No-Reference Image Quality Assessment in the Spatial Domain

Anish Mittal, Anush Krishna Moorthy, and Alan Conrad Bovik, *Fellow, IEEE*

Abstract—We propose a natural scene statistic-based distortion-generic blind/no-reference (NR) image quality assessment (IQA) model that operates in the spatial domain. The new model, dubbed blind/referenceless image spatial quality evaluator (BRISQUE) does not compute distortion-specific features, such as ringing, blur, or blocking, but instead uses scene statistics of locally normalized luminance coefficients to quantify possible losses of "naturalness" in the image due to the presence of distortions, thereby leading to a holistic measure of quality. The underlying features used derive from the empirical distribution of locally normalized luminances and products of locally normalized luminances under a spatial natural scene statistic model. No transformation to another coordinate frame (DCT, wavelet, etc.) is required, distinguishing it from prior NR IQA approaches. Despite its simplicity, we are able to show that BRISQUE is statistically better than the full-reference peak signal-to-noise ratio and the structural similarity index, and is highly competitive with respect to all present-day distortion-generic NR IQA algorithms. BRISQUE has very low computational complexity, making it well suited for real time applications. BRISQUE features may be used for distortion-identification as well. To illustrate a new practical application of BRISQUE, we describe how a nonblind image denoising algorithm can be augmented with BRISQUE in order to perform blind image denoising. Results show that BRISQUE augmentation leads to performance improvements over state-of-the-art methods. A software release of BRISQUE is available online: http://live.ece.utexas.edu/research/quality/BRISQUE_release.zip for public use and evaluation.

Index Terms—Blind quality assessment, denoising, natural scene statistics, no reference image quality assessment, spatial domain.

I. INTRODUCTION

WITH THE launch of networked handheld devices which can capture, store, compress, send and display a variety of audiovisual stimuli; high definition television (HDTV); streaming Internet protocol TV (IPTV) and websites such as Youtube, Facebook and Flickr etc., an enormous amount of visual data of visual data is making its way to consumers. Because of this, considerable time and resources are being

The authors are with the Laboratory for Image and Video Engineering, Department of Electrical and Computer Engineering, University of Texas at Austin, Austin, TX 78712 USA (e-mail: mittal.anish@gmail.com; anushmoorthy@gmail.com; bovik@ece.utexas.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.2012.2214050

expanded to ensure that the end user is presented with a satisfactory quality of experience (QoE) [1]. While traditional QoE methods have focused on optimizing delivery networks with respect to throughput, buffer-lengths and capacity, perceptually optimized delivery of multimedia services is also fast gaining importance. This is especially timely given the explosive growth in (especially wireless) video traffic and expected shortfalls in bandwidth. These perceptual approaches attempt to deliver an optimized QoE to the end-user by utilizing objective measures of visual quality.

Objective blind or No-reference (NR) image quality assessment (IQA) refers to automatic quality assessment of an image using an algorithm such that the only information that the algorithm receives before it makes a prediction on quality is the distorted image whose quality is being assessed. On the other end of the spectrum lie full-reference (FR) algorithms that require as input not only the distorted image, but also a 'clean', pristine reference image with respect to which the quality of the distorted image is assessed. Somewhere between these two extremes lie reduced-reference (RR) approaches that possess some information regarding the reference image (eg., a watermark), but not the actual reference image itself, apart from the distorted image [1]–[3].

Our approach to NR IQA is based on the principle that natural images[1] possess certain *regular* statistical properties that are measurably modified by the presence of distortions. Figure 1(a) and (b) shows examples of natural and artificial images from the TID database [4] respectively. The normalized luminance coefficients (explained later) of the natural image closely follow Gaussian-like distribution, as shown in Fig. 1(c) while the same doesnot hold for the empirical distribution of the artificial image shown in Fig. 1(d).

Deviations from the regularity of natural statistics, when quantified appropriately, enable the design of algorithms capable of assessing the perceptual quality of an image without the need for any reference image. By quantifying natural image statistics and refraining from an explicit characterization of distortions, our approach to quality assessment is not limited by the type of distortions that afflict the image. Such approaches to NR IQA are significant since most current approaches are distortion-specific [5]–[11], i.e., they are capable of performing blind IQA only if the distortion

---

[1] 'Natural' images are not necessarily images of natural environments such as trees or skies. Any natural light image that is captured by an optical camera and is not subjected to artificial processing on a computer is regarded as a natural image. Of course, image sensors may capture natural radiation other than visible light, but the images formed may obey different NSS than those considered here.
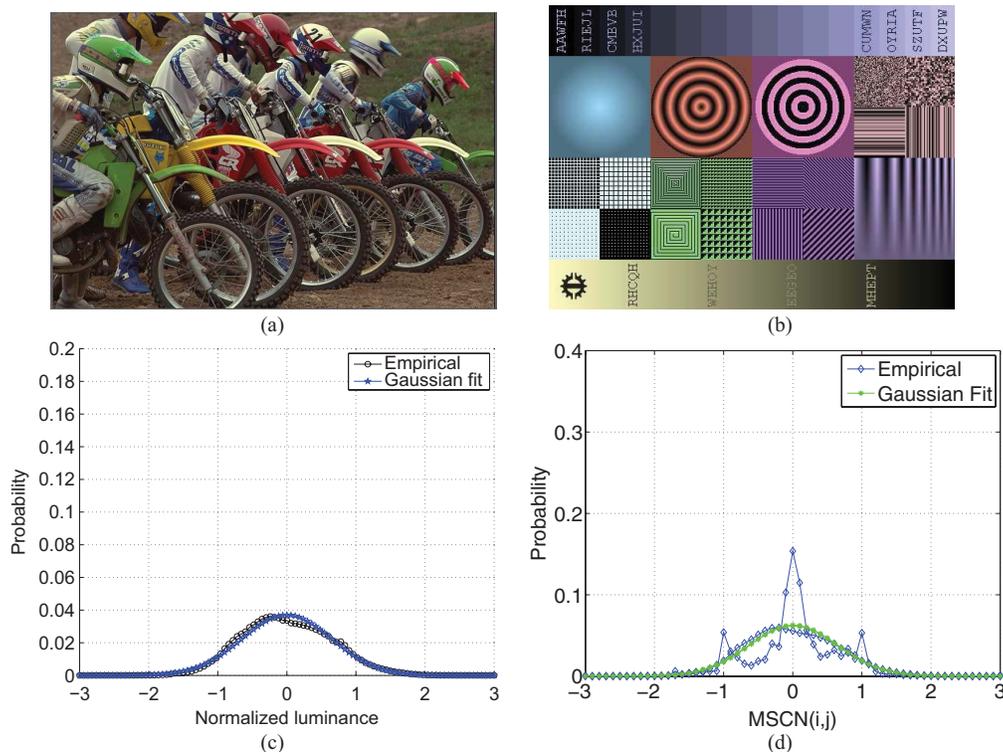
Fig. 1. Underlying Gaussianity of natural images. Examples of (a) natural images and (b) artificial images from the TID database [4]. (c) shows that normalized luminance coefficients follow a nearly Gaussian distribution for the natural image (a). (d) shows that this property does not hold true for the empirical distribution of the artificial image (b).

that afflicts the image is known beforehand, e.g., blur or noise or compression and so on (see below). Previously, we have proposed other NSS-based distortion-generic approaches to NR IQA that statistically model images in the wavelet domain [12] and in the DCT-domain [13]. Our contribution here is a new NR IQA model that is purely spatial; that relies on a spatial NSS model which does not require a mapping to a different co-ordinate domain (wavelet, DCT, etc.) and so is 'transform-free'; that demonstrates better ability to predict human judgments of quality than other popular FR and NR IQA models; that is highly efficient; and that is useful for perceptually optimizing image processing algorithms such as denoising.

While the presence of a reference image or information regarding the reference simplifies the problem of quality assessment, practical applications of such algorithms are limited in real-world scenarios where reference information is generally unavailable at nodes where quality computation is undertaken. Further, it can be argued that FR and to a large-extent RR approaches are not quality measures in the true sense, since these approaches measure *fidelity* relative to a reference image. Moreover, the assumption of a pristine nature of any reference is questionable, since all images are ostensibly distorted [14].

The performance of any IQA model is best gauged by its correlation with human subjective judgements of quality, since the human is the ultimate receiver of the visual signal. Such human opinions of visual quality are generally obtained by conducting large-scale human studies, referred

to as subjective quality assessment, where human observers rate a large number of distorted (and possibly reference) signals. When the individual opinions are averaged across the subjects, a mean opinion score (MOS) or differential mean opinion score (DMOS) is obtained for each of the visual signals in the study, where the MOS/DMOS is representative of the perceptual quality of the visual signal. The goal of an objective quality assessment (QA) algorithm is to predict quality scores for these signals such that the scores produced by the algorithm correlate well with human opinions of signal quality (MOS/DMOS). Practical application of QA algorithms requires that these algorithms compute perceptual quality *efficiently*.

The regularity of natural scene statistics (NSS) has been well established in the visual science literature, where regularity has been demonstrated in the spatial domain [15], and in the wavelet domain [16]. For example, it is well known that the power spectrum of natural images is a function of frequency and takes the form $1/f^\gamma$, where γ is an exponent that varies over a small range across natural images.

The product of our research is the *Blind/Referenceless Image Spatial QUality Evaluator (BRISQUE)* which utilizes an NSS model framework of locally normalized luminance coefficients and quantifies 'naturalness' using the parameters of the model. BRISQUE introduces a new model of the statistics of pair-wise products of neighboring (locally normalized) luminance values. The parameters of this model further quantify the naturalness of the image. Our claim is that characterizing locally normalized luminance coefficients

in this way is sufficient not only to quantify naturalness, but also to quantify quality in the presence of distortion.

