No-Reference Quality Assessment of JPEG Images via a Quality Relevance Map

S. Alireza Golestaneh, *Student Member, IEEE*, and Damon M. Chandler, *Senior Member, IEEE*

***Abstract*—This letter presents a no-reference quality assessment algorithm for JPEG compressed images (NJQA). Our method does not specifically aim to measure blockiness. Instead, quality is estimated by first counting the number of zero-valued DCT coefficients within each block, and then using a map, which we call the *quality relevance map*, to weight these counts. The quality relevance map for an image is a map that indicates which blocks are naturally uniform (or near-uniform) vs. which blocks have been made uniform (or near-uniform) via JPEG compression. Testing on various image-quality databases demonstrates that NJQA is either competitive with or outperforms modern competing methods on JPEG images.**

***Index Terms*—Discrete Cosine Transform (DCT), JPEG compressed images, No Reference Image Quality Assessment.**

I. Introduction

ALGORITHMS for image quality assessment (IQA) seek to predict the qualities of images in a manner that is consistent with quality ratings provided by human subjects. Of recent interest is the topic of no-reference IQA (NR IQA), in which the assessment is performed without the aid of the original, reference image. In this letter, we specifically address NR IQA of images which have undergone JPEG compression.

Over the past two decades, numerous algorithms have been proposed for NR IQA of JPEG images. A characteristic distortion in JPEG images is the presence of blocking artifacts; accordingly, the vast majority of NR IQA algorithms attempt to measure blockiness [1]–[16]. One common approach estimates the visibility of blocking artifacts based on block boundary edge-strength measurements (e.g.,[1]–[12]). Various transforms, such as the Discrete Cosine Transform (DCT) and Discrete Fourier Transform (DFT), are also commonly used for quantifying blockiness (e.g.,[13]–[16]).

In [1], Wu and Yuen estimated quality by measuring impairments at the block edges. Bailey *et al.* [4] measured blocking artifacts by comparing measured edge activity with an estimated blockiness-free activity. Pan *et al.* [5] estimated blockiness based on directional information measured for edges. Liu and Heynderickx [8] computed blockiness by weighting the local pixel-based distortion by its local visibility estimated via a model of visual masking. Wang *et al.* [9] presented an NR measure of blockiness which models blocky images as a non-blocky image corrupted by a pure blocky signal; they compute the quality of the image based on the energy of the blocky signal. Pan *et al.*[10] proposed a method which takes into account both the blockiness and the flatness to measure blocking artifacts. Perra *et al.* [11] exploited properties of the Sobel operator to generate a NR blockiness index that separately quantifies the luminance variation of block boundary pixels and the remaining pixels. Zhang [12] *et al.* presented a NR blockiness measure which calculates the image's luminance gradient matrix by using the Sobel operator. This matrix was used with perceptual adjustments (luminance adaptation and texture masking) to estimate the severity of blocking artifacts and the annoyance of large flat regions.

Blockiness has also been estimated via the DCT and other transforms. In [13], Bovik and Liu presented an NR measure of blockiness which operates in the DCT domain; blocking artifacts were first located via detection of 2D step functions, and then a human-vision-based measurement of blocking impairment was employed. Park *et al.* [14] presented an NR measure for blocking artifacts by modeling abrupt changes between adjacent blocks in both the pixel domain and the DCT domain. Tan and Ghanbari [15] applied a 2D DFT to small blocks segmented from the gradient image; a blockiness measure was developed by examining both the amplitude information and the phase information of the harmonics. Chen and Bloom [16] presented an NR DFT-based measure of blockiness. The DFT was employed to derive separate measures of horizontal and vertical blockiness, and then the overall blockiness was estimated via a combination of these two directional measures.

In this letter, we propose a NR IQA algorithm for JPEG images that does not specifically aim to measure blockiness. Instead, quality is estimated via a hybrid approach that uses both the DFT and DCT to estimate the degradation within each block. Specifically, in JPEG compression, as the compression ratio increases, the number of zero-valued DCT coefficients within each 8×8 block generally increases, thus making each block more uniform. Therefore, one potentially effective technique for NR IQA of JPEG images is to count the number of zero-valued DCT coefficients. However, such an approach cannot distinguish between naturally uniform/near-uniform blocks and distorted blocks that have been made uniform/near-uniform via quantization. Here, to overcome this problem, we use the DFT to compute what we call a *quality relevance map*.

