

RESEARCH

Open Access

A no-reference objective image quality metric based on perceptually weighted local noise

Tong Zhu and Lina Karam*

Abstract

This work proposes a perceptual based no-reference objective image quality metric by integrating perceptually weighted local noise into a probability summation model. Unlike existing objective metrics, the proposed no-reference metric is able to predict the relative amount of noise perceived in images with different content, without a reference. Results are reported on both the LIVE and TID2008 databases. The proposed no-reference metric achieves consistently a good performance across noise types and across databases as compared to many of the best very recent no-reference quality metrics. The proposed metric is able to predict with high accuracy the relative amount of perceived noise in images of different content.

Introduction

Reliable assessment of image quality plays an important role in meeting the promised quality of service (QoS) and in improving the end user's quality of experience (QoE). There is a growing interest to develop objective quality assessment algorithms that can predict perceived image quality automatically. These methods are highly useful in various image processing applications, such as image compression, transmission, restoration, enhancement, and display. For example, the quality metric can be used to evaluate and control the performance of individual system components in image/video processing and transmission systems.

One direct way to evaluate video quality is through subjective tests. In these tests, a group of human subjects are asked to judge the quality under a predefined viewing condition. The scores given by observers are averaged to produce the mean opinion score (MOS). However, subjective tests are time-consuming, laborious, and expensive. Objective image quality (IQA) assessment methods can be categorized as full reference (FR), reduced reference (RR), and no reference (NR) depending on whether a reference, partial information about a reference, or no reference is used for calculation. Quality assessment without a reference is challenging. A no-reference metric is not relative to a reference image, but rather an absolute value

is computed based on some characteristics of the test image.

Of particular interest to this work is the no-reference noisiness objective metric. Noisiness and blurriness are two key distortions in multiple applications, and typically there is a tradeoff to balance between noisiness and blurriness. For example, in soft-thresholding for image denoising [1], the image could be blurry when the threshold is high, while the image could remain noisy when the threshold is low. Also, in Wiener-based super-resolution [2], too much regularization will result in less noise at the expense of more blur. The reconstructed image could be blurry when the auto-correlation function is modeled to be too flat, while the reconstructed image could be noisy when the auto-correlation function is modeled to be too sharp. No-reference image sharpness/blur metrics have been widely discussed [3,4]. However, these image sharpness/blur metrics typically fail in the presence of noise. The sharpness metric may increase when noise increases. A no-reference noise-immune image sharpness metric was also proposed [5]. Furthermore, all the edge-based sharpness metrics can be easily applied in the wavelet domain as described in [5] to provide resilience to noise. Still, it lacks the ability to assess the impairment due to noise. For visual quality assessment of noisiness, many full-reference metrics are presented in [6], such as peak signal-to-noise ratio (PSNR), multi-scale structural similarity (MS-SSIM) [7], noise quality measure (NQM) [8], and information fidelity criterion (IFC) [9].

*Correspondence: karam@asu.edu
School of Electrical, Computer & Energy Engineering, Arizona State University,
Tempe, AZ 85287, USA

However, these full-reference metrics require the reference image for calculation. There is a need to develop a no-reference noisiness quality metric. Furthermore, such noisiness metric could be further combined with the no-reference blur metrics [3,4] to provide a better prediction of image quality for several applications including super-resolution, image restoration, and other multiply distorted images. A global estimate of image noise variance was used as a no-reference noisiness metric in [10]. The histogram of the local noise variances is used to derive the global estimate. However, the locally perceived visibility of noise is not considered. Similarly in [11], noisiness is expressed by the sum of estimated noise amplitudes and the ratio of noise pixels. Both the metrics of [10,11] do not account for the effects of locally varying noise on the perceived noise impairment and they do not exploit the characteristics of the human visual system (HVS).

To tackle this issue, this paper firstly presents a full-reference image noisiness metric which integrates perceptually weighted local noise into a probability summation model. This proposed metric can predict the perceptual noisiness in images with high accuracy. In addition, a no-reference objective noisiness metric is derived based on local noise standard deviation, local perceptual weighting, and probability summation. The experimental results show that the proposed FR and NR metrics show better and more consistent performance across databases and distortion types, when compared with several very recent FR and NR metrics.

The remainder of this paper is organized as follows. A perceived noisiness model based on probability summation is presented first followed by details on the contrast sensitivity thresholds computation. A full-reference perceptually weighted noise (FR-PWN) metric is proposed next based on perceptual weighting using the computed contrast sensitivity thresholds and probability summation. After that, a no-reference perceptually weighted noise (NR-PWN) metric is further derived. Performance results and comparison with existing metrics are presented followed by a conclusion.