In this article, we detail the statistical model of locally normalized luminance coefficients in the spatial domain, as well as the model for pairwise products of these coefficients. We describe the statistical features that are used from the model and demonstrate that these features correlate well with human judgements of quality. We then describe how we learn a mapping from features to quality space to produce an automatic blind measure of perceptual quality. We thoroughly evaluate the performance of BRISQUE, and statistically compare BRISQUE performance to state-of-the-art FR and NR IQA approaches. We demonstrate that BRISQUE is highly competitive to these NR IQA approaches, and also statistically better than the popular full-reference peak signal-to-noise-ratio (PSNR) and structural similarity index (SSIM). We show that BRISQUE performs well on independent databases, analyze its complexity and compare it with other NR IQA approaches. Finally, to further illustrate the practical relevance of BRISQUE, we describe how a non-blind image denoising algorithm can be augmented with BRISQUE in order to improve blind image denoising. Results show that BRISQUE augmentation leads to significant performance improvements over the state-of-the-art. Before we describe BRISQUE in detail, we first briefly review relevant prior work in the area of blind IQA.

II. Previous Work

Most existing blind IQA models proposed in the past assume that the image whose quality is being assessed is afflicted by a particular kind of distortion [5]–[11], [17]. These approaches extract distortion-specific features that relate to loss of visual quality, such as edge-strength at block-boundaries. However, a few general purpose approaches for NR IQA have been proposed recently.

Li devised a set of heuristic measures to characterize visual quality in terms of edge sharpness, random noise and structural noise [18] while Gabarda and Cristobal, modeled anisotropies in images using Renyi entropy [19]. The authors in [20] use gabor filter based local appearance descriptors to form a visual codebook, and learn DMOS score vector, associating each word with a quality score. However, in the process of visual codebook formation, each feature vector associated with an image patch is labeled by DMOS asigned to the entire image. This is questionable as each image patch can present a different level of quality depending on the distortion process the image is afflicted with. In particular, local distortions such as packet loss might afflict only a few image patches. Also, the approach is computationally expensive limiting its applicability in real time applications.

Tang *et al.* [21] proposed an approach which learns an ensemble of regressors trained on three different groups of features - natural image statistics, distortion texture statistics and blur/noise statistics. Another approach [22] is based on a hybrid of curvelet, wavelet and cosine transforms. Although these approaches work on a variety of distortions, each set of features (in the first approach) and transforms (in the second) caters only to certain kinds of distortion processes. This limits the applicability of their framework to new distortions.

We have also developed previous NR QA models in the past, following our philosophy, first fully developed in [23], that NSS models provide powerful tools for probing human judgements of visual distortions. Our work on NSS based FR QA algorithms [9], [23], [24], more recent RR models [3] and very recent work on NSS based NR QA [12], [13], [25] have led us to the conclusion that visual features derived from NSS lead to particularly potent and simple QA models [26].

Our recently proposed NSS based NR IQA model, dubbed the Distortion Identification-based Image INtegrity and Verity Evaluation (DIIVINE) index, deploys summary statistics derived from an NSS wavelet coefficient model, using a two stage framework for QA: distortion-identification followed by distortion-specific QA [12]. The DIIVINE index performs quite well on the LIVE IQA database [27], achieving statistical parity with the full-reference structural similarity (SSIM) index [28].

A complementary approach developed at the same time, named BLind Image Notator using DCT Statistics (BLIINDS-II index) is a pragmatic approach to NR IQA that operates in the DCT domain, where a small number of features are computed from an NSS model of block DCT coefficients [13]. Efficient NSS features are calculated and fed to a regression function that delivers accurate QA predictions. BLIINDS-II is a single-stage algorithm that also delivers highly competitive QA prediction power. Although BLIINDS-II index is multi-scale, the small number of feature types (4) allow for efficient computation of visual quality and hence the index is attractive for practical applications.

While both DIIVINE and BLIINDS-II deliver top NR IQA performance (to date), each of them has certain limitations. The large number of features that DIIVINE computes implies that it may be difficult to compute in real time. Although BLIINDS-II is more efficient than DIIVINE, it requires non-linear sorting of block based NSS features, which slows it considerably.

In our continued search for fast and efficient high performance NSS based NR QA indices, we have recently studied the possibility of developing transform-free models that operate directly on the spatial pixel data. Our inspiration for thinking we may succeed is the pioneering work by Ruderman [15] on spatial natural scene modeling, and the success of the spatial multi-scale SSIM index [29], which competes well with transform domain IQA models.

III. Blind Spatial Image Quality Assessment

Much recent work has focused on modeling the statistics of responses of natural images using multiscale transforms (eg., Gabor filters, wavelets etc.) [16]. Given that neuronal responses in area V1 of visual cortex perform scale-space-orientation decompositions of visual data, transform domain models seem like natural approaches, particularly in view of the energy compaction (sparsity) and decorrelating properties

of these transforms when combined with divisive normalization [26], [30]. However, successful models of spatial luminance statistics have also received attention from vision researchers [15].

A. Natural Scene Statistics in the Spatial Domain

The spatial approach to NR IQA that we have developed can be summarized as follows. Given a (possibly distorted) image, first compute locally normalized luminances via local mean subtraction and divisive normalization [15]. Ruderman observed that applying a local *non-linear* operation to log-contrast luminances to remove local mean displacements from zero log-contrast and to normalize the local variance of the log contrast has a decorrelating effect [15]. Such an operation may be applied to a given intensity image $I(i, j)$ to produce:

$$\hat{I}(i, j) = \frac{I(i, j) - \mu(i, j)}{\sigma(i, j) + C} \quad (1)$$

where, $i \in 1, 2 \ldots M$, $j \in 1, 2 \ldots N$ are spatial indices, M, N are the image height and width respectively, $C = 1$ is a constant that prevents instabilities from occurring when the denominator tends to zero (eg., in the case of an image patch corresponding to the plain sky) and

$$\mu(i, j) = \sum_{k=-K}^{K} \sum_{l=-L}^{L} w_{k,l} I_{k,l}(i, j) \quad (2)$$

$$\sigma(i, j) = \sqrt{\sum_{k=-K}^{K} \sum_{l=-L}^{L} w_{k,l}(I_{k,l}(i, j) - \mu(i, j))^2} \quad (3)$$

where $w = \{w_{k,l} | k = -K, \ldots, K, l = -L, \ldots L\}$ is a 2D circularly-symmetric Gaussian weighting function sampled out to 3 standard deviations and rescaled to unit volume. In our implementation, $K = L = 3$. We show how performance varies with changes in the window size in the performance evaluation section.

Ruderman also observed that these normalized luminance values strongly tend towards a unit normal Gaussian characteristic [15] for natural images. Such an operation can be used to model the contrast-gain masking process in early human vision [30], [31]. We utilize the pre-processing model (1) in our QA model development and refer to the transformed luminances $\hat{I}(i, j)$ as mean subtracted contrast normalized (MSCN) coefficients. As illustrated in the left column of Fig. 2, there is high correlation between surrounding pixels because image functions are generally piecewise smooth aside from sparse edge discontinuities. Hence we observe a diagonal kind of structure in the plots shown in the left column. The normalization procedure greatly reduces dependencies between neighboring coefficients as is apparent in the plots shown in the right column.

In order to help the reader visualize what the non-linear transformation does to an image, Figure 3 plots an image from the LIVE IQA database [27], its local mean field $\mu(i, j)$ and local variance field, $\sigma(i, j)$ and the MSCN field. The variance field highlights object boundaries and other local high contrast phenomenon. The MSCN field, while clearly not entirely decorrelated, exhibits a largely homogeneous appearance with a few low-energy residual object boundaries.

Our hypothesis is that the MSCN coefficients have characteristic statistical properties that are changed by the presence of distortion, and that quantifying these changes will make it possible to predict the type of distortion affecting an image as well as its perceptual quality. In order to visualize how the MSCN coefficient distributions vary as a function of distortion, Fig. 4 plots a histogram of MSCN coefficients for a natural undistorted image and for various distorted versions of it. Notice how the reference image exhibits a Gaussian-like appearance, as observed by Ruderman [15], while each distortion modifies the statistics in its own characteristic way. For example, blur creates a more Laplacian appearance, while white-noise distortion appears to reduce the weight of the tail of the histogram. We have found that a generalized Gaussian distribution (GGD) can be used to effectively capture a broader spectrum of distorted image statistics, which often exhibit changes in the tail behaviour (i.e. kurtosis) of the empirical coefficient distributions [32] where the GGD with zero mean is given by:

$$f(x; \alpha, \sigma^2) = \frac{\alpha}{2\beta\Gamma(1/\alpha)} \exp\left(-\left(\frac{|x|}{\beta}\right)^\alpha\right) \quad (4)$$

where

$$\beta = \sigma \sqrt{\frac{\Gamma(1/\alpha)}{\Gamma(3/\alpha)}} \quad (5)$$

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

$$\Gamma(a) = \int_0^\infty t^{a-1} e^{-t} dt \quad a > 0. \quad (6)$$

The shape parameter α controls the 'shape' of the distribution while σ^2 control the variance. We choose the zero mean distribution, since (generally) MSCN coefficient distributions are symmetric. The parameters of the GGD (α, σ^2), are estimated using the moment-matching based approach proposed in [32].

