RR-SSIM	**0.9597**	**0.9632**	**0.9448**	0.9761	**0.9772**	0.9154	0.9315	**0.9194**
	SRCC							
Wavelet marginal [24]	0.9370	0.9419	0.8109	0.8936	0.8600	0.8757	0.9212	0.8270
DNT marginal [25]	0.9439	**0.9556**	0.8246	0.8853	0.9508	**0.9599**	**0.9431**	0.8882
βW-SCM [58]	0.9495	0.9517	0.8535	0.8705	0.9715	0.9371	0.9258	0.8391
Zhang *et al.* [59]	0.9134	0.9495	0.9105	**0.9294**	0.8417	0.9265	0.9365	0.8832
Ma *et al.* [60]	0.7945	0.8717	0.8042	0.9100	0.8619	0.9214	0.8866	0.8807
RR-SSIM	**0.9555**	0.9539	**0.9493**	0.8978	**0.9642**	0.8692	0.9137	**0.9129**

TABLE VIII

COMPARISON OF COMPUTATION TIME USING LIVE DATABASE (SECONDS/IMAGE)

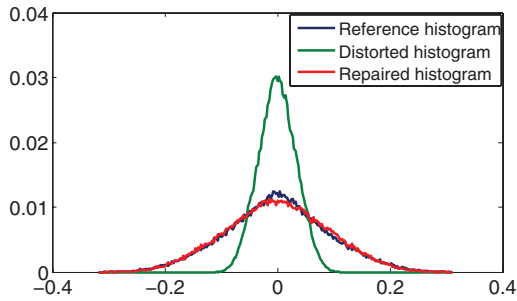| Model	Wavelet marginal [24]	DNT marginal [25]	βW-SCM [58]	Zhang *et al.* [59]	Ma *et al.* [60]	RR-SSIM
Time	6.3719	10.3843	6.6258	3.4937	18	11.2309

Fig. 5. DNT-coefficient histograms of original, distorted, and repaired images.

desired scale ratio in the DNT domain can be well matched by scaling in the wavelet domain. To ensure this, we apply our approach in an iterative manner, and the resulting algorithm is given by Algorithm 1. In our experiment, we find that this iterative algorithm converges quickly, and typically three iterations are enough to reconstruct a stable repaired image (and thus $J = 3$ in Algorithm 1) that matches the DNT-domain statistics quite well. This is demonstrated in Fig. 5, which compares the subband histograms of the reference, distorted, and repaired DNT coefficients. It can be observed that the histogram of the scaled DNT coefficients very well approximates that of the reference image. A similar design philosophy of iteratively synthesizing images by matching desirable statistical features has been used before in the literature for texture synthesis, e.g., [62].

Algorithm 1 Iterative image repairing algorithm

1) *Initialization:* Let $j = 0$, $\hat{x}_{(0)} = y$, where y is the distorted image
2) *Repeat J times*
   a) *Wavelet transform:* Compute wavelet transform of $\hat{x}_{(j)}$, resulting in wavelet coefficients ω
   b) *DNT stage:* Compute DNT from ω, resulting in DNT coefficients v; For all i, in the ith subband, calculate std of DNT coefficients σ_v^i
   c) *Scaling factor calculation:* For all i, in the ith subband, compute the scale factor $s^i = \sigma_r^i/\sigma_v^i$, where σ_r^i is the std of DNT coefficients of the reference image (obtained as RR features)
   d) *Wavelet coefficient scaling:* For all i, in the ith subband, let $\omega_{\text{new}} = s^i \omega$
   e) *Image reconstruction:* Compute inverse wavelet transform of ω_{new}, resulting in $\hat{x}_{(j+1)}$
   f) Increase j by 1
3) *Report reconstructed image:* $\hat{x} = \hat{x}_{(J)}$

