SCREEN CONTENT IMAGE QUALITY ASSESSMENT USING EDGE MODEL

Zhangkai Ni[1], *Lin Ma*[2], *Huanqiang Zeng*[1,*], *Canhui Cai*[1], *and Kai-Kuang Ma*[3]

[1]School of Information Science and Engineering, Huaqiao University, Xiamen, China 361021
[2]Huawei Noah's Ark Lab, Hong Kong
[3]School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore 639798

ABSTRACT

Since the *human visual system* (HVS) is highly sensitive to edges, a novel *image quality assessment* (IQA) metric for assessing *screen content images* (SCIs) is proposed in this paper. The turnkey novelty lies in the use of an existing parametric edge model to extract two types of salient attributes—namely, *edge contrast* and *edge width*, for the distorted SCI under assessment and its original SCI, respectively. The extracted information is subject to conduct similarity measurements on each attribute, independently. The obtained similarity scores are then combined using our proposed *edge-width pooling* strategy to generate the final IQA score. Hopefully, this score is consistent with the judgment made by the HVS. Experimental results have shown that the proposed IQA metric produces higher consistency with that of the HVS on the evaluation of the image quality of the distorted SCI than that of other state-of-the-art IQA metrics.

Index Terms— Image quality assessment, screen content image, edge model.

1. INTRODUCTION

Unlike a camera-captured natural image, a typical *screen content image* (SCI) consists of a mixture of natural images, graphics, text, and even some icons generated and rendered by electronic devices. In recent years, how to objectively evaluate the image quality of SCIs has been receiving close attention due to the fact that such kind of images is often encountered in various multimedia applications and services, such as online education, video conference, electronic catalogs and advertisements, to name a few [1, 2].

One common issue shared among them is that the degradations of *perceptual* quality, which could be incurred through various image processing stages (e.g., compression, transmission, post-processing), might lead to unacceptable image quality. Therefore, an objective *image quality assessment* (IQA) metric is effective to guide and facilitate various image processing tasks, which has been intensively investigated in the past (e.g., [3]). Since most existing IQA metrics are developed for evaluating natural images, therefore it is imperative to develop an IQA metric specifically for SCIs. For that, a novel and accurate IQA metric is proposed in this paper.

For evaluating natural images, the simplest and widely-used IQA metric is the *peak signal-to-noise ratio* (PSNR) and the *mean square error* (MSE). However, it is also quite well recognized that these metrics are not highly consistent with the quality perceived by the *human visual system* (HVS). To address this problem, many IQA metrics have been developed by taking the HVS characteristics into account [4]. For example, the *structural similarity* (SSIM) [5] considers the degradation of structural information, which is more sensitive to the HVS. More such edge-information-based metrics can be found in [6, 7, 8, 9, 10].

To conduct image quality assessment for SCIs, Wang *et al.* [1] proposed an IQA metric by incorporating visual field adaptation and information content weighting. On the other hand, Yang *et al.* [2] considered the visual difference of textual and pictorial regions. In this paper, a novel and accurate IQA metric for SCIs using an existing parametric edge model [11] is proposed, which is motivated by the facts that (1) a typical SCI contains abundant of edge information; and (2) the HVS is highly sensitive to such information.

The remaining of this paper is organized as follows. In Section 2, the proposed IQA metric for evaluating SCIs is described in detail. Performance comparisons with other state-of-the-art IQA metrics are documented and discussed in Section 3. Finally, Section 4 draws the conclusion.

2. PROPOSED IQA METRIC FOR SCI USING EDGE MODEL

As shown in Fig. 1, the proposed IQA metric for SCIs