Perceptual noisiness model based on probability summation

The PSNR simply calculates the difference point by point. However, the human visual system should be taken into consideration since the visual impairment due to the same noise could be perceived differently based on the local characteristics of the visual content. Contrast is a key concept in vision science because the information in the visual system is represented in terms of contrast and not in terms of the absolute level of light. So, the relative changes in luminance are important rather than the absolute ones [3]. The contrast sensitivity threshold measures the smallest contrast or the just-noticeable difference (JND) that yields

a visible signal over a uniform background. The proposed metric makes use of JND for calculating the probability of noise detection. Even when the noise is uniform, the impact of the noise will be more visible in image regions with a relatively lower JND. Consider the noisy signal y as

$$y(i, j) = y'(i, j) + \text{error}(i, j) \quad (1)$$

where $y'(i, j)$ is the original undistorted image. The probability of detecting a noise distortion at location (i, j) can be modeled as an exponential having the following form

$$P(i, j) = 1 - \exp\left(-\left|\frac{\text{error}(i, j)}{\text{JND}(i, j)}\right|^\beta\right) \quad (2)$$

where $\text{JND}(i, j)$ is the JND value at (i, j) and it depends on the mean intensity in a local neighborhood region surrounding pixel (i, j) . β is a parameter whose value is chosen to maximize the correspondence of (2) with the experimentally determined psychometric function for noise detection. In psychophysical experiments that examine summation over space, a value of about 4 has been observed to correspond well to probability summation [12].

A less-localized probability of noise detection can be computed by adopting the ‘probability summation’ hypothesis which pools the localized detection probabilities over a region of interest, R [13]. The probability summation hypothesis is based on the following two assumptions: (1) A noise distortion is detected if and only if at least one detector senses the presence of a noise distortion; (2) The probabilities of detection are independent; i.e., the probability that a particular detector will signal the presence of a distortion is independent of the probability that any other detector will. The measurement of noise detection in a region R is then given by

$$P_{\text{noise}}(R) = 1 - \prod_{i, j \in R} (1 - P(i, j)). \quad (3)$$

Substituting (2) into (3) yields

$$P_{\text{noise}}(R) = 1 - \exp(-D_R^\beta) \quad (4)$$

where

$$D_R = \left(\sum_{i, j \in R} \left|\frac{\text{error}(i, j)}{\text{JND}(i, j)}\right|^\beta\right)^{1/\beta} \quad (5)$$

From (4), it can be seen that $P_{\text{noise}}(R)$ increases if D_R increases and vice versa. So D_R can be used as a noisiness metric over region R . However, the probability of noise detection does not directly translate to noise annoyance level. In this work, the β parameter in (4) and (5) is

replaced with $\alpha = \beta \times s$, which has the effect of steering the slope of the psychometric function in order to translate noise detection levels into noise annoyance levels. The factor s was found experimentally to be 1/16 resulting in a value of 0.25 for α . More details about how JND(i, j) is computed is given in the Section 'Perceptual contrast sensitivity threshold model and JND computation'.

Perceptual contrast sensitivity threshold model and JND computation

Multiple parameters including screen resolution, the viewing distance, the minimum display luminance, and the maximum display luminance are considered in the contrast sensitivity model [14]. The thresholds are computed locally for each block. Firstly, the contrast sensitivity threshold t_{128} is generated for a region with a mean grayscale value of 128 as follows:

$$t_{128} = \frac{TM_g}{L_{\max} - L_{\min}} \quad (6)$$

where L_{\min} and L_{\max} are the minimum and maximum display luminances, M_g is the total number of gray scale levels, and T is given by the following parabolic approximation [15]:

$$T = \min(10^{g_{0,1}}, 10^{g_{1,0}}), \quad (7)$$

$$g_{0,1} = \log_{10} T_{\min} + K \left(\log_{10} \frac{1}{2N\omega_y} - \log_{10} f_{\min} \right)^2, \quad (8)$$

$$g_{1,0} = \log_{10} T_{\min} + K \left(\log_{10} \frac{1}{2N\omega_x} - \log_{10} f_{\min} \right)^2. \quad (9)$$

In (8) and (9), T_{\min} is the luminance threshold at frequency, f_{\min} , where the threshold is minimum. ω_x and ω_y represent, respectively, the horizontal width and the vertical height of a pixel in degrees of visual angle, K is the steepness of the parabola. N is the local neighborhood size and is set to 8. T_{\min} , f_{\min} , and K can be computed as [15]:

$$T_{\min} = \left\{ \begin{array}{l} \frac{L_T}{S_0} \left(\frac{L}{L_T} \right)^{\alpha_T}, L \leq L_T \\ \frac{L}{S_0}, L > L_T \end{array} \right\} \quad (10)$$

$$f_{\min} = \left\{ \begin{array}{l} f_0 \left(\frac{L}{L_f} \right)^{\alpha_f}, L \leq L_f \\ f_0, L > L_f \end{array} \right\} \quad (11)$$

$$K = \left\{ \begin{array}{l} K_0 \left(\frac{L}{L_K} \right)^{\alpha_K}, L \leq L_K \\ K_0, L > L_K \end{array} \right\} \quad (12)$$