The quality relevance map for an image is a map that indicates which blocks are naturally uniform (or near-uniform) vs. which blocks have been made uniform (or near-uniform) via

Manuscript received August 08, 2013; revised December 05, 2013; accepted December 16, 2013. Date of publication December 23, 2013; date of current version December 26, 2013. This work was supported by National Science Foundation Award 1054612. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Marco Mattavelli.

The authors are with the Laboratory of Computational Perception and Image Quality (CPIQ), School of Electrical and Computer Engineering, Oklahoma State University, Stillwater, OK 74078, USA. (e-mail: golesta@okstate.edu, damon.chandler@okstate.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/LSP.2013.2296038

JPEG compression. We estimate quality by first counting the number of zero-valued DCT coefficients within each block, and then use the quality relevance map to weight these counts. The quality relevance map is obtained by using a block-based DFT to compute the local slope of the rotationally averaged magnitude spectrum. We have previously used this DFT-based local slope measurement in [17] to develop a sharpness estimator for natural images; there, we showed that the spectral slope is useful for quantifying the perceived sharpness. In [17], the goal was not to estimate quality but instead to estimate local perceived sharpness. Here, we use the DFT-based local slope measurement to help distinguish between naturally vs. compression-induced uniform blocks across a wide range of JPEG compression ratios. As we will demonstrate, our resulting algorithm, which we call *NJQA*, performs well at IQA for a wide variety of JPEG images.

This letter is organized as follows: In Section II, we provide details of the NJQA algorithm. Section III presents and discusses the results of NJQA on the JPEG-compressed subsets of three image-quality databases. General conclusions are presented in Section IV.

II. Algorithm

Two stages are used in NJQA to predict the quality of the input image. The first stage generates the quality relevance map of the image. The second stage estimates quality by first counting the number of zero-valued DCT coefficients within each 8×8 block, and then summing and weighting these counts based on the quality relevance map.

Let X denote the $N_1 \times N_2$-pixel input image. If the input is a color image, we first convert it to grayscale via $X = 0.2989R + 0.5870G + 0.1140B$, where R, G, and B denote the 8-bit red, green, and blue intensities, respectively.

A. Quality Relevance Map

In natural images, uniform blocks rarely occur in spatial isolation. Rather, if a block is naturally uniform, it is usually surrounded by other uniform or nearly uniforms blocks. In contrast, blocks that originally contained structure but which have now been made uniform or nearly uniform via JPEG compression are almost always surrounded by non-uniform blocks. Thus, to distinguish between these two types of blocks (naturally uniform vs. compression-induced uniform), we examine a larger $M \times M$ region around each 8×8 block, where $M > 8$. If the local region is devoid of structure, then the block is deemed a naturally uniform block.

One limitation of the above approach is that the larger $M \times M$ region, which is comprised of multiple 8×8 blocks, might contain compression-induced blocking artifacts which can be confused with naturally occurring structure. One simple, yet effective technique to circumvent this confusion is to apply a motion-blurring filter to the image before examining each $M \times M$ region. Unlike standard deblocking algorithms which seek to maximize visual quality, the motion-blurring filter satisfies two important criteria: (1) When applied at an angle, the motion-blurring filter can very effectively eliminate horizontal and vertical blocking artifacts across a wide range of JPEG compression ratios. (2) The motion-blurring filter serves to fill-in uniform blocks with structure from surrounding blocks as long as the surrounding blocks also contain structure. Thus, the motion-blurring filter facilitates the process of distinguishing naturally uniform blocks from compression-induced uniform blocks. Blocks which are naturally uniform (e.g., sky) will remain uniform after motion blurring, whereas blocks which are uniform due to compression will contain a small amount of structure after motion blurring.