An interesting feature of the above image deblurring process is that it does not require any information about the blur kernel. Depending on the nature of the blur process, the energy reductions at different subbands are different. For example, out-of-focus blur may lead to uniform energy reduction in
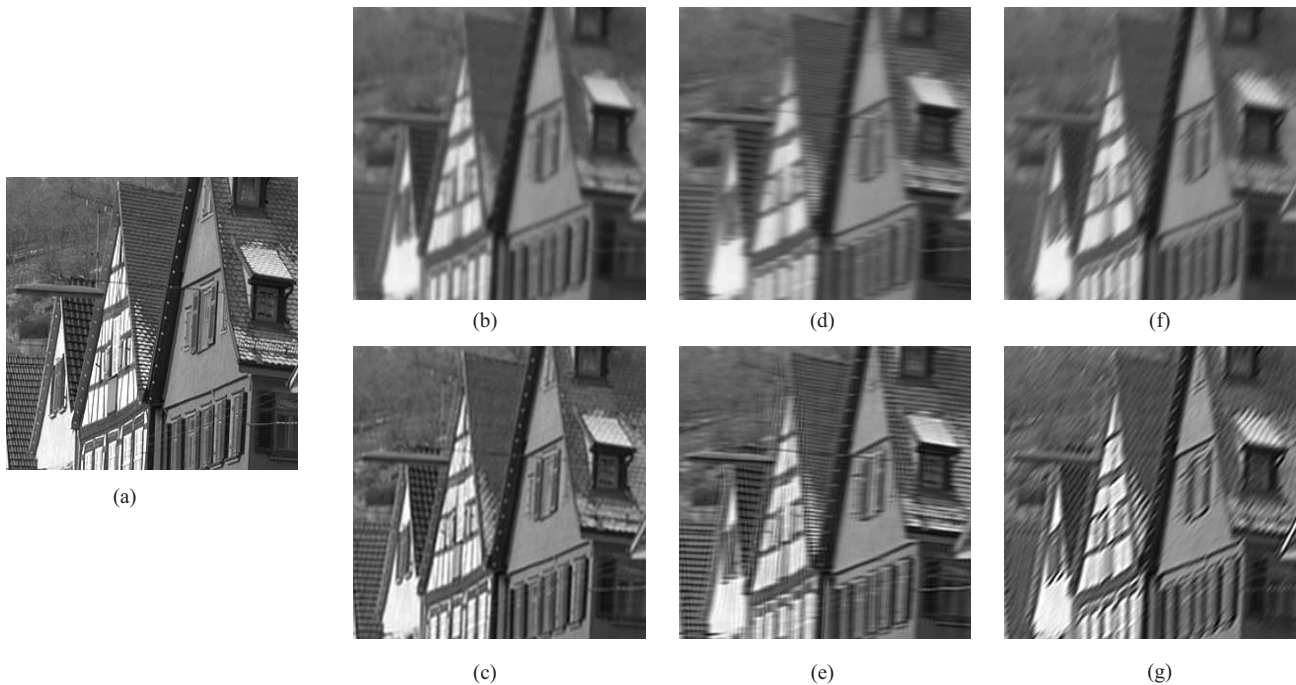
Fig. 6. Repairing homogeneously and directionally burred images using RR features. (a) Original "building" image (cropped for visibility). (b) Homogeneously blurred image, SSIM = 0.7389, $(\hat{S})$ = 0.7118. (c) Repaired image SSIM = 0.9142, $(\hat{S})$ = 0.9327. (d) Directionally blurred image (0 degree), SSIM = 0.6734, $(\hat{S})$ = 0.6821. (e) Repaired image SSIM = 0.7991, $(\hat{S})$ = 0.8063. (f) Directionally blurred image (45 degree), SSIM = 0.6612, $(\hat{S})$ = 0.6324. (g) Repaired image SSIM = 0.7896, $(\hat{S})$ = 0.8135.

all orientation subbands, while motion blur could result in more significant energy reduction along one orientation against another. Since the scale factor s in our algorithm is computed for individual subbands independently, it could automatically adapt the energy correction factors based on the energy reduction occurred in individual subbands. Fig. 6 provides an example, where the homogeneously Gaussian blurred and directionally motion blurred images at different angles are deblurred using exactly the same image repairing algorithm described above. All repaired images appear to be much sharper and have higher contrast than their blurred versions. The visual effect is also reflected by both FR-SSIM and the proposed RR-SSIM evaluations.

One needs to be aware that the RR features only provide limited amount of additional information about the reference image and such information is global in the current implementation (due to the nature of the extracted RR features). Therefore, the same repairing process may or may not work as effectively as we observe in Fig. 6 for the types of image distortions other than linear blur. In the future, more advanced image repairing methods may be developed that make the best use of the RR features as side information in the image repairing process, though these methods are beyond the scope of this paper.