The values of the constants in (10) - (12) are [15] $L_T = 13.45 \text{ cd/m}^2$, $S_0 = 94.7$, $\alpha_T = 0.649$, $\alpha_f = 0.182$, $f_0 = 6.78 \text{ cycle/deg}$, $L_f = 300 \text{ cd/m}^2$, $K_0 = 3.125$, $\alpha_K = 0.0706$ and $L_K = 300 \text{ cd/m}^2$. Equations 10 to 12 give T_{\min} , f_{\min} , and K as functions of local background luminance L . For a background intensity value of 128, given a gamma-corrected display, the corresponding local background luminance is computed as follows:

$$L = L_{\min} + 128 \frac{L_{\max} - L_{\min}}{M_g} \quad (13)$$

where L_{\min} and L_{\max} denote the minimum and maximum luminances of the display. Once the JND for a region with mean grayscale value of 128, t_{128} , is calculated using (6), the JND for regions with other mean grayscale values are approximated as follows [16]:

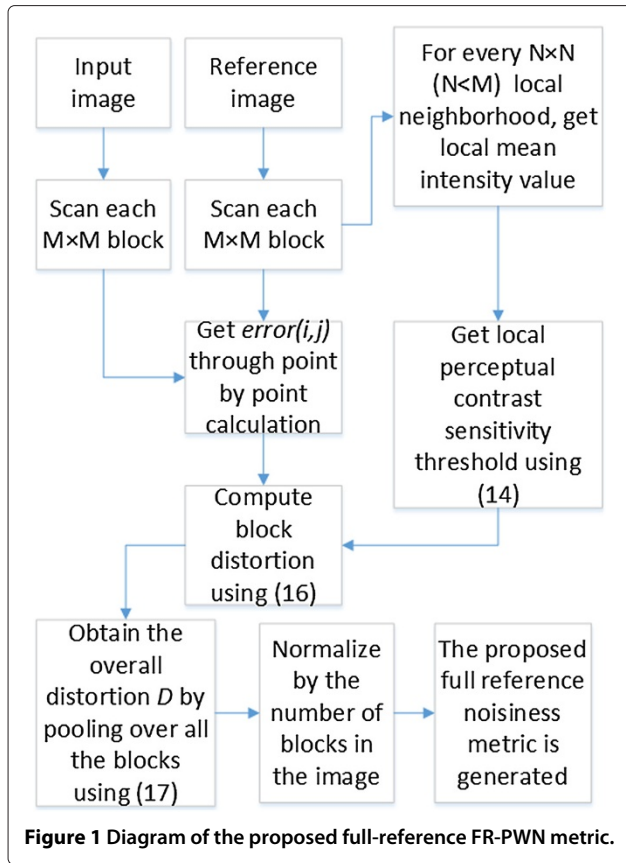
$$\begin{aligned} \text{JND}(i, j) &= t_{128} \left(\frac{\sum_{n_1=0}^{N-1} \sum_{n_2=0}^{N-1} I_{n_1, n_2}}{N^2(128)} \right)^{\alpha_T} \\ &= t_{128} \left(\frac{\text{Mean}(I_{n_1, n_2})}{128} \right)^{\alpha_T} \end{aligned} \quad (14)$$

where I_{n_1, n_2} is the intensity level at pixel location (n_1, n_2) in a $N \times N$ region surrounding pixel (i, j) . It should be noted that the indices (n_1, n_2) are used to denote the location with respect to the top left corner of the $N \times N$ region, while the indices (i, j) are used to denote the location with respect to the top left corner of the whole image. $\text{Mean}(I_{n_1, n_2})$ is the mean value over the considered $N \times N$ region surrounding pixel (i, j) . α_T is a correction exponent that controls the degree to which luminance masking occurs and is set to $\alpha_T = 0.649$, as given in [16]. JND(i, j) in (5) is computed using (14). In our implementation, $N = 8$ was used for the $N \times N$ region.

Full-reference noisiness metric

This work firstly presents a full-reference noisiness metric based on the probability summation model presented in the previous sections. Figure 1 shows the block diagram of the proposed full-reference FR-PWN metric. The input image is first divided into blocks of $M \times M$. The block will be the region of interest R_b . The block size is chosen to correspond with the foveal region. Let r be the visual resolution of the display in pixels per degree, v the viewing distance in centimeters, and d the display resolution in pixels per centimeter. Then the visual resolution can be calculated as follows [17]:

$$r = d \cdot v \cdot \tan(\pi/180) \approx d \frac{v\pi}{180} \approx d \frac{v}{57.3}. \quad (15)$$



In the HVS, the foveal region has the highest visual acuity and corresponds to about 2° of visual angle. The number of pixels contained in the foveal region can be computed as $(2 \lfloor r \rfloor)^2$ [17]. For example, for a viewing distance of 60 cm and 31.5 pixels/cm display, the number of pixels contained in the foveal region is $(64)^2$, corresponding to a block size of 64×64 . Using (5), the perceived noise distortion within a block R_b is given by

$$D_{R_b} = \left(\sum_{i,j \in R_b} \left| \frac{\text{error}(i,j)}{\text{JND}(i,j)} \right|^\alpha \right)^{1/\alpha} \quad (16)$$

where $\text{JND}(i,j)$ is the JND at location (i,j) and is computed using (14). Using the probability summation model as discussed previously, the noisiness measure D for the whole image I is obtained by using a Minkowski metric for inter-block pooling as follows:

$$D = \left(\sum_{R_b} |D_{R_b}|^\alpha \right)^{1/\alpha} \quad (17)$$

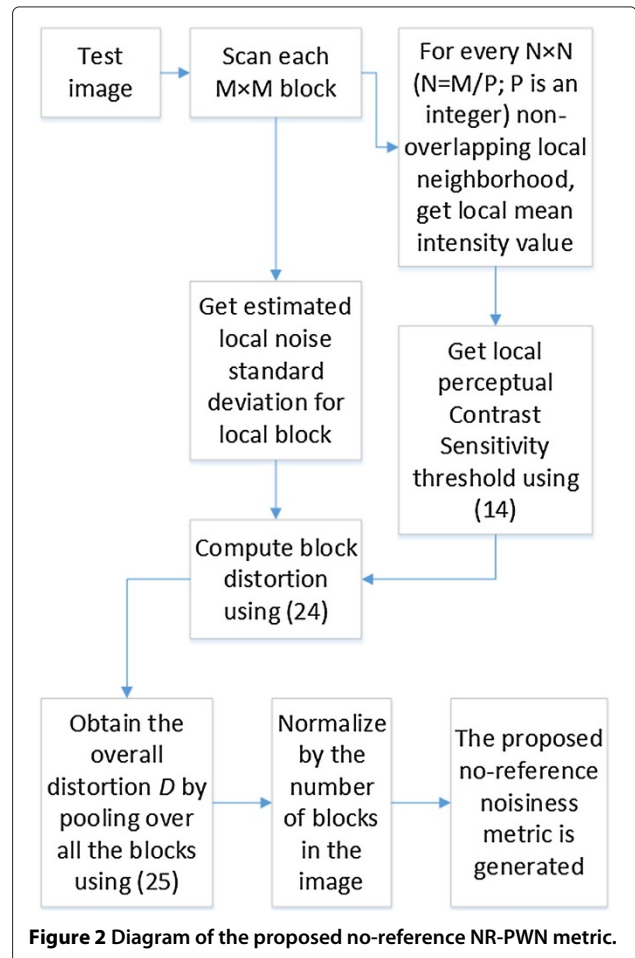
The resulting distortion measure, D , normalized by the number of blocks, is adopted as the proposed full-reference metric FR-PWN. This full-reference metric not

only works for noisiness, but could also work for other additive distortions.

No-reference noisiness metric

In the previous section, a full-reference quality metric is presented based on the probability summation model and JND. However, in many cases, the reference image is not available, so $\text{error}(i,j)$ in (16) can not be computed. Therefore, there is a need to develop a no-reference noisiness quality metric. Figure 2 shows the block diagram which summarizes the proposed no-reference NR-PWN metric. From (14), it can be seen that $\text{JND}(i,j)$ depends on the local mean of the neighborhood surrounding (i,j) . For the proposed NR metric, the local mean for a pixel (i,j) belonging to a region R_N is taken to be the mean of region R_N and is denoted by $\text{mean}(R_N)$. Consequently, Equation 14 can be written as follows:

$$\begin{aligned} \text{JND}(i,j) &= \text{JND}(R_N) \\ &= t_{128} \left(\frac{\text{Mean}(R_N)}{128} \right)^{\alpha_T}, \text{ for all } (i,j) \text{ belongs to } R_N. \end{aligned} \quad (18)$$



Now only one $JND(R_N)$ will be calculated for all pixel (i, j) belonging to the same R_N , and different $JND(R_N)$ will be calculated separately for each R_N within the considered region of interest block R_b . The size of the block R_b is chosen to approximate a foveal region (e.g., 64×64 as discussed previously). Using p, q as the indices within a local neighborhood R_N , the proposed NR metric is derived from the presented FR metric (16) as follows:

$$D_{R_b} = \left(\sum_{R_N \in R_b} \sum_{p, q \in R_N} \left| \frac{\text{error}(p, q)}{JND(p, q)} \right|^\alpha \right)^{1/\alpha} \quad (19)$$

$$= \left(\sum_{R_N \in R_b} \frac{\sum_{p, q \in R_N} |\text{error}(p, q)|^\alpha}{(JND(R_N))^\alpha} \right)^{1/\alpha}$$

In (19), $\sum_{p, q \in R_N} |\text{error}(p, q)|^\alpha$ can be approximated by $N^2 E[|\text{error}(p, q)|^\alpha]$ under the ergodicity assumption, where $N \times N$ is the size of each local neighborhood R_N . Also, if $\text{error}(p, q)$ is a Gaussian distribution process with a mean of 0 and a standard deviation of σ_{R_N} , using the central absolute moments of a Gaussian distribution process [18], it can be shown that

$$E[|\text{error}(p, q)|^\alpha] = \sigma_{R_N}^\alpha \frac{2^{\alpha/2} \Gamma(\frac{\alpha+1}{2})}{\pi^{1/2}}, \text{ for } \alpha > -1 \quad (20)$$

where $\Gamma(t)$ is the gamma function

$$\Gamma(t) = \int_0^\infty x^{t-1} e^{-x} dx. \quad (21)$$

Using (20), D_{R_b} in (19) can be written as follows:

$$D_{R_b} = \left(\sum_{R_N \in R_b} \frac{N^2 \sigma_{R_N}^\alpha \frac{2^{\alpha/2} \Gamma(\frac{\alpha+1}{2})}{\pi^{1/2}}}{(JND(R_N))^\alpha} \right)^{1/\alpha} \quad (22)$$