We deploy this parametric model to fit the MSCN empirical distributions from distorted images as well as undistorted ones. For each image, we estimate 2 parameters (α, σ^2) from a GGD fit of the MSCN coefficients. These form the first set of features that will be used to capture image distortion. To show that pristine and distorted images are well separated in GGD parameter space, we took a set of pristine images from the Berkeley image segmentation database [33]. Similar kinds of distortions as present in the LIVE IQA database [27] - JPEG 2000, JPEG, white noise, Gaussian blur, and fast fading channel errors were introduced in each image at varying degrees of severity to form the distorted image set. As shown in Fig. 5(a), pristine and distorted images occupy different regions in this parameter space. White noise is very clearly separated from the pristine image set making it one of the easiest to gauge the quality of JPEG2000 and fast fading have a high degree of overlap as fast fading images in LIVE database are actually multidistorted, first compressed into a bitstream using a JPEG2000 codec, then passed through a Rayleigh fast fading channel to simulate packet loss [27].
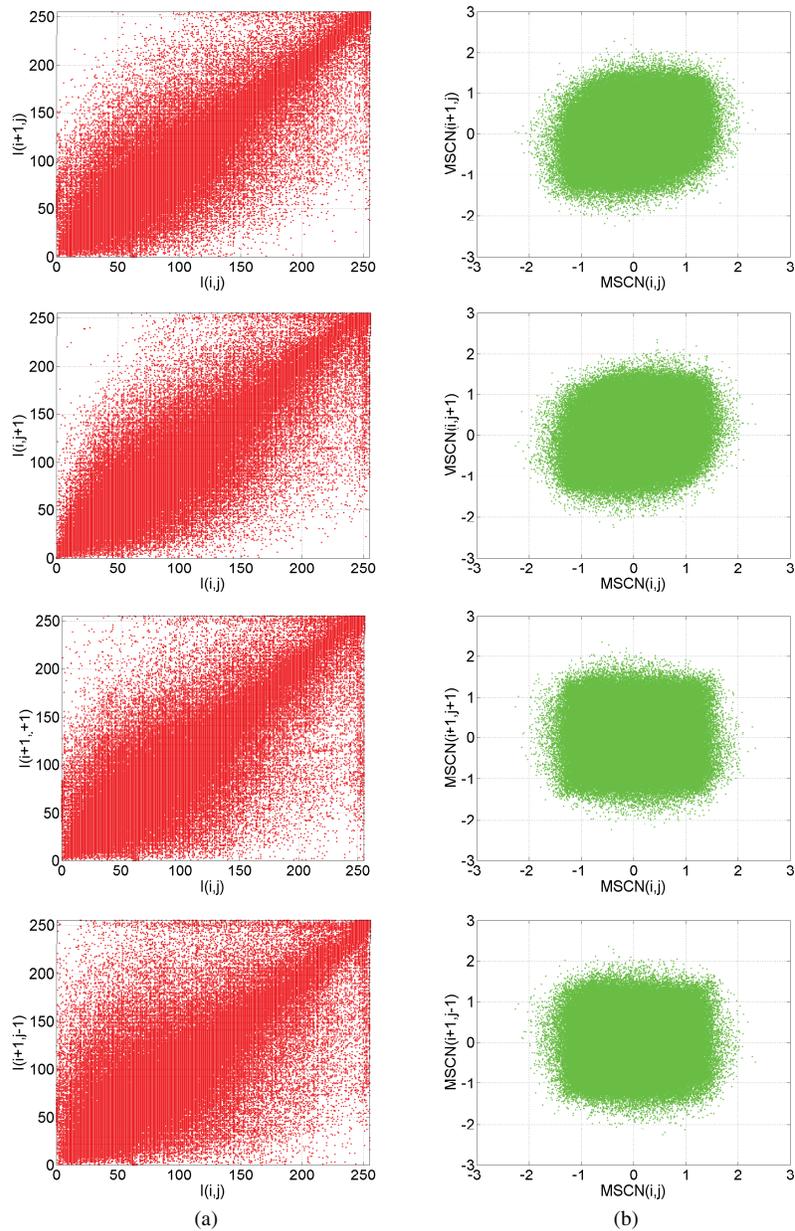
Fig. 2. Scatter plot between neighboring values of (a) Original luminance coefficients and (b) MSCN coefficients. Rows from top to bottom illustrate horizontal, vertical, main diagonal, and secondary diagonal neighbors. Notice a high correlation between surrounding pixels with a diagonal structure in the plots shown in (a). The normalization procedure greatly reduces these dependencies as is apparent in the plots shown in (b).

We also model the statistical relationships between neighboring pixels. While MSCN coefficients are definitely more homogenous for pristine images, the signs of adjacent coefficients also exhibit a regular structure, which gets disturbed in the presence of distortion. We model this structure using the empirical distributions of pairwise products of neighboring MSCN coefficients along four orientations – horizontal (H), vertical (V), main-diagonal ($D1$) and secondary-diagonal ($D2$), as illustrated in Fig. 6. Specifically,

$$H(i, j) = \hat{I}(i, j)\hat{I}(i, j + 1) \tag{7}$$

$$V(i, j) = \hat{I}(i, j)\hat{I}(i + 1, j) \tag{8}$$

$$D1(i, j) = \hat{I}(i, j)\hat{I}(i + 1, j + 1) \tag{9}$$

$$D2(i, j) = \hat{I}(i, j)\hat{I}(i + 1, j - 1) \tag{10}$$

for $i \in \{1, 2 \ldots M\}$ and $j \in \{1, 2 \ldots N\}$.

Under the Gaussian coefficient model, and assuming the MSCN coeffficients are zero mean and unit variance, these products obey the following distribution in the absence of distortion [34]:

$$f(x, \rho) = \frac{\exp\left(\frac{|x|\rho}{1 - \rho^2}\right) K_0\left(\frac{|x|}{1 - \rho^2}\right)}{\pi\sqrt{(1 - \rho^2)}} \tag{11}$$

where f is an asymmetric probability density function, ρ denotes the correlation coefficient of adjacent coefficents, and K_0 is the modified bessel function of the second kind.
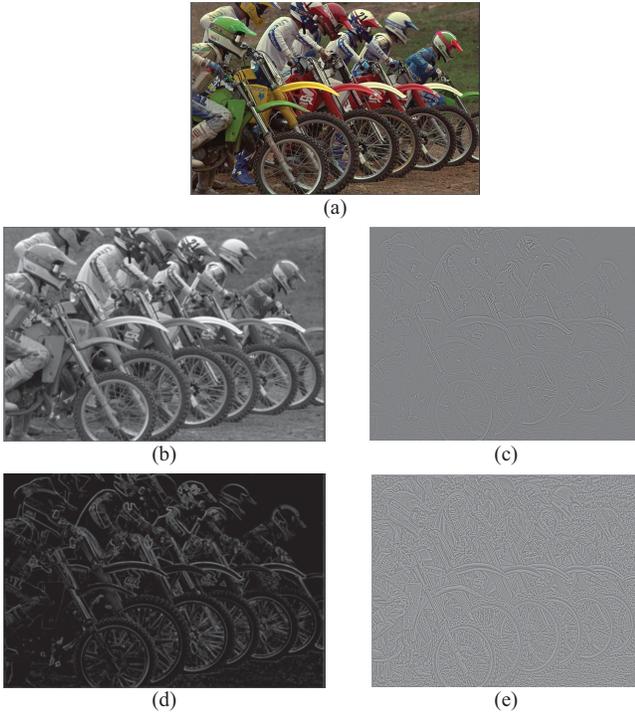
Fig. 3. Effect of the normalization procedure. (a) Original image I. (b) Local mean field μ. (c) $I - \mu$. (d) Local variance field σ. (e) MSCN coefficients $((I - \mu)/\sigma)$.
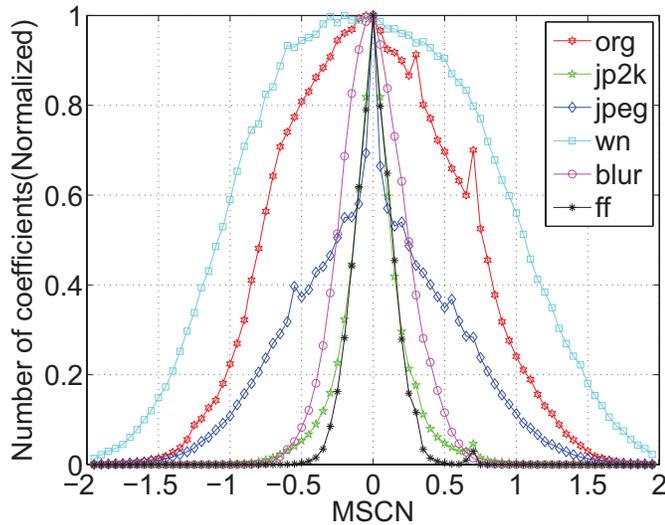

Fig. 4. Histogram of MSCN coefficients for a natural undistorted image and its various distorted versions. Distortions from the LIVE IQA database [27]. jp2k: JPEG2000. jpeg: JPEG compression. WN: additive white Gaussian noise. blur: Gaussian blur. ff: Rayleigh fast-fading channel simulation.

While we have found that this density function is a good model of the empirical histograms of products of adjacent normalized coefficients, it has only a single parameter, and as such, does not provide a good fit to the empirical histograms of coefficient products (Fig. 2) from *distorted* images. Further, it is not finite at the origin. Hence, as a practical alternative, we adopt the very general asymmetric generalized Gaussian distribution (AGGD) model [35]. In order to visualize how paired products

vary in the presence of distortion, in Fig. 7, we plot histograms of paired products along each of four orientations, for a reference image and for distorted versions of it.