V. CONCLUSION

We proposed an RR-IQA algorithm in an attempt to approximate FR-SSIM by making use of DNT-domain image statistical properties and the design principle of the SSIM approach. Experimental results using six publicly available subject-rated image databases showed that the proposed RR-SSIM method exhibits good correlations with not only FR-SSIM but also subjective evaluations of image quality over a wide variety of image distortions. We also demonstrated the concept of image repairing by iteratively matching the DNT-domain statistical properties (available as RR features) of the reference image. The proposed method has a fairly low RR data rate (36 scalar features per image in the current implementation) and has good potential to be employed in visual communications applications for quality monitoring, streaming, and image repairing tasks.

ACKNOWLEDGMENT

REFERENCES

[1] Z. Wang, A. Bovik, H. Sheikh, and E. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600–612, Apr. 2004.

[2] Z. Wang, E. P. Simoncelli, and A. C. Bovik, "Multiscale structural similarity for image quality assessment," in *Proc. IEEE 37th Asilomar Conf. Signals, Syst., Comput.*, Pacific Grove, CA, Nov. 2003, pp. 1398–1402.

[3] Z. Wang and Q. Li, "Information content weighting for perceptual image quality assessment," *IEEE Trans. Image Process.*, vol. 20, no. 5, pp. 1185–1198, May 2011.

[4] H. Sheikh, A. Bovik, and G. Veciana, "An information fidelity criterion for image quality assessment using natural scene statistics," *IEEE Trans. Image Process.*, vol. 14, no. 12, pp. 2117–2128, Dec. 2005.

[5] D. Chandler and 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.

[6] 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. 011006-1–011006-21, Jan.–Mar. 2010.

[7] Z. Wang and A. Bovik, "Mean squared error: Love it or leave it? A new look at signal fidelity measures," *IEEE Signal Process. Mag.*, vol. 26, no. 1, pp. 98–117, Jan. 2009.

[8] D. Brunet, E. R. Vrscay, and Z. Wang, "On the mathematical properties of the structural similarity index," *IEEE Trans. Image Process.*, vol. 24, no. 4, pp. 1488–1499, Apr. 2012.

[9] S. Wang, A. Rehman, Z. Wang, S. Ma, and W. Gao, "SSIM-motivated rate distortion optimization for video coding," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 22, no. 4, pp. 516–529, Apr. 2012.

[10] S. Wang, A. Rehman, Z. Wang, S. Ma, and W. Gao, "SSIM-inspired divisive normalization for perceptual video coding," in *Proc. 18th IEEE Int. Conf. Image Process.*, Brussels, Belgium, Sep. 2011, pp. 1657–1660.

[11] Y. Gao, A. Rehman, and Z. Wang, "CW-SSIM based image classification," in *Proc. 18th IEEE Int. Conf. Image Process.*, Brussels, Belgium, Sep. 2011, pp. 1249–1252.

[12] A. Rehman and Z. Wang, "SSIM-based non-local means image denoising," in *Proc. 18th IEEE Int. Conf. Image Process.*, Brussels, Belgium, Sep. 2011, pp. 217–220.

[13] A. Rehman, Z. Wang, D. Brunet, and E. Vrscay, "SSIM-inspired image denoising using sparse representations," in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process.*, May 2011, pp. 1121–1124.

[14] Z. Wang and A. C. Bovik, *Modern Image Quality Assessment.* San Mateo, CA: Morgan Kauffmann, Mar. 2006.

[15] H. R. Wu and M. Yuen, "A generalized block-edge impairment metric for video coding," *IEEE Signal Process. Lett.*, vol. 4, no. 11, pp. 317–320, Nov. 1997.

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

[17] Z. Yu, H. R. Wu, S. Winkler, and T. Chen, "Vision-model-based impairment metric to evaluate blocking artifact in digital video," *Proc. IEEE*, vol. 90, no. 1, pp. 154–169, Jan. 2002.

[18] Z. Wang, H. R. Sheikh, and A. C. Bovik, "No-reference perceptual quality assessment of JPEG compressed images," in *Proc. IEEE Int. Conf. Image Process.*, Sep. 2002, pp. I-477–I-480.

[19] 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. 11, pp. 1918–1927, Nov. 2005.

[20] P. Marziliano, F. Dufaux, S. Winkler, and T. Ebrahimi, "Perceptual blur and ringing metrics: Application to JPEG2000," *Signal Process.: Image Commun.*, vol. 19, no. 2, pp. 163–172, 2004.