For a given α , define a constant C as

$$C = \frac{2^{\alpha/2} \Gamma(\frac{\alpha+1}{2})}{\pi^{1/2}}. \quad (23)$$

Then, the proposed NR noisiness metric over the region R_b is given by

$$D_{R_b} = \left(\sum_{R_N \in R_b} \frac{C \cdot N^2 \cdot \sigma_{R_N}^\alpha}{(JND(R_N))^\alpha} \right)^{1/\alpha}. \quad (24)$$

As in (17), the noisiness metric over the image I can be computed as follows:

$$D = \left(\sum_{R_b} |D_{R_b}|^\alpha \right)^{1/\alpha}. \quad (25)$$

The resulting noise measure D , normalized by the number of blocks, is adopted as the proposed no-reference NR-PWN metric.

In (24), the noise variance σ_{R_N} is estimated directly from the test image, without the reference image. Multiple methods are available to estimate the noise variance, such as fast noise variance estimation (FNV) [19] and generalized cross validation (GCV)-based method [20,21]. In our implementation, the GCV method was used for computing the local noise variance. Similar results were also obtained using the FNV [19] noise estimation method.

Performance results

The performance of the proposed FR-PWN and NR-PWN metrics is assessed using the LIVE [6] and TID2008 [22] databases. The LIVE database [6] consists of 29 RGB color image. The images are distorted using different distortion types: JPEG2000, JPEG, Gaussian blur, white noise, and bit errors. The difference mean opinion score (DMOS) for each image is provided. The white noise part of the LIVE database includes 174 images with a noise standard deviation ranging from 0 to 2. White noise was added to the RGB components of images after scaling between 0 and 1. All of the white noise images (174 images) from the LIVE database are used in our experiments. The TID2008 database [22] consists of 25 reference images (512×384) and 1,700 distorted images. The images are distorted using 17 types of distortions, including additive Gaussian noise, high-frequency noise, JPEG2000, and Gaussian blur. The MOS was obtained using a total of 838 observers with 256,428 comparisons of the visual quality of distorted images. All of the additive Gaussian noise image (100 images) and high-frequency noise images (100 images) from the TID2008 database are used in our experiments. As mentioned in [22], additive zero-mean noise is often present in images and it is commonly modeled as a white Gaussian noise. This type of distortion is included in most studies of quality metric effectiveness. High-frequency noise is an additive non-white noise which can be used for analyzing spatial frequency sensitivity of the HVS [23]. High-frequency noise is typical in lossy image compression and watermarking.

To measure how well the proposed metrics correlate with the provided subjective scores, the correlation coefficients adopted by VQEG [24] are used, including the Pearson's linear correlation coefficient (PLCC) and the Spearman rank-order correlation coefficient (SROCC). A four-parameter logistic function as suggested in [24] is used prior to computing the Pearson's linear correlation coefficient:

$$MOS_{P_i} = \frac{\beta_1 - \beta_2}{1 + \exp\left(\frac{M_i - \beta_3}{|\beta_4|}\right)} + \beta_2 \quad (26)$$

where M_i is the quality metric for image i , MOS_{P_i} is the predicted MOS or DMOS. Figure 3 shows the DMOS

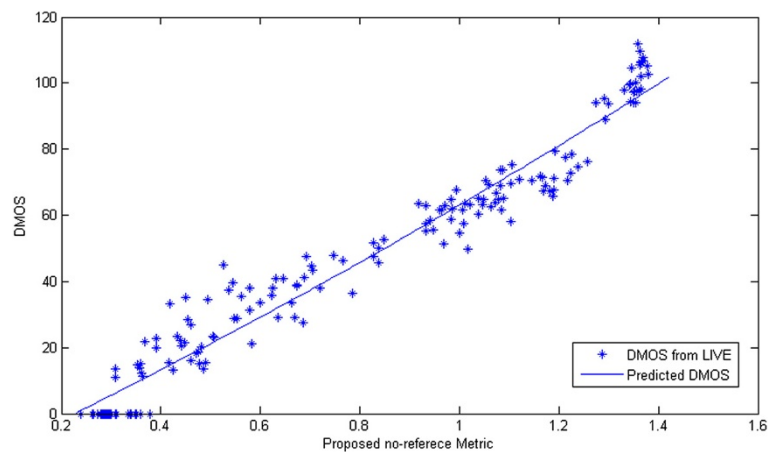


Figure 3 Correlation between the predicted score of NR-PWN and DMOS using the LIVE database.

score and predicted DMOS obtained using NR-PWN for the LIVE database.