The AGGD with zero mode is given by:

$$f(x; \nu, \sigma_l^2, \sigma_r^2) = \begin{cases} \dfrac{\nu}{(\beta_l + \beta_r)\Gamma\left(\frac{1}{\nu}\right)} \exp\left(-\left(\frac{-x}{\beta_l}\right)^\nu\right) & x < 0 \\[2em] \dfrac{\nu}{(\beta_l + \beta_r)\Gamma\left(\frac{1}{\nu}\right)} \exp\left(-\left(\frac{x}{\beta_r}\right)^\nu\right) & x \geq 0 \end{cases}$$

$$(12)$$

where

$$\beta_l = \sigma_l \sqrt{\frac{\Gamma\left(\frac{1}{\nu}\right)}{\Gamma\left(\frac{3}{\nu}\right)}} \qquad (13)$$

$$\beta_r = \sigma_r \sqrt{\frac{\Gamma\left(\frac{1}{\nu}\right)}{\Gamma\left(\frac{3}{\nu}\right)}} \qquad (14)$$

The shape parameter ν controls the 'shape' of the distribution while σ_l^2 and σ_r^2 are scale parameters that control the spread on each side of the mode, respectively. The AGGD further generalizes the generalized Gaussian distribution (GGD) [32] and subsumes it by allowing for asymmetry in the distribution. The skew of the distribution is a function of the left and right scale parameters. If $\sigma_l^2 = \sigma_r^2$, then the AGGD reduces to the GGD. Although the AGGD is infrequently used, it has been deployed to model skewed heavy-tailed distributions of image texture [35]. The parameters of the AGGD ($\nu, \sigma_l^2, \sigma_r^2$), are estimated using the moment-matching based approach proposed in [35]. Figure 5(b) shows the 3-D scatter plot between ($\nu, \sigma_l^2, \sigma_r^2$) for horizontal paired products using the same set of images as used for showing separation in GGD parameter space. It can be visualized that different distortions occupy different parts of the space. Also, we expect images to have a better separation when modeled in the high dimensional space of parameters obtained by fitting AGGD distributions to paired products from different orientations and scales together. This figure also motives the use of (12) to better capture the finite empirical density function.

The parameters ($\eta, \nu, \sigma_l^2, \sigma_r^2$) of the best AGGD fit are extracted where η is given by:

$$\eta = (\beta_r - \beta_l) \frac{\Gamma\left(\frac{2}{\nu}\right)}{\Gamma\left(\frac{1}{\nu}\right)}. \qquad (15)$$

Thus for each paired product, 16 parameters (4 parameters/orientation × 4 orientations) are computed, yielding the next set of features. Table I summarizes the features utilized.

Images are naturally multiscale, and distortions affect image structure across scales. Further, as research in quality assessment has demonstrated, incorporating multiscale information when assessing quality produces QA algorithms that perform better in terms of correlation with human perception [13], [29]. Hence, we extract all features listed in Table I at two scales - the original image scale, and at a reduced resolution (low pass filtered and downsampled by a factor of 2). We observed that increasing the number of scales beyond 2 did not contribute to performance much. Thus, a total of
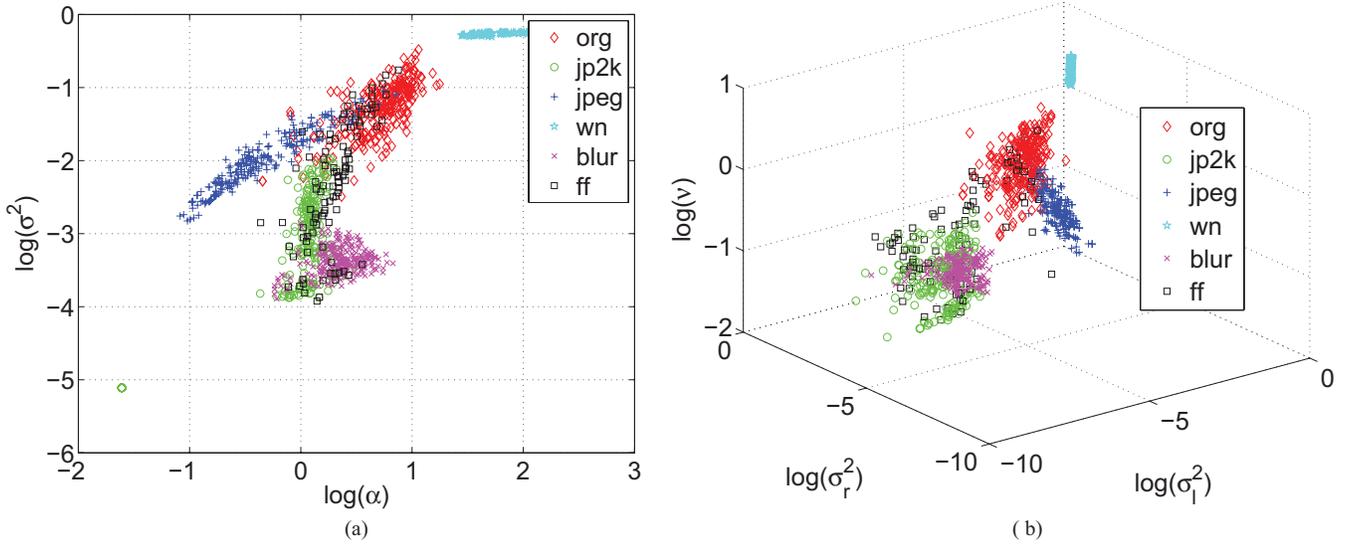
Fig. 5. (a) 2-D scatter plot between shape and scale parameters obtained by fitting GGD to the empirical distributions of MSCN coefficients of pristine images of Berkeley image segmentation database [33] and simulated distorted images, where similar kinds of distortions as those present in the LIVE IQA database [27] (JPEG 2000, JPEG, White Noise, Gaussian Blur, and Fast fading channel errors) were introduced in each image at varying degrees of severity. (b) 3-D scatter plot between shape, left scale, and right scale obtained by fitting AGGD to horizontal paired products using the same set of images as (a).

TABLE I

SUMMARY OF FEATURES EXTRACTED IN ORDER TO CLASSIFY AND QUANTIFY DISTORTIONS

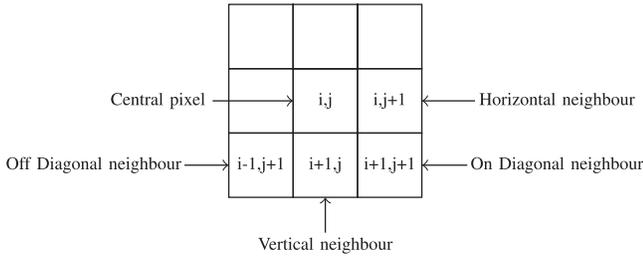| Feature ID	Feature Description	Computation Procedure
$f_1 - f_2$	Shape and variance	Fit GGD [32] to MSCN coefficients
$f_3 - f_6$	Shape, mean, left variance, right variance	Fit AGGD [35] to H pairwise products
$f_7 - f_{10}$	Shape, mean, left variance, right variance	Fit AGGD [35] to V pairwise products
$f_{11} - f_{14}$	Shape, mean, left variance, right variance	Fit AGGD [35] to $D1$ pairwise products
$f_{15} - f_{18}$	Shape, mean, left variance, right variance	Fit AGGD [35] to $D2$ pairwise products

Fig. 6. Various paired products computed in order to quantify neighboring statistical relationships. Pairwise products are computed along four orientations—horizontal, vertical, main-diagonal, and secondary-diagonal—at a distance of 1 pixel.

36 features – 18 at each scale, are used to identify distortions and to perform distortion-specific quality assessment. In Fig. 8, we plot the Spearman's rank ordered correlation coefficient (SROCC) between each of these features and human DMOS from the LIVE IQA database, for each of the distortions in the database – JPEG and JPEG2000 compression, additive white Gaussian noise, Gaussian blur and a Rayleigh fast fading channel distortion, to ascertain how well the features correlate with human judgments of quality. Note that no training is undertaken here, the plot is simply to illustrate that each feature captures quality information and to show that images are affected differently by different distortions.

B. Quality Evaluation

A mapping is learned from feature space to quality scores using a regression module, yielding a measure of image quality. The framework is generic enough to allow for the use of any regressor. In our implementation, a support vector machine (SVM) regressor (SVR) [36] is used. SVR has previously been applied to image quality assessment problems [12], [37], [38]. For example, a learning driven feature pooling approach using SVR was proposed in [38]. Wavelet-domain NSS and singular value decomposition features have been used to map quality to human ratings via SVR in [12] and [37] respectively. SVR is generally noted for being able to handle high dimensional data [39]. We utilize the LIBSVM package [40] to implement the SVR with a radial basis function (RBF) kernel.