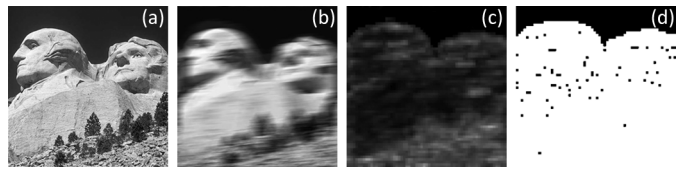

Fig. 1. Example results of the Quality Relevance Map stage. (a) Input image. (b) Motion-blurred version of (a). (c) The slope of the rotationally averaged magnitude spectrum of (b). (d) Quality relevance map. Note that (c) and (d) have been resized for visualization purposes only.

Here, we apply a motion-blurring filter[1] (see Fig. 1(b)) at an angle of $\theta = 5^\circ$ and with a *Length* of 50 pixels. These parameters were empirically selected to satisfy the aforementioned two criteria across a wide variety of images and compression ratios.

Next, to determine whether each $M \times M$ block is uniform or nearly uniform, we examine the local magnitude spectrum. Specifically, we compute the slope of the rotationally averaged magnitude spectrum for each 32×32 block with 24 pixels of overlap between neighboring blocks. Let $\tilde{X}$ denote the output of the motion blurring filter, and let $\tilde{x}$ denote a 32×32 block of $\tilde{X}$. We compute the 2D DFT of $\tilde{x}$, denoted as $y_{\tilde{x}}(f, \tilde{\theta})$, where f is the radial frequency and $\tilde{\theta}$ is the orientation. The magnitude spectrum, summed across all orientations of $y_{\tilde{x}}(f, \tilde{\theta})$, is given by $E_{\tilde{x}}(f) = \sum_{\tilde{\theta}} |y_{\tilde{x}}(f, \tilde{\theta})|$.

The slope of the magnitude spectrum of $\tilde{x}$, denoted as $\alpha_{\tilde{x}}$, is computed via $\alpha_{\tilde{x}} = \arg\min_\alpha \|f^{-\alpha} - E_{\tilde{x}}(f)\|_2^2$. Where the $L2$-norm is taken over all radial frequency $f > 0$. The overall normalized slope is given by:

$$S_l(\tilde{x}) = 1 - \frac{1}{1 + e^{-\tau_1(\alpha_{\tilde{x}} - \tau_2)}}, \tag{1}$$

where S denotes the normalized slope map, and $S_l(\tilde{x})$ denotes the lth pixel value of S (see Fig. 1(c)). Note that the map S has dimensions $\frac{N_1}{8} \times \frac{N_2}{8}$ due to the overlapping blocks. Again, this slope map is used here to identify blocks which appear uniform or nearly uniform. The sigmoid normalizes the resulting map to the range [0,1], where values near zero indicate more uniform areas. The constants $\tau_1 = 3$ and $\tau_2 = 2$ follow from our work in [17] where they were chosen to map the raw slope values to estimates of local perceived smoothness.

Finally, the binary quality relevance map, R, is formed via the following equation:

$$R_l = \begin{cases} 0, & \text{if } S_l(\tilde{x}) < \frac{1}{16} \\ 1 & \text{otherwise} \end{cases}, \tag{2}$$

where R_l indicates the lth pixel value of the quality relevance map (see Fig. 1(d)). In Equation (2), if the lth pixel value of the S map is less than $\frac{1}{16}$, the pixel is replaced by a black pixel. Otherwise, the pixel will be replaced with a white pixel. The threshold value is set low so that even subtle structure will be

[1]Using the MATLAB command `fspecial('motion', len, theta)`.

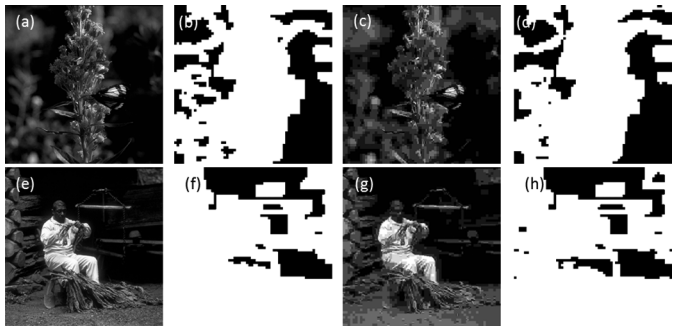

Fig. 2. Example quality relevance maps for high and low quality images. (a) and (e) Input images, $\mathrm{DMOS} = 0.023$ and 0.024. (b) and (f) Quality relevance maps of (a) and (e), respectively. (c) and (g) Input images, $\mathrm{DMOS} = 0.808$ and 0.666. (d) and (h) Quality relevance map of (c) and (g), respectively. Note that the maps have been resized for visualization purposes only.

considered relevant during the quality assessment process. The result is a map R, that differentiates between naturally uniform blocks (black areas) and blocks which have been made uniform by compression or contain structure (white areas).

The quality relevance map serves to distinguish between naturally nearly uniform blocks and those blocks which may have been made uniform during compression. Fig. 2 shows the similarity of quality relevance map for different versions of the same image with vastly different amounts of compression.

B. Quality Assessment

In the second stage of NJQA, we estimate quality by first counting the number of zero-valued DCT coefficients within each 8×8 block, and then use the quality relevance map to weight these counts. In general, as the quality decreases, the number of zero-valued DCT coefficients increases. However, a high quality image can also have areas with little spatial content resulting in many zeros in the DCT coefficients. Thus, it is necessary to employ our quality relevance map when considering the number of zero-valued DCT coefficients.

The quality assessment considers each 8×8 block of the input image X, and its corresponding quality relevance map R. Let $x_l^{8\times 8}$ denote the lth block of size 8×8 of the image X. The zero-valued DCT coefficients of the lth 8×8 image block are categorized according to the lth pixel value of the quality relevance map R. If the corresponding pixel value in the map is zero, the result goes in the lth element of a vector, b_l; otherwise, the result goes in the lth element of a vector, w_l; as follows:

$$b_l = \begin{cases} Z_s(x_l^{8\times 8}), & \text{if } R_l = 0 \\ 0 & \text{if } R_l = 1 \end{cases}, \qquad w_l = \begin{cases} Z_s(x_l^{8\times 8}), & \text{if } R_l = 1 \\ 0 & \text{if } R_l = 0 \end{cases}, \tag{3}$$

where $Z_s(x_l^{8\times 8})$ indicates the number of zero coefficients of the 2D DCT of $x_l^{8\times 8}$.

Finally, quality degradation is estimated via the following summation, weighted based on the quality relevance map:

$$NJQA = \frac{\sum_l w_l + 0.2 \sum_l b_l}{N_1 \times N_2}, \tag{4}$$

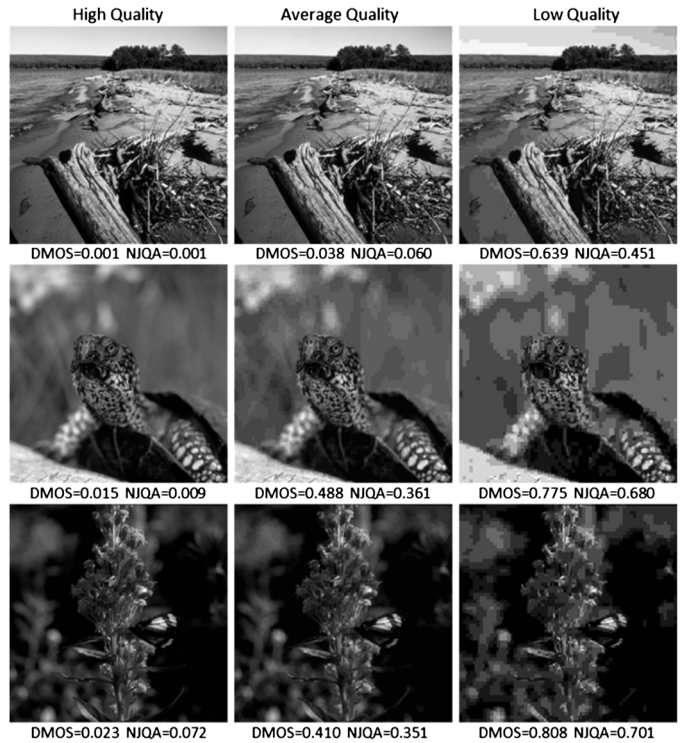

Fig. 3. Qualitative results of our proposed algorithm; additional qualitative results are available online at [18].

where $NJQA \in [0, 1]$, where 0 denotes perfect quality. We assume that areas with structures are more important than nearly uniform areas in estimating image quality. Therefore, in Equation (4) we give lesser weight to the $\sum_l b_l$ term.

We acknowledge that there are many empirical parameters used in this algorithm. However, as we demonstrate in the online supplement to this letter [18] the algorithm is robust to small changes in these parameters.

III. Results

This section analyses the performance of NJQA in terms of qualitative and quantitative results.

A. Qualitative Results

Fig. 3 provides results of our algorithm on three images with different qualities. As shown in Fig. 3, our algorithm can estimate the qualities of images over a range of different qualities in a manner that is consistent with human quality judgments (DMOS). Notice that as we move from left to right within each row, DMOS increases and NJQA follows a similar trend. In terms of the across-image quality assessment, as we move from top to bottom, DMOS increases and NJQA follows a similar trend. Additional qualitative results are available online at [18].

B. Quantitative Results

To quantify the performance of our algorithm, we applied NJQA to the JPEG image subsets of three image-quality databases: LIVE [19], CSIQ [20], and TID [21]. The JPEG subsets in the LIVE, CSIQ, and TID databases contain 204, 180, and 125 images, respectively. In these databases, 29 of 204, 30 of 180, and 25 of 125 were reference images. The

TABLE I
COMPARISON OF NJQA VS. VARIOUS NR IQA ALGORITHMS ON THE JPEG SUBSET OF THE LIVE DATABASE. BOLD ENTRIES ARE THE BEST AND SECOND-BEST PERFORMERS

	Ref. [1]	Ref. [4]	Ref. [5]	Ref. [6]	Ref. [8]	Ref. [13]	Ref. [10]	Ref. [14]	Ref. [11]	Ref. [12]	Ref. [16]	NJQA
PCC	0.956	0.938	0.932	0.941	0.940	0.935	0.923	0.922	0.901	0.928	**0.963**	**0.970**
SROCC	0.952	0.927	0.904	0.925	0.904	**0.955**	0.902	0.866	0.883	0.921	0.944	**0.967**

TABLE II
COMPARISON OF NJQA VS. GENERAL-PURPOSE NR IQA AND FR-IQA ALGORITHMS ON JPEG SUBSETS OF LIVE, CSIQ, AND TID. BOLD ENTRIES ARE THE BEST AND SECOND-BEST PERFORMERS

		PSNR	SSIM	MAD	DIIVINE	BLINDS-II	NJQA
LIVE	PCC	0.917	0.940	**0.944**	Trained	Trained	**0.963**
	SROCC	0.881	**0.966**	0.949	Trained	Trained	**0.956**
CSIQ	PCC	0.891	0.940	**0.983**	0.697	0.912	**0.954**
	SROCC	0.888	0.922	**0.966**	0.704	0.881	**0.925**
TID	PCC	0.887	0.937	**0.961**	0.899	0.928	**0.938**
	SROCC	0.871	0.897	**0.925**	0.871	0.889	**0.899**

performance of predicting subjective ratings of quality was measured in terms of the Pearson linear correlation coefficient (PCC) (following nonlinear regression; see [22]), and the Spearman rank-order correlation coefficient (SROCC) for gauging prediction monotonicity.

Table I provides the comparison between our results and various modern NR IQA algorithms designed specifically for JPEG images on the JPEG subset (including reference images) of the LIVE database. The results show that our algorithm yields high correlation with the subjective quality ratings. Moreover, as shown in Table I, our proposed method is the only method which yields the best results in terms of both PCC and SROCC.

We also analyzed our algorithm on the JPEG subsets of the LIVE, CSIQ, and TID databases (excluding reference images) in comparison to some modern general-purpose NR IQA methods that employ training (DIIVINE [23] and BLIINDS-II [24]), Peak Signal-to Noise Ratio (PSNR), the Structural Similarity Index (SSIM), and Most Apparent Distortion (MAD) [20]. The results, which are provided in Table II, show that our algorithm is competitive with these FR and trained NR IQA algorithms.

IV. CONCLUSION

This letter presented an algorithm for no-reference quality assessment of JPEG images. Our algorithm, called NJQA, estimates quality by first counting the number of zero-valued DCT coefficients within each block, and then using a quality relevance map to weight these counts. Testing on various image-quality databases has shown that NJQA is either competitive with or outperforms modern competing methods. A Matlab implementation of NJQA is available at [18].

REFERENCES

[1] H. Wu and M. Yuen, "A generalized block-edge impairment metric for video coding," *IEEE Signal Process. Lett.*, vol. 4, no. 11, 1997.

[2] L. Meesters and J.-B. Martens, "A single-ended blockiness measure for jpeg-coded images," *Signal Process.*, vol. 82, no. 3, 2002.

[3] Z. Wang, H. R. Sheikh, and A. C. Bovik, "No-reference perceptual quality assessment of jpeg compressed images," in *Proc. Int. Conf. Image Process.*, 2002, vol. 1, IEEE.

[4] D. Bailey, M. Carli, M. Farias, and S. Mitra, "Quality assessment for block-based compressed images and videos with regard to blockiness artifacts," in *Proc. Tyrrhenian Int. Workshop Digit. Commun.*, Capri, Italy, 2002.

[5] F. Pan, X. Lin, S. Rahardja, E. Ong, and W. Lin, "Using edge direction information for measuring blocking artifacts of images," *Multidimensional Syst. Signal Process.*, vol. 18, 2007.

[6] 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," *Adv. Multimedia Inf. Process.*, 2008.

[7] S. Suthaharan, "No-reference visually significant blocking artifact metric for natural scene images," *Signal Process.*, vol. 89, no. 8, 2009.

[8] H. Liu and I. Heynderickx, "A no-reference perceptual blockiness metric," in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process.*, 2008, IEEE.

[9] Z. Wang, A. C. Bovik, and B. Evan, "Blind measurement of blocking artifacts in images," in *Proc. Int. Conf. Image Process.*, 2000, vol. 3, IEEE.

[10] F. Pan, X. Lin, S. Rahardja, W. Lin, E. Ong, and S. Yao *et al.*, "A locally adaptive algorithm for measuring blocking artifacts in images and videos," *Signal Process. Image Commun.*, vol. 19, 2004.

[11] C. Perra, F. Massidda, and D. Giusto, "Image blockiness evaluation based on sobel operator," in *Proc. IEEE Int. Conf. Image Process. ICIP*, 2005, vol. 1, IEEE.

[12] Z. Hua, Z. Yiran, and T. Xiang, "A weighted sobel operator-based no-reference blockiness metric," in *Proc. Pacific-Asia Workshop Comput. Intell. Ind. Appl. PACIIA*, 2008, IEEE.

[13] A. C. Bovik and S. Liu, "DCT-domain blind measurement of blocking artifacts in dct-coded images," in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process. (ICASSP'01)*, 2001, vol. 3.

[14] C.-S. Park, J.-H. Kim, and S.-J. Ko, "Fast blind measurement of blocking artifacts in both pixel and dct domains," *J. Math. Imag. Vis.*, vol. 28, no. 3, 2007.

[15] K. Tan and M. Ghanbari, "Frequency domain measurement of blockiness in mpeg-2 coded video," in *Proc. Int. Conf. Image Process.*, 2000, vol. 3, IEEE.

[16] B. J. A. Chen and C. , "A blind reference-free blockiness measure," *Adv. Multimedia Inf. Process.*, 2010.

[17] C. T. Vu, T. D. Phan, and D. M. Chandler, "S3: A spectral and spatial measure of local perceived sharpness in natural images," *IEEE Trans. Image Process.*, vol. 21, no. 3, 2012.

[18] Online Supplement, "NJQA," [Online]. Available: http://vision.okstate.edu/njqa/njqa.htm

[19] 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, 2004.

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

[21] 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. Mod. Radioelectron.*, vol. 10, no. 4, pp. 30–45, 2009.

[22] Video Quality Experts Group [Online]. Available: http://www.vqeg.org

[23] 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, 2011.

[24] M. A. Saad, A. C. Bovik, and C. Charrier, "Blind image quality assessment: A natural scene statistics approach in the dct domain," *IEEE Trans. Image Process.*, vol. 21, no. 8, 2012.