Table 1 shows the evaluation results for the LIVE database. In addition to the proposed FR-PWN and NR-PWN metrics, the performance results of various existing metrics are presented for comparison, including seven full-reference metrics, DCTune [25], picture quality scale (PQS) [26], NQM [8], Fuzzy S7 [27], blockwise spectral distance measure (BSDM) [28], MS-SSIM [7],

IFC [9], one reduced reference metric quality-aware images (QAI) [29], and seven no-reference metrics, blind image integrity notator using DCT statistics (BLINDS-II) (SVM) [30], BLINDS-II (Prob.) [30], hybrid no-reference (HNR) [31], blind/referenceless image spatial quality evaluator (BRISQUE) [32], naturalness image quality evaluator (NIQE) [33], blind image quality index (BIQI) [34], and learning a blind measure of perceptual image quality (LBIQ) [35]. The benchmarks of full-reference metrics are obtained from [6], and the others are obtained from their respective authors or available implementations. The shown 'N/A' in Table 1 means the value is not provided in the literature.

Table 2 shows the performance of the proposed FR-PWN and NR-PWN metrics using images with different types of distortion as provided by the TID2008

Table 1 Performance evaluation for the LIVE database

	Metrics	PLCC	SROCC
FR	DCTune [25]	0.9288	0.9324
	PQS [26]	0.9603	0.9535
	NQM [8]	0.9885	0.9854
	Fuzzy S7 [27]	0.9038	0.9199
	BSDM (S4) [28]	0.9559	0.9327
	MS-SSIM [7]	0.9737	0.9805
	IFC [9]	0.9766	0.9625
	<i>FR-PWN (proposed)</i>	0.9846	0.9835
RR	QAI [29]	0.8889	0.8639
NR	BLINDS-II(SVM) [30]	0.9799	0.9691
	BLINDS-II(Prob.) [30]	0.9854	0.9783
	HNR [31]	0.962	N/A
	BRISQUE [32]	0.9851	0.9786
	NIQE [33]	0.9773	0.9662
	BIQI [34]	0.9538	0.9510
	LBIQ [35]	0.9761	0.9702
	Estimated noise standard deviation	0.9497	0.9713
		<i>NR-PWN (proposed)</i>	0.9770

Table 2 Performance evaluation using SROCC for the TID2008 database

	Metrics	Additive Gaussian noise	High-frequency noise
FR	MS-SSIM [7]	0.8094	0.8685
	DCTune [25]	0.8415	0.8721
	NQM [8]	0.7679	0.9015
	<i>FR-PWN (proposed)</i>	0.8818	0.9194
NR	BLINDS-II (SVM) [30]	0.6600	N/A
	BLINDS-II (Prob.) [30]	0.6956	0.7454
	BRISQUE [32]	0.829	0.6234
	NIQE [33]	0.7775	0.8539
	GRNN [36]	0.7532	N/A
	Li et al. [37]	0.7043	N/A
	<i>NR-PWN (proposed)</i>	0.8020	0.9136

database [22]. The proposed metric is compared with three full-reference metrics DCTune [25], NQM [8], MS-SSIM [7], and six very recent no-reference metrics that reported results for TID2008: BLINDS-II (SVM) [30], BLINDS-II (Prob.) [30], BRISQUE [32], NIQE [33], general regression neural network (GRNN) [36], and Li et al. [37]. The benchmarks of full-reference metrics are obtained from [22], and the others are obtained from their respective authors or available implementations. The shown N/A in Table 2 means the value is not provided in the literature. The proposed metric uses the same parameters as used with the LIVE database without any training.

From Table 1, it can be observed that the proposed FR-PWN metric outperforms the existing FR metrics for the LIVE database while achieving a similar performance as the NQM [8] metric. Table 2 shows that the proposed FR-PWN metric outperforms the existing FR metrics for the TID2008 database, on both Gaussian noise and high-frequency noise. The proposed NR-PWN metric comes close in performance to the proposed FR-PWN metric for both the LIVE and the TID2008 databases. In particular, Table 1 shows that the proposed NR-PWN metric performs better than existing NR metrics except for the Blinds-II and BRISQUE metrics in terms of PLCC. The proposed NR-PWN metric outperforms all the considered NR metrics in terms of SROCC and even existing FR metrics except the full-reference NQM [8] for the LIVE database. Table 2 shows that the proposed NR-PWN metric surpasses existing NR metrics except BRISQUE [32] for additive Gaussian noise, and that it significantly outperforms existing FR and NR metrics for high-frequency noise. Particularly, it should be noted that the performance of BRISQUE [32] drops dramatically on high-frequency noise and is significantly lower than the proposed metric. In addition, many of the shown state-of-the-art metrics including BLINDS-II [30], NIQE [33], and BRISQUE [32] use 80% of the data for training [30,32,33]. Consequently, these may not perform well on new distortions outside the training set, such as high-frequency noise (Table 2). In contrast, the proposed NR-PWN does not require training and still performs well on this new distortion.

Furthermore, it is worth indicating that as shown in Tables 1 and 2, the existing metrics exhibit differences in performance across different databases and types of distortions. It is noted in [38] that the performance of many image quality metrics could be quite different across databases. The difference in performance can be attributed to the differences in quality range, distortions, and contents across databases. Despite this, the results obtained show that the proposed FR-PWN and NR-PWN metrics achieve consistently a good performance across noise types (white noise and high-frequency noise) and

across databases as compared to the existing quality metrics. For example, the proposed FR-PWN metric exhibits a performance similar to NQM [8] for the LIVE database, while it significantly outperforms NQM [8] for white noise images from TID2008. Also, the existing BLINDS-II [30] performs fairly well for the LIVE database, but its performance significantly decreases when applied to TID2008. It is also interesting to note that although the mathematical derivations for the proposed NR-PWN is based on white noise, the proposed NR-PWN metric performs consistently well for high-frequency noise, a non-white noise.