IV. PERFORMANCE EVALUATION

A. Correlation With Human Opinions

We used the LIVE IQA database [27] to test the performance of BRISQUE, which consists of 29 reference images with 779 distorted images spanning five different distortion categories – JPEG2000 (JP2K) and JPEG compression, additive white Gaussian noise (WN), Gaussian blur (Blur), and a Rayleigh fast-fading channel simulation (FF). Each of the distorted images has an associated difference mean opinion
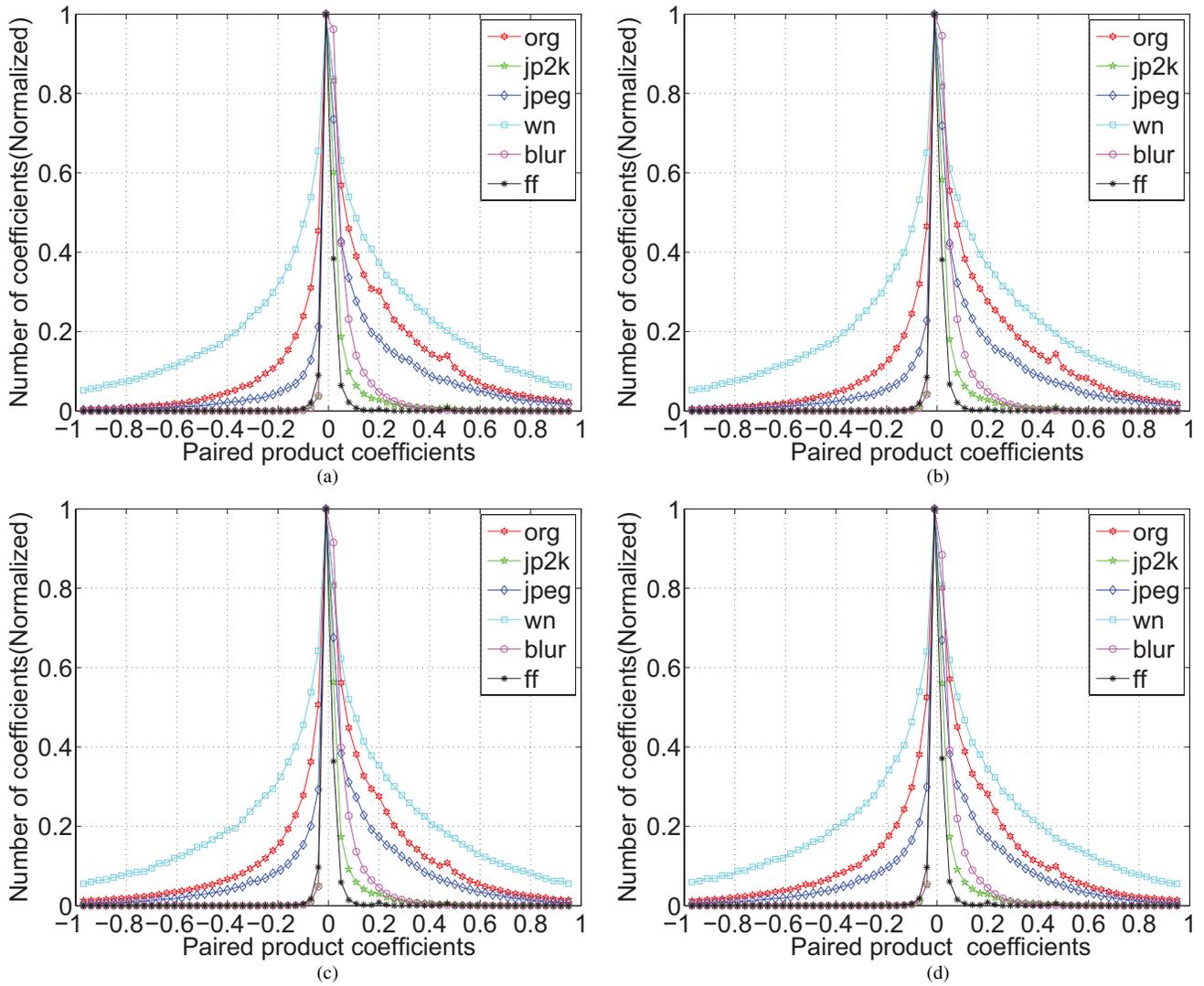
Fig. 7. Histograms of paired products of MSCN coefficients of a natural undistorted image and various distorted versions of it. (a) Horizontal. (b) Vertical. (c) Main-diagonal. (d) Secondary-diagonal. Distortions from the LIVE IQA database [27]. jp2k: JPEG2000. jpeg: JPEG compression. WN: additive white Gaussian noise. blur: Gaussian blur. ff: Rayleigh fast-fading channel simulation.

score (DMOS) which represents the subjective quality of the image.

Since the BRISQUE approach requires a training procedure to calibrate the regressor module, we divide the LIVE database into two randomly chosen subsets – 80% training and 20% testing – such that no overlap between train and test content occurs. We do this to ensure that the reported results do not depend on features extracted from known spatial content, which can artifically improve performance. Further, we repeat this random train-test procedure 1000 times and report the median of the performance across these 1000 iterations, in order to eliminate performance bias.

The Spearman's rank ordered correlation coefficient (SROCC) and Pearson's (linear) correlation coefficient (LCC) between the predicted score from the algorithm and DMOS were used to access QA performance. Before computing LCC, the algorithm scores were passed through a logistic non-linearity as described in [27]. A value close to 1 for SROCC

and LCC indicate good performance in terms of correlation with human opinion. These performance indices are tabulated in Tables II and III respectively[2].

We also tabulated the performance of three full-reference indices: peak-signal-to-noise ratio (PSNR), structural similarity index (SSIM) [28] and multi-scale structural similarity index (MS-SSIM) [29]. Although PSNR is a poor measure of perceptual quality, it is often used to benchmark for QA algorithms [41], [42]. The SSIM and MS-SSIM indices are popular owing to their performance and simplicity. We also include the performance of the previously summarized general purpose no-reference algorithms - CBIQ [20], LBIQ [21], BLIINDS-II [13] and DIIVINE index [12]. We requested quality scores from authors for CBIQ [20] and

---

[2]Further, note that due to randomness of the 1000 trials, there may be a slight discrepancy between results reported here and elsewhere, however, these differences in correlations are not statistically significant, and are simply an artifact of the random train-test sampling.
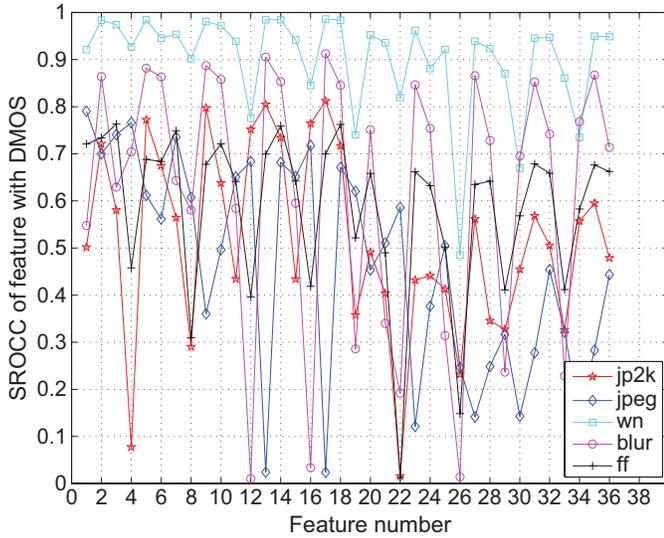
Fig. 8. Correlation of features with human judgments of quality (DMOS) for different distortions.

TABLE II

MEDIAN SPEARMAN RANK ORDERED CORRELATION COEFFICIENT (SROCC) ACROSS 1000 TRAIN-TEST COMBINATIONS ON THE LIVE IQA DATABASE. ITALICS INDICATE NO-REFERENCE ALGORITHMS

	JP2K	JPEG	WN	Blur	FF	All
PSNR	0.8646	0.8831	0.9410	0.7515	0.8736	0.8636
SSIM	0.9389	0.9466	0.9635	0.9046	0.9393	0.9129
MS-SSIM	0.9627	0.9785	0.9773	0.9542	0.9386	0.9535
CBIQ	*0.8935*	*0.9418*	*0.9582*	*0.9324*	*0.8727*	*0.8954*
LBIQ	*0.9040*	*0.9291*	*0.9702*	*0.8983*	*0.8222*	*0.9063*
BLIINDS-II	*0.9323*	*0.9331*	*0.9463*	*0.8912*	*0.8519*	*0.9124*
DIIVINE	*0.9123*	*0.9208*	*0.9818*	*0.9373*	*0.8694*	*0.9250*
Pointwise	*0.7957*	*0.8593*	*0.9608*	*0.8759*	*0.7773*	*0.8297*
Pairwise	*0.9007*	*0.9510*	*0.9773*	*0.8759*	*0.8741*	*0.9302*
BRISQUE	*0.9139*	*0.9647*	*0.9786*	*0.9511*	*0.8768*	*0.9395*

LBIQ [21]. Implementations of other indices are available online [43]–[46]. We also reported the correlations obtained by modeling empirical distributions of MSCN coefficients (pointwise) alone and pairwise products alone to compare their relative importance.

B. Variation With Window Size

As observed from the Table IV, the performance of BRISQUE remains relatively stable with respect to variation in the window size used to compute the local mean and variances. However, the performance starts to decrease when it becomes fairly large as the computations become non-local.

C. Statistical Significance and Hypothesis Testing

Figure 9 plots the mean SROCC across the 1000 trials and the standard deviations of performance across these 1000 trials for each of the algorithms considered here.

TABLE III

MEDIAN LINEAR CORRELATION COEFFICIENT ACROSS 1000 TRAIN-TEST COMBINATIONS ON THE LIVE IQA DATABASE. ITALICS INDICATE NO-REFERENCE ALGORITHMS

	JP2K	JPEG	WN	Blur	FF	All
PSNR	0.8762	0.9029	0.9173	0.7801	0.8795	0.8592
SSIM	0.9405	0.9462	0.9824	0.9004	0.9514	0.9066
MS-SSIM	0.9746	0.9793	0.9883	0.9645	0.9488	0.9511
CBIQ	*0.8898*	*0.9454*	*0.9533*	*0.9338*	*0.8951*	*0.8955*
LBIQ	*0.9103*	*0.9345*	*0.9761*	*0.9104*	*0.8382*	*0.9087*
BLIINDS-II	*0.9386*	*0.9426*	*0.9635*	*0.8994*	*0.8790*	*0.9164*
DIIVINE	*0.9233*	*0.9347*	*0.9867*	*0.9370*	*0.8916*	*0.9270*
Pointwise	*0.7947*	*0.8447*	*0.9711*	*0.8670*	*0.8151*	*0.8258*
Pairwise	*0.8968*	*0.9571*	*0.9830*	*0.8670*	*0.8952*	*0.9309*
BRISQUE	*0.9229*	*0.9734*	*0.9851*	*0.9506*	*0.9030*	*0.9424*

TABLE IV

MEDIAN SPEARMAN RANK ORDERED CORRELATION COEFFICIENT (SROCC) ACROSS 1000 TRAIN-TEST COMBINATIONS ON THE LIVE IQA DATABASE FOR DIFFERENT WINDOW SIZES. ITALICS INDICATE NO-REFERENCE ALGORITHMS

K,L	JPEG2000	JPEG	White noise	Gaussian Blur	Fast fading	Overall
4	0.9120	0.9581	0.9764	0.9535	0.8839	0.9388
5	0.9083	0.9510	0.9742	0.9497	0.8790	0.9360
6	0.9043	0.9483	0.9706	0.9417	0.8725	0.9309
7	0.9040	0.9482	0.9700	0.9407	0.8720	0.9305
8	0.8950	0.9405	0.9631	0.9321	0.8683	0.9208

Fig. 9. Mean SROCC and standard error bars for various algorithms across the 1000 train-test trials on LIVE IQA database.