[21] X. Li, "Blind image quality assessment," in *Proc. IEEE Int. Conf. Image Process.*, vol. 1. Sep. 2002, pp. 449–452.

[22] Z. Wang, H. R. Sheikh, and A. C. Bovik, "Objective video quality assessment," in *The Handbook of Video Databases: Design and Applications*, B. Furht and O. Marques, Eds. Boca Raton, FL: CRC Press, Sep. 2003, pp. 1041–1078.

[23] Z. Wang and E. P. Simoncelli, "Reduced-reference image quality assessment using a wavelet-domain natural image statistic model," *Proc. SPIE*, vol. 5666, pp. 149–159, Jan. 2005.

[24] 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.

[25] 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.

[26] S. Wolf and M. H. Pinson, "Spatio-temporal distortion metrics for in-service quality monitoring of any digital video system," *Proc. SPIE*, vol. 3845, pp. 266–277, 1999.

[27] T. 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.*, Oct. 2003, pp. 71–74.

[28] I. P. Gunawan and M. Ghanbari, "Reduced reference picture quality estimation by using local harmonic amplitude information," in *Proc. London Commun. Symp.*, Sep. 2003, pp. 137–140.

[29] 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 Exposit.*, Apr. 2008, pp. 609–612.

[30] 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.*, vol. 3. Sep. 2003, pp. III-185–III-188.

[31] M. Carnec, P. Le Callet, and D. Barba, "Visual features for image quality assessment with reduced reference," in *Proc. IEEE Int. Conf. Image Process.*, vol. 1. Sep. 2005, pp. I-421–I-424.

[32] H. B. Barlow, "Possible principles underlying the transformation of sensory messages," in *Sensory Communication*, W. A. Rosenblith, Ed. Cambridge, MA: MIT Press, 1961, pp. 217–234.

[33] E. P. Simoncelli and B. Olshausen, "Natural image statistics and neural representation," *Annu. Rev. Neurosci.*, vol. 24, pp. 1193–1216, May 2001.

[34] M. Tagliasacchi, G. Valenzise, M. Naccari, and S. Tubaro, "A reduced-reference structural similarity approximation for videos corrupted by channel errors," *Multim. Tools Appl.*, vol. 48, no. 3, pp. 471–492, 2010.

[35] D. J. Heeger, "Normalization of cell responses in cat striate cortex," *Visual Neural Sci.*, vol. 9, no. 2, pp. 181–198, 1992.

[36] 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, Mar. 1998.

[37] M. J. Wainwright, "Visual adaptation as optimal information transmission," *Vis. Res.*, vol. 39, no. 23, pp. 3960–3974, 1999.

[38] J. Foley, "Human luminance pattern mechanisms: Masking experiments require a new model," *J. Opt. Soc. Amer.*, vol. 11, no. 6, pp. 1710–1719, 1994.

[39] A. B. Watson and J. A. Solomon, "Model of visual contrast gain control and pattern masking," *J. Opt. Soc. Amer.*, vol. 14, no. 9, pp. 2379–2391, 1997.

[40] M. J. Wainwright and E. P. Simoncelli, "Scale mixtures of Gaussians and the statistics of natural images," in *Advanced Neural Information Processing Systems*. Cambridge, MA: MIT Press, 2000, pp. 855–861.

[41] 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.

[42] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. New York: Wiley, 1991.

[43] 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.

[44] H. R. Sheikh, Z. Wang, A. C. Bovik, and L. K. Cormack. (2006). *Image and Video Quality Assessment Research at LIVE* [Online]. Available: http://live.ece.utexas.edu/research/quality/

[45] S. M. Weiss and N. Indurkhya, "Rule-based machine learning methods for functional prediction," *J. Artif. Intell. Res.*, vol. 3, no. 1, pp. 383–403, 1995.

[46] L. Breiman, "Random forests," *Mach. Learn.*, vol. 45, no. 1, pp. 5–32, 2001.

[47] E. P. Simoncelli, W. T. Freeman, E. H. Adelson, and D. J. Heeger, "Shiftable multiscale transforms," *IEEE Trans. Inf. Theory*, vol. 38, no. 2, pp. 587–607, Mar. 1992.

[48] 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.