The performance results presented in Tables 1 and 2 for the proposed NR-PWN metric are obtained using the GCV method [20,21] for local variance estimation. If the local variance is estimated using the FNV method [19], the resulting SROCC values are 0.9627 for the LIVE database additive Gaussian noise, 0.7850 for the TID2008 database additive Gaussian noise, and 0.9210 for the TID2008 database high-frequency noise, respectively.

Finally, the calculation of the proposed FR-PWN and NR-PWN metrics involves parameters of viewing conditions such as maximum luminance L_{max} of the monitor. However, the performance of the proposed metrics are resilient to different L_{max} values. In Tables 1 and 2, the proposed metrics are calculated using $L_{max} = 175 \text{ cd/m}^2$. The L_{max} in real viewing conditions may vary from 100 cd/m^2 for CRT monitors to 300 cd/m^2 for LCD monitors. Table 3 shows the performance of the proposed metric in terms of SROCC using different values of L_{max} , for both the LIVE and the TID2008 databases. It can be observed that the proposed metrics are not sensitive to the selection of L_{max} .

Conclusions

This paper proposed both a full-reference and a no-reference noisiness metrics. The no-reference noisiness metric is derived from the proposed full-reference metric

Table 3 SROCC of the proposed metrics using different L_{max}

		L_{max} (cd/m ²)	100	175	300
LIVE additive	FR-PWN		0.9835	0.9835	0.9835
	NR-PWN		0.9816	0.9816	0.9816
TID2008 additive	FR-PWN		0.8816	0.8818	0.8818
	NR-PWN		0.8020	0.8020	0.8020
TID2008 high-frequency noise	FR-PWN		0.9194	0.9194	0.9197
	NR-PWN		0.9136	0.9136	0.9136

and integrates noise variance estimation and perceptual contrast sensitivity thresholds into a probability summation model. The proposed metrics can predict the relative noisiness in images based on the probability of noise detection. Results show that the proposed metrics achieve a consistently good performance across noise types and across databases as compared to the existing quality metrics. Further work can be performed to develop a no-reference quality metric for multiply distorted images.

Competing interests

The authors declare that they have no competing interests.