Although there exist differences in the median correlations between the different algorithms (see Table II), these differences may not be statistically relevant. Hence, to evaluate the statistical significance of performance of each of the algorithms considered, we performed hypothesis testing based on the *t*-test [47] on the SROCC values obtained from the 1000 train-test trials, and we tabulated the results in Table V.

TABLE V

RESULTS OF ONE-SIDED T-TEST PERFORMED BETWEEN SROCC VALUES OF VARIOUS IQA ALGORITHMS. A VALUE OF "1" INDICATES THAT THE ROW ALGORITHM IS STATICALLY SUPERIOR TO THE COLUMN ALGORITHM; "−1" INDICATES THAT THE ROW IS WORSE THAN THE COLUMN; A VALUE OF "0" GIVES INDICATES THAT THE TWO ALGORITHMS ARE STATISTICALLY INDISTINGUISHABLE. ITALICS INDICATE NO-REFERENCE ALGORITHMS

	PSNR	SSIM	MSSSIM	CBIQ	LBIQ	BLIINDS-II	DIIVINE	BRISQUE
PSNR	0	−1	−1	−1	−1	1	−1	−1
SSIM	1	0	−1	1	1	1	−1	−1
MSSSIM	1	1	0	1	1	1	1	1
CBIQ	*1*	*−1*	*−1*	*0*	*−1*	*1*	*−1*	*−1*
LBIQ	*1*	*−1*	*−1*	*1*	*0*	*1*	*−1*	*−1*
BLIINDS-II	*1*	*−1*	*−1*	*1*	*1*	*0*	*−1*	*−1*
DIIVINE	*1*	*1*	*−1*	*1*	*1*	*1*	*0*	*−1*
BRISQUE	*1*	*1*	*−1*	*1*	*1*	*1*	*1*	*0*

TABLE VI

MEDIAN CLASSIFICATION ACCURACY ACROSS 1000 TRAIN-TEST TRIALS

	JP2kK	JPEG	WN	Blur	FF	All
Classification Accuracy (%)	82.9	88.9	100.0	96.7	83.3	88.6

Fig. 10. Mean confusion matrix for classifier across 1000 trials illustrates which row(distortion) is confused with which column (distortion). Higher number indicates greater confusion.

TABLE VII

MEDIAN SPEARMAN RANK ORDERED CORRELATION COEFFICIENT (SROCC) ACROSS 1000 TRAIN-TEST COMBINATIONS ON THE LIVE IQA DATABASE. ITALICS INDICATE NO-REFERENCE ALGORITHMS

	JP2K	JPEG	WN	Blur	FF	All
BRISQUE	*0.9139*	*0.9647*	*0.9786*	*0.9511*	*0.8768*	*0.9395*
BRISQUE (2-stage)	*0.8991*	*0.9439*	*0.9849*	*0.9479*	*0.8825*	*0.9315*

The null hypothesis is that the mean correlation for the (row) algorithm is equal to mean correlation for the (column) algorithm with a confidence of 95%. The alternate hypothesis is that the mean correlation of row is greater than or lesser than the mean correlation of the column. A value of '1' in the table indicates that the row algorithm is statically superior to the column algorithm, whereas a '−1' indicates that the row is statistically worse than the column. A value of '0' indicates that the row and column are statistically indistinguishable (or equivalent), i.e., we could not reject the null hypothesis at the 95% confidence level.

From Table V we conclude that BRISQUE is highly competitive with all no reference algorithms tested and statistically better than the full reference algorithms PSNR and SSIM. Given that these measures require additional information in the form of the reference image, this is by no means a small achievement. This result suggests that to the extent distortions can be trained on, one can replace full reference algorithms such as SSIM with the proposed BRISQUE without any loss of performance. We note that BRISQUE remains slightly inferior to the FR MS-SSIM, indicating that there may still be some room for improvement in performance.

D. Classification Accuracy

In order to demonstrate that BRISQUE features can also be used for explicit distortion-identification [48], we report the median classification accuracy of the classifier for each of the distortions in the LIVE database, as well as across all distortions in Table VI.

Further, in order to visualize which distortions are 'confused' the most, Fig. 10 plots the confusion matrix for each of the distortions, where the sum of each row in the confusion matrix is 1 and actual values represent the mean confusion percentage across the 1000 train-test trials. We see from Fig. 10 that FF and JP2K are most confused with each other which is not surprising, since FF distortion is a combination of JP2K followed by packet-loss errors. JP2K and JPEG are also confused sometimes. WN and Blur are generally not confused with other distortions.

E. Two-Stage Performance

We also investigated the possibility of replacing the one stage framework, where features are directly mapped to quality, with a two-stage framework, similar to that proposed in [48]. In this approach, the same set of features are used to identify the distortion afflicting the image as are then used for distortion-specific QA. Such a two-stage approach was used with recent success for NSS-based blind IQA [12]. In Table VII, we tabulate the median SROCC value across 1000 trials for the two-stage realization of BRISQUE. We also list the performances of BRISQUE for comparison purposes. The slight dip in the performance can be attributed to imperfect distortion identification in the first stage of the two-stage framework.

TABLE VIII

SPEARMAN'S RANK ORDERED CORRELATION COEFFICIENT (SROCC) ON THE TID2008 DATABASE. ITALICIZED ALGORITHMS ARE NR IQA ALGORITHMS, OTHERS ARE FR IQA ALGORITHMS

	JP2K	JPEG	WN	Gblur	All
PSNR	0.825	0.876	0.918	0.934	0.870
SSIM	0.963	0.935	0.817	0.960	0.902
BRISQUE	*0.832*	*0.924*	*0.829*	*0.881*	*0.896*

TABLE IX

INFORMAL COMPLEXITY ANALYSIS OF BRISQUE. TABULATED VALUES REFLECT THE PERCENTAGE OF TIME DEVOTED TO EACH OF THE STEPS IN BRISQUE

Step	Percentage of time
MSCN	50.9
GGD	8.6
Pairwise Products and AGGD	40.6

F. Database Independence

Having evaluated BRISQUE on the LIVE IQA database, we now demonstrate that the performance of BRISQUE is not bound by the database on which it is tested. To show this, we trained BRISQUE on the entire LIVE IQA database and then applied BRISQUE to the TID2008 database [4].

The TID database consists of 25 reference images and 1700 distorted images over 17 distortion categories [4]. Since there are only 24 natural images, and our algorithm is based on the statistics of natural images, we test our approach only on these 24 images. Further, although there exist 17 distortion categories, we tested BRISQUE only on these distortions that it is trained for: JPEG, JPEG2000 compression (JP2K), additive white noise (WN) and Gaussian Blur (blur) – FF distortion does not exist in the TID database. The results of applying BRISQUE on TID are tabulated in Table VIII, where we also list the performance of PSNR and SSIM for comparison purposes. It should be clear that BRISQUE performs well in terms of correlation with human perception of quality and that the performance does not depend on the database.

G. Computational Complexity

Our description of BRISQUE focused on the relationship of the statistical features to natural scene statistics and the effect that distortions have on such statistics. However, given the small number of features that are extracted (18 per scale) and the fact that parameter estimation needs to be performed only 5 times for an entire image, in comparison to parameter estimation for each block as in BLIINDS-II [13], the reader will appreciate the fact that BRISQUE is extremely efficient. Having demonstrated that BRISQUE performs well in terms of correlation with human perception, we also now show that BRISQUE has low complexity. In Table IX we list the relative percentage of time each of the stages of BRISQUE uses as a percentage of the time taken to compute the quality of an image (once trained).

We also compare the overall computational complexity of BRISQUE with the FR PSNR and the NR BLIINDS-II

TABLE X

COMPLEXITY ANALYSIS OF BRISQUE: A COMPARISON OF THE AMOUNT OF TIME TAKEN TO COMPUTE VARIOUS QUALITY MEASURES FOR A 512×768 IMAGE ON A 1.8-GHz SINGLE-CORE PC WITH 2 GB OF RAM

Algorithm	Time (seconds)
PSNR	0.05
DIIVINE	149
BLIINDS-II	70
BRISQUE	1

and DIIVINE, and in Table X, we list the time taken (in seconds) to compute each quality measure on an image of resolution 512×768 on a 1.8 Ghz single-core PC with 2 GB of RAM. We use unoptimized MATLAB code for all of these algorithms in order to ensure a fair comparison. We also list the efficiency as a fraction of the time taken to compute PSNR, to allow for a machine-independent comparison across algorithms. As Table X demonstrates, BRISQUE is quite efficient, outperforming the DIIVINE index and the BLIINDS-II index by a large amount. This suggests that the spatial-domain BRISQUE an ideal candidate for real-time blind assessment of visual quality.

V. APPLICATION TO BLIND IMAGE DENOISING

The computational efficiency and excellent quality prediction performance makes BRISQUE an attractive option for practical applications. One such application could be using a quality measure to augment the performance of image repair algorithms. In this section, we describe one such approach, where the BRISQUE features are used to transform a non-blind image denoising algorithm into a blind image denoising algorithm.

Blind image denoising algorithms seek to reduce the amount of noise present in corrupted images, without any additional information such as the noise variance. Although image denoising is a well studied problem in image processing [50]–[54], blind image denoising remains relatively underexplored [49], [54]. The proposed algorithms typically address parameter estimation in an *ad-hoc* fashion without regard to natural scene statistics. Here, we demonstrate a systematic perception-based parameter estimation approach that results in better denoising performance. We augment a state-of-the-art image denoising algorithm by using BRISQUE feature-based parameter prediction to improve performance.

The work closest in concept to this approach is the one proposed in [49] where image content measures were used to predict the noise variance in the image, which was then used for image denoising; however the approach is computationally intensive and the measure of content in the image may not be the ideal measure to predict noise variance. In [49], the noisy image is denoised multiple times and quality is estimated using their proposed no-reference content evaluation algorithm. Amongst the large set of denoised images produced, the image with the best content-quality is selected as the denoised image. As an alternative, we propose a learning based
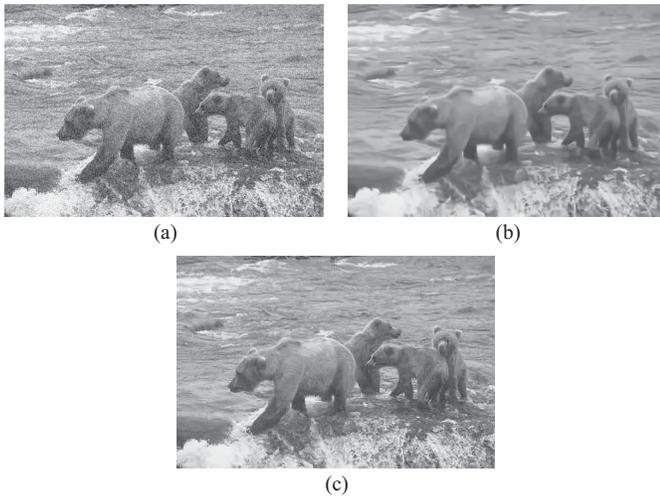
Fig. 11. Accurate noise variance as input to the algorithm in [49] produces poorer quality denoised images: (a) Noisy image ($\sigma = 0.0158$, MS-SSIM = 0.9063), (b) denoised with $\sigma = 0.0158$ (MS-SSIM = 0.9176), and (c) denoised with $\sigma = 0.0040$ (MS-SSIM = 0.9480).
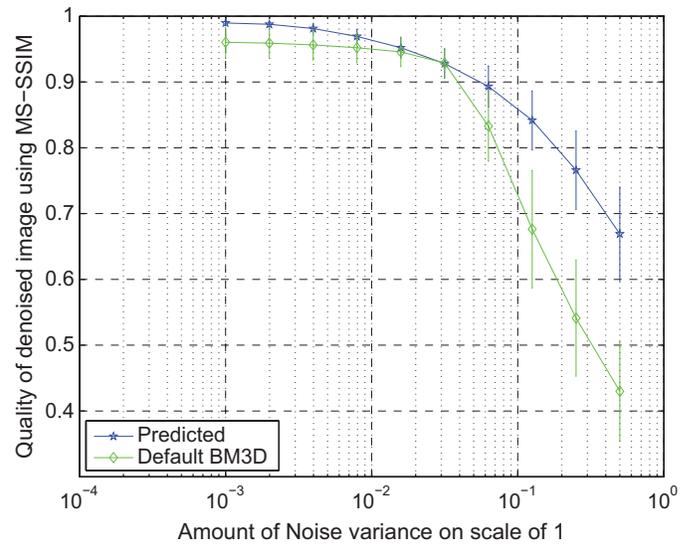

Fig. 12. Mean quality and associated errors at each noise level across 2000 test images for our approach as well as the reference implementation of BM3D.

framework where noise parameters are estimated using natural scene statistics based on BRISQUE features.

The denoising algorithm that we use is the one proposed in [54], which requires as input the noise variance in the image. However, our experiments suggest that when the algorithm is fed the true accurate noise-variance, the performance of the denoiser is sub-par. The performance of the algorithm drastically improves, if a (systematically) different parameter selected based on perceptual quality is fed as input to the algorithm. In order to demonstrate this, in Fig. 11, we plot an image denoised using the true noise variance and that arrived at using the noise variance from our approach (described below). Notice that our approach produces better visual quality, and better objective quality, as gauged by the multi-scale structural similarity index (MS-SSIM) [29].

We design our training framework to account for this discrepency and to ensure that the denoised image attains the highest visual quality. Our approach proceeds as follows. Given a large set of noisy images afflicted with different levels of noise, we denoise each image using the denoising algorithm – BM3D [54] – by providing as input images distorted with various values of noise variance. The denoised images so obtained are judged for their quality using MS-SSIM and the noise parameter corresponding to the image with the maximum denoised quality is set as the input to the algorithm. These noise variances are then used in a training phase, where BRISQUE features are mapped on to the noise-prediction parameter, using SVM regression as before [40]. Once trained, the automatic parameter prediction approach is capable of predicting the level of input noise to BM3D, so that the output denoised image has the highest visual quality. We note that our training approach resembles that of [49].

Given a new (unseen) test noisy image, the BRISQUE augmented BM3D approach predicts the accurate input to BM3D and denoises the image with (as we shall soon see) much higher visual quality than the baseline. Notice that BRISQUE augmentation is not limited to the BM3D algorithm; and any

non-blind algorithm could be improved by using BRISQUE natural scene features to produce a blind image denoiser.

To show the effectiveness of our algorithm and to demonstrate its robustness across a large variety of images and distortion levels, we created a noisy image dataset from the 300 images present in the Berkeley image segmentation database [33]. We introduced 10 different levels of Gaussian noise to each image yielding a total of 3000 noisy images. The noise variance ranged from 0.001 to 0.5, uniformly sampled on a logarithmic scale. 1000 images were then used for training and 2000 for testing thereby ensuring no content overlap between the two sets. The regression model described above was trained on 1000 training images and then used to predict the input parameter on the test images.

Once denoised images are obtained, we compare their quality (using MS-SSIM) using our approach as well for the default implementation of the BM3D algorithm and in Fig. 12, we plot the mean quality and the associated standard errors at each noise level across the 2000 test images for both these approaches. It is clear that BRISQUE augmented BM3D produces much higher quality images than the baseline BM3D. We also analyzed whether the differences observed in the quality of the denoised images between our approach and the reference BM3D implementation are statistically significant using the *t*-test [55]. Our analysis indicates that for all noise variances simulated in the present data, our approach is *statistically superior* to the reference BM3D implementation in terms of perceived visual quality at the 95% confidence level, excepting when the noise variance is a tiny 0.0316 - where the two approaches become statistically indistinguishable.

VI. CONCLUSION

We proposed a natural scene statistic based distortion-generic blind/no-reference (NR) quality assessment algorithm– the Blind/Referenceless Image Spatial QUality Evaluator (BRISQUE) – which operates in the spatial domain. No

distortion specific features such as ringing, blur or blocking were modeled in the algorithm. The algorithm only quantifies the 'naturalness' (or lack thereof) in the image due to presence of distortions.

We detailed the algorithm and the statistical features extracted, and demonstrated how each of these features correlate with human perception. We then undertook a thorough evaluation of the BRISQUE index in terms of correlation with human perception and demonstrated that BRISQUE is statistically better than FR PSNR and SSIM as well as highly competitive to all NR algorithms compared with. We demonstrated that BRISQUE performance is independent of database content and BRISQUE features may be used for distortion-identification as well. Further, we also showed that BRISQUE is computationally efficient and that its efficiency is superior to other distortion-generic approaches to NR IQA, thus making BRISQUE an attractive option for practical applications like image denoising. We demonstrated this application by augmenting non-blind image denoising algorithms using the BRISQUE features to produce blind image denoising algorithms.

REFERENCES

[1] A. C. Bovik, *Handbook of Image and Video Processing*. New York: Academic, 2005.

[2] 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.

[3] 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. 2011.

[4] 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. 10, pp. 30–45, 2009.

[5] 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.

[6] N. D. Narvekar and L. J. Karam, "A no-reference perceptual image sharpness metric based on a cumulative probability of blur detection," in *Proc. IEEE Int. Workshop Qual. Multimedia Experience*, Jul. 2009, pp. 87–91.

[7] S. Varadarajan and L. J. Karam, "An improved perception-based no-reference objective image sharpness metric using iterative edge refinement," in *Proc. IEEE Int. Conf. Image Process.*, Oct. 2008, pp. 401–404.

[8] N. G. Sadaka, L. J. Karam, R. Ferzli, and G. P. Abousleman, "A no-reference perceptual image sharpness metric based on saliency-weighted foveal pooling," in *Proc. IEEE Int. Conf. Image Process.*, Oct. 2008, pp. 369–372.

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

[10] J. Chen, Y. Zhang, L. Liang, S. Ma, R. Wang, and W. Gao, "A no-reference blocking artifacts metric using selective gradient and plainness measures," in *Proc. Pacific Rim Conf. Multimedia, Adv. Multimedia Inf. Process.*, Nov. 2008, pp. 894–897.

[11] S. Suthaharan, "No-reference visually significant blocking artifact metric for natural scene images," *J. Signal Process.*, vol. 89, no. 8, pp. 1647–1652, 2009.

[12] A. K. Moorthy and A. C. Bovik, "Blind image quality assessment: From natural scene statistics to perceptual quality," *IEEE Trans. Image Process.*, vol. 20, no. 12, pp. 3350–3364, Dec. 2011.

[13] M. 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.

[14] A. C. Bovik, "Perceptual image processing: Seeing the future," *Proc. IEEE*, vol. 98, no. 11, pp. 1799–1803, Nov. 2010.

[15] D. L. Ruderman, "The statistics of natural images," *Netw. Comput. Neural Syst.*, vol. 5, no. 4, pp. 517–548, 1994.

[16] A. Srivastava, A. B. Lee, E. P. Simoncelli, and S. C. Zhu, "On advances in statistical modeling of natural images," *J. Math. Imag. Vis.*, vol. 18, no. 1, pp. 17–33, 2003.

[17] R. Barland and A. Saadane, "Reference free quality metric using a region-based attention model for JPEG-2000 compressed images," *Proc. SPIE*, vol. 6059, pp. 605905-1–605905-10, Jan. 2006.

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

[19] S. Gabarda and G. Cristóbal, "Blind image quality assessment through anisotropy," *J. Opt. Soc. Amer.*, vol. 24, no. 12, pp. 42–51, 2007.

[20] P. Ye and D. Doermann, "No-reference image quality assessment using visual codebooks," in *Proc. IEEE Int. Conf. Image Process.*, Jul. 2011, pp. 3129–3138.

[21] H. Tang, N. Joshi, and A. Kapoor, "Learning a blind measure of perceptual image quality," in *Proc. Int. Conf. Comput. Vis. Pattern Recognit.*, Jun. 2011, pp. 305–312.

[22] J. Shen, Q. Li, and G. Erlebacher, "Hybrid no-reference natural image quality assessment of noisy, blurry, JPEG2000, and JPEG images," *IEEE Trans. Image Process.*, vol. 20, no. 8, pp. 2089–2098, Aug. 2011.

[23] 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.

[24] H. R. Sheikh, A. C. Bovik, and G. De 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.

[25] A. K. Moorthy and A. C. Bovik, "A two-step framework for constructing blind image quality indices," *IEEE Signal Process. Lett.*, vol. 17, no. 5, pp. 513–516, May 2010.

[26] Z. Wang and A. C. Bovik, "Reduced- and no-reference image quality assessment," *IEEE Signal Process. Mag.*, vol. 28, no. 6, pp. 29–40, Nov. 2011.

[27] 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.

[28] 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.

[29] Z. Wang, E. P. Simoncelli, and A. C. Bovik, "Multiscale structural similarity for image quality assessment," in *Proc. Asilomar Conf. Signals, Syst. Comput.*, vol. 2. 2003, pp. 1398–1402.

[30] M. Carandini, D. J. Heeger, and J. A. Movshon, "Linearity and normalization in simple cells of the macaque primary visual cortex," *J. Neurosci.*, vol. 17, no. 21, pp. 8621–8644, 1997.

[31] M. J. Wainwright, O. Schwartz, and E. P. Simoncelli, "Natural image statistics and divisive normalization: Modeling nonlinearities and adaptation in cortical neurons," in *Statistical Theories of the Brain*. Cambridge, MA: MIT Press, 2002, pp. 203–222.

[32] K. Sharifi and A. Leon-Garcia, "Estimation of shape parameter for generalized Gaussian distributions in subband decompositions of video," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 5, no. 1, pp. 52–56, Feb. 1995.

[33] D. Martin, C. Fowlkes, D. Tal, and J. Malik, "A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics," in *Proc. 8th Int. Conf. Comput. Vis.*, vol. 2. 2001, pp. 416–423.

[34] A. H. Nuttall, "Accurate efficient evaluation of cumulative or exceedance probability distributions directly from characteristic functions," Naval Underwater Systems Center, New London, CT, Tech. Rep. AD-A133703, 1983.

[35] N. E. Lasmar, Y. Stitou, and Y. Berthoumieu, "Multiscale skewed heavy tailed model for texture analysis," in *Proc. IEEE Int. Conf. Image Process.*, Nov. 2009, pp. 2281–2284.

[36] B. Schölkopf, A. J. Smola, R. C. Williamson, and P. L. Bartlett, "New support vector algorithms," *Neural Comput.*, vol. 12, no. 5, pp. 1207–1245, 2000.

[37] M. Narwaria and W. Lin, "Objective image quality assessment based on support vector regression," *IEEE Trans. Neural Netw.*, vol. 21, no. 3, pp. 515–519, Mar. 2010.

[38] M. Narwaria and W. Lin, "SVD-based quality metric for image and video using machine learning," *IEEE Trans. Syst., Man, Cybern., B, Cybern.*, vol. 42, no. 2, pp. 347–364, Apr. 2012.

[39] C. J. C. Burges, "A tutorial on support vector machines for pattern recognition," *Data Mining Knowl. Discovery*, vol. 2, no. 2, pp. 121–167, 1998.

[40] C. C. Chang and C. J. Lin. (2001). *LIBSVM: A Library for Support Vector Machines* [Online]. Available: http://www.csie.ntu.edu.tw/~cjlin/libsvm/

[41] Z. Wang and A. C. 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.

[42] B. Girod, "What's wrong with mean squared error?" in *Visual Factors of Electronic Image Communications*, A. B. Watson Ed. Cambridge, MA: MIT Press, 1993, pp. 207–220.

[43] Z. Wang, A. C. Bovik, E. P. Sheikh, and H. R. Simoncelli. (2004). *SSIM Code* [Online]. Available: http://live.ece.utexas.edu/research/Quality/index.htm

[44] Z. Wang, E. P. Simoncelli, and A. C. Bovik. (2003). *MS-SSIM Code* [Online]. Available: http://live.ece.utexas.edu/research/Quality/index.htm

[45] A. K. Moorthy and A. C. Bovik. (2011). *DIIVINE Code Implementation* [Online]. Available: http://live.ece.utexas.edu/research/quality/DIIVINE_release.zip

[46] M. A. Saad and A. C. Bovik. (2012). *BLIINDS Code* [Online]. Available: http://live.ece.utexas.edu/research/Quality/bliinds2_release.zip

[47] D. Sheskin, *Handbook of Parametric and Nonparametric Statistical Procedures*. London, U.K.: Chapman & Hall, 2004.

[48] A. K. Moorthy and A. C. Bovik, "A two-stage framework for blind image quality assessment," in *Proc. IEEE Int. Conf. Image Process.*, Sep. 2010, pp. 2481–2484.

[49] X. Zhu and P. Milanfar, "Automatic parameter selection for denoising algorithms using a no-reference measure of image content," *IEEE Trans. Image Process.*, vol. 19, no. 12, pp. 3116–3132, Dec. 2010.

[50] J. Portilla, V. Strela, M. J. Wainwright, and E. P. Simoncelli, "Image denoising using scale mixtures of Gaussians in the wavelet domain," *IEEE Trans. Image Process.*, vol. 12, no. 11, pp. 1338–1351, Nov. 2003.

[51] A. Buades, B. Coll, and J. M. Morel, "A non-local algorithm for image denoising," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, vol. 2. Jun. 2005, pp. 60–65.

[52] M. Elad and M. Aharon, "Image denoising via sparse and redundant representations over learned dictionaries," *IEEE Trans. Image Process.*, vol. 15, no. 12, pp. 3736–3745, Dec. 2006.

[53] F. Luisier, T. Blu, and M. Unser, "A new SURE approach to image denoising: Interscale orthonormal wavelet thresholding," *IEEE Trans. Image Process.*, vol. 16, no. 3, pp. 593–606, Mar. 2007.

[54] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, "Image denoising by sparse 3-D transform-domain collaborative filtering," *IEEE Trans. Image Process.*, vol. 16, no. 8, pp. 2080–2095, Aug. 2007.

[55] L. V. Hedges and I. Olkin, *Statistical Methods for Meta-Analysis*. New York: Academic, 1985.

Anish Mittal received the B.Tech. degree in electrical engineering from the Indian Institute of Technology Roorkee, Roorkee, India, and the M.S. degree in computer engineering from The University of Texas at Austin, Austin, in 2009 and 2011, respectively.

He joined the Laboratory for Image and Video Engineering, The University of Texas at Austin, in 2009, where he has been an Assistant Director since January 2012. His research interests include image and video processing, machine learning, and computer vision.

Mr. Mittal was a recipient of the Professional Development Award from The University of Texas at Austin in 2011.

Anush Krishna Moorthy received the B.E. degree in electronics and telecommunication from the University of Pune, Pune, India, in 2007, and the M.S. degree in electrical engineering and the Ph.D. degree from The University of Texas at Austin, Austin, in 2009 and 2012, respectively.

He is currently an Advanced Imaging Engineer at Texas Instruments, Dallas, Texas. He joined the Laboratory for Image and Video Engineering (LIVE), The University of Texas at Austin, in 2008, where he was an Assistant Director from 2008 to 2011.

Dr. Moorthy was a recipient of a Silver Medal from the University of Pune in 2007, the Continuing Graduate Fellowship for 2010-2011, the Professional Development Award in 2009 and 2010, the Center for Perceptual Systems Travel Grant from The University of Texas at Austin in 2010, and the TATA Scholarship for Higher Education Abroad.

Alan Conrad Bovik is the Curry/Cullen Trust Endowed Chair Professor with The University of Texas at Austin, Austin, where he is also the Director of the Laboratory for Image and Video Engineering (LIVE) and a Faculty Member in the Department of Electrical and Computer Engineering and the Center for Perceptual Systems, Institute for Neuroscience. He has authored or co-authored more than 650 technical articles and holds two U.S. patents. His several books include the recent companion volumes *The Essential Guides to Image and Video Processing* (Academic Press, 2009). His current research interests include image and video processing, computational vision, and visual perception.

He was a recipient of a Journal Paper Award from the International Pattern Recognition Society in 1988 and 1993, the IEEE Third Millennium Medal in 2000, the Hocott Award for Distinguished Engineering Research from the University of Texas at Austin, the Distinguished Alumni Award from the University of Illinois at Champaign-Urbana in 2008, the SPIE/IS&T Imaging Scientist Award in 2011, and a number of major awards from the IEEE Signal Processing Society, including the Best Paper Award in 2009, the Education Award in 2007, the Technical Achievement Award in 2005, and the Meritorious Service Award in 1998. He is a fellow of the IEEE, the Optical Society of America, the Society of Photo-Optical and Instrumentation Engineers, and the American Institute of Medical and Biomedical Engineering. He was on the Board of Governors of the IEEE Signal Processing Society from 1996 to 1998, and a Co-Founder and the Editor-in-Chief of the IEEE TRANSACTIONS ON IMAGE PROCESSING from 1996 to 2002. He was on the Editorial Board of The Proceedings of the IEEE from 1998 to 2004, and he was the Founding General Chairman of the First IEEE International Conference on Image Processing, Austin, Texas, in November, 1994. He has been a Series Editor of *Image, Video, and Multimedia Processing* since 2003. He is a Registered Professional Engineer in the State of Texas and a frequent consultant to legal, industrial, and academic institutions.