[49] P. Le Callet and F. Autrusseau. (2005). *Subjective Quality Assessment IRCCYN/IVC Database* [Online]. Available: http://www.irccyn.ec-nantes.fr/ivcdb/

[50] A. Ninassi, P. L. Callet, and F. Autrusseau, "Pseudo no reference image quality metric using perceptual data hiding," *Proc. SPIE*, vol. 6057, no. 1, pp. 60570G-1–60570G-12, 2006.

[51] Y. Horita, K. Shibata, Y. Kawayoke, and Z. M. Parvez. (2010). *Mict Image Quality Evaluation Database* [Online]. Available: http://mict.eng.u-toyama.ac.jp/mictdb

[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, pp. 30–45, 2009.

[53] N. Ponomarenko and K. Egiazarian. (2008). *Tampere Image Database TID2008* [Online]. Available: http://www.ponomarenko.info/tid2008

[54] E. C. Larson and D. M. Chandler. (2010). *Categorical Image Quality (CSIQ) Database* [Online]. Available: http://vision.okstate.edu/csiq

[55] H. Sheikh, M. Sabir, and A. 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.

[56] "Final report from the video quality experts group on the validation of objective models of video quality assessment," Video Quality Experts Group (VQEG), San Jose, CA, Tech. Rep. COM 9-80-E, Apr. 2000.

[57] N. Ponomarenko, F. Battisti, K. Egiazarian, J. Astola, and V. Lukin, "Metrics performance comparison for color image database," in *Proc. 4th Int. Workshop Video Process. Qual. Metrics Consumer Electron.*, Scottsdale, AZ, Jan. 2009, pp. 1–6.

[58] W. Xue and X. Mou, "Reduced reference image quality assessment based on Weibull statistics," in *Proc. 2nd Int. Workshop Quality Multim. Experience*, Jun. 2010, pp. 1–6.

[59] M. Zhang, W. Xue, and X. Mou, "Reduced reference image quality assessment based on statistics of edge," *Proc. SPIE*, vol. 7876, no. 1, p. 787611, Jan. 2011.

[60] 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.

[61] J. Malo, I. Epifanio, R. Navarro, and E. P. Simoncelli, "Non-linear image representation for efficient perceptual coding," *IEEE Trans. Image Process.*, vol. 15, no. 1, pp. 68–80, Jan. 2006.

[62] 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. 49–71, Dec. 2000.

Abdul Rehman (S'10) received the B.S. degree in electrical engineering from the National University of Sciences and Technology, Rawalpindi, Pakistan, in 2007, and the M.Sc. degree in communications engineering from Technical University Munich, Munich, Germany, in 2009. He is currently pursuing the Ph.D. degree in electrical and computer engineering from the University of Waterloo, Waterloo, ON, Canada.

He has been a Research Assistant with the Department of Electrical and Computer Engineering, University of Waterloo, since 2009. In 2011, he was with the Video Compression Research Group, Research in Motion, Waterloo. From 2007 to 2009, he was a Research and Teaching Assistant with the Department of Electrical Engineering and Information Technology, Technical University Munich. His current research interests include image and video processing, coding, communication and quality assessment, machine learning, and compressed sensing.

Zhou Wang (S'97–A'01–M'02) received the Ph.D. degree in electrical and computer engineering from the University of Texas, Austin, in 2001.

He is currently an Associate Professor with the Department of Electrical and Computer Engineering, University of Waterloo, Waterloo, ON, Canada. He has published more than 100 publications with over 10 000 citations (Google Scholar). He is a co-author of *Modern Image Quality Assessment* (Morgan & Claypool, 2006). His current research interests include image processing, coding and quality assessment, computational vision and pattern analysis, multimedia communications, and biomedical signal processing.

Dr. Wang served as an Associate Editor of the IEEE TRANSACTIONS ON IMAGE PROCESSING in 2009, the IEEE SIGNAL PROCESSING LETTERS from 2006 to 2010, and *Pattern Recognition* in 2006. He was a Guest Editor of the IEEE JOURNAL OF SELECTED TOPICS IN SIGNAL PROCESSING from 2007 to 2009, the *EURASIP Journal of Image and Video Processing* from 2009 to 2010, and *Signal, Image, and Video Processing* in 2011. He was a recipient of the IEEE Signal Processing Best Paper Award in 2009, the ICIP IBM Best Student Paper Award in 2008 as a senior author, and the Ontario Early Researcher Award in 2009.