Received: 15 April 2013 Accepted: 23 December 2013

Published: 16 January 2014

References

1. D Donoho, De-noising by soft-thresholding. *IEEE Trans. Inf. Theory*. **41**(3), 613–627 (1995)
2. R Hardie, A fast image super-resolution algorithm using an adaptive Wiener Filter. *IEEE Trans. Image Process.* **16**(12), 2953–2964 (2007)
3. R Ferzli, L Karam, A no-reference objective image sharpness metric based on the notion of just noticeable blur (JNB). *IEEE Trans. Image Process.* **18**(4), 717–728 (2009)
4. N Narvekar, L Karam, A no-reference image blur metric based on the cumulative probability of blur detection (CPBD). *IEEE Trans. Image Process.* **20**(9), 2678–2683 (2011)
5. R Ferzli, L Karam, No-reference objective wavelet based noise immune image sharpness metric, in *IEEE International Conference on Image Processing, ICIP 2005*, vol. 1 (IEEE, Piscataway, 2005), pp. 1–405–8
6. H Sheikh, M Sabir, A Bovik, A statistical evaluation of recent full reference image quality assessment algorithms. *IEEE Trans. Image Process.* **15**(11), 3440–3451 (2006)
7. Z Wang, E Simoncelli, A Bovik, Multiscale structural similarity for image quality assessment, in *Proceedings of the Thirty-Seventh Asilomar Conference on Signals, Systems and Computers, 2004*, vol. 2 (Pacific Grove, 9–12 November 2003), pp. 1398–1402
8. N Damera-Venkata, T Kite, W Geisler, B Evans, A Bovik, Image quality assessment based on a degradation model. *IEEE Trans. Image Process.* **9**(4), 636–650 (2000)
9. H Sheikh, A Bovik, G De Veciana, An information fidelity criterion for image quality assessment using natural scene statistics. *IEEE Trans. Image Process.* **14**(12), 2117–2128 (2005)
10. MCQ Farias, S Mitra, No-reference video quality metric based on artifact measurements, in *IEEE International Conference on Image Processing, ICIP 2005*, vol. 3 (IEEE, Piscataway, 2005), pp. III–141–144
11. MG Choi, JH Jung, JW Jeon, No-reference image quality assessment using blur and noise. *Int. J. Electrical Electron. Eng.* **3**, 318–322 (2009)
12. I Hontsch, L Karam, Locally adaptive perceptual image coding. *IEEE Trans. Image Process.* **9**(9), 1472–1483 (2000)
13. JG Robson, N Graham, Probability summation and regional variation in contrast sensitivity across the visual field. *Vis. Res.* **21**, 409–418 (1981)
14. L Karam, N Sadaka, R Ferzli, Z Ivanovski, An efficient selective perceptual-based super-resolution estimator. *IEEE Trans. Image Process.* **20**(12), 3470–3482 (2011)
15. A Ahumada, H Peterson, Luminance-model-based DCT quantization for color image compression. *Human Vision, Visual Processing, and Digital Display III*, SPIE Proc. **1666**, 365–374 (1992)
16. AB Watson, DCT quantization matrices visually optimized for individual images. *Human Vision, Visual Processing, and Digital Display IV*, SPIE Proc. **1913**, 202–216. (SPIE, San Jose, 1993)
17. Z Liu, L Karam, A Watson, JPEG2000 encoding with perceptual distortion control. *IEEE Trans. Image Process.* **15**(7), 1763–1778 (2006)
18. A Wilkelbauer, Moments and absolute moments of the normal distribution. (2012), <http://arxiv.org/abs/1209.4340>. (Accessed 15 April 2013)
19. J Immerkær, Fast noise variance estimation. *Comput. Vis. Image Underst.* **64**(2), 300–302 (1996). <http://www.sciencedirect.com/science/article/pii/S1077314296900600>.
20. G Wahba, Estimating the smoothing parameter, in *Spline Models for Observational Data* (Society for Industrial Mathematics, Philadelphia, 1990), pp. 45–65
21. D Garcia, Robust smoothing of gridded data in one and higher dimensions with missing values. *Comput. Stat. Data Anal.* **54**(4), 1167–1178 (2010)
22. N Ponomarenko, V Lukin, A Zelensky, K Egiazarian, M Carli, F Battisti, TID2008-A database for evaluation of full-reference visual quality assessment metrics. *Adv. Mod. Radioelectronics.* **10**, 30–45 (2009)
23. N Ponomarenko, F Silvestri, K Egiazarian, M Carli, J Astola, V Lukin, On between-coefficient contrast masking of DCT basis functions, in *Proceedings of the Third International Workshop on Video Processing and Quality Metrics*, vol. 2007 (Scottsdale, January 2007)
24. VQEG, Final report from the Video Quality Experts Group on the validation of objective models of video quality assessment (2000). <http://www.vqeg.org/>. (Accessed 15 April 2013)
25. AB Watson, DCTune: A technique for visual optimization of DCT quantization matrices for individual images. *Soc. Inf. Display Dig. Tech. Papers.* **24**, 946–949 (1993)
26. M Miyahara, K Kotani, V Algazi, Objective picture quality scale (PQS) for image coding. *IEEE Trans. Commun.* **46**(9), 1215–1226 (1998)
27. DV Weken, M Nachtegaal, EE Kerre, Using similarity measures and homogeneity for the comparison of images. *Image Vis. Comput.* **22**(9), 695–702 (2004)
28. I Avcibas, I Avcibas, B Sankur, K Sayood, Statistical evaluation of image quality measures. *J. Electron. Imaging.* **11**, 206–223 (2002)
29. Z Wang, G Wu, H Sheikh, E Simoncelli, EH Yang, A Bovik, Quality-aware images. *IEEE Trans. Image Process.* **15**(6), 1680–1689 (2006)
30. M Saad, A Bovik, C Charrier, Blind image quality assessment: a natural scene statistics approach in the DCT domain. *IEEE Trans. Image Process.* **21**(8), 3339–3352 (2012)
31. J Shen, Q Li, G Erlebacher, Hybrid no-reference natural image quality assessment of noisy, blurry, JPEG2000, and JPEG images. *IEEE Trans. Image Process.* **20**(8), 2089–2098 (2011)
32. A Mittal, A Moorthy, A Bovik, No-reference image quality assessment in the spatial domain. *IEEE Trans. Image Process.* **21**(12), 4695–4708 (2012)
33. A Mittal, R Soundararajan, A Bovik, Making a completely blind image quality analyzer. *IEEE Signal Process. Lett.* **20**(3), 209–212 (2013)
34. A Moorthy, A Bovik, A two-step framework for constructing blind image quality indices. *IEEE Signal Process. Lett.* **17**(5), 513–516 (2010)
35. H Tang, N Joshi, A Kapoor, Learning a blind measure of perceptual image quality, in *2011 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (IEEE, Piscataway, 2011), pp. 305–312
36. C Li, A Bovik, X Wu, Blind image quality assessment using a general regression neural network. *IEEE Trans. Neural Netw.* **22**(5), 793–799 (2011)
37. C Li, G Tang, X Wu, Y Ju, No-reference image quality assessment with learning phase congruency feature. *J. Electrical Inf. Technol.* **35**(2), 484–488 (2012)
38. S Tourancheau, F Atrousseau, Z Sazzad, Y Horita, Impact of subjective dataset on the performance of image quality metrics, in *15th IEEE International Conference on Image Processing* (IEEE Piscataway, 2008), pp. 365–368

doi:10.1186/1687-5281-2014-5

Cite this article as: Zhu and Karam: A no-reference objective image quality metric based on perceptually weighted local noise. *EURASIP Journal on Image and Video Processing* 2014 **2014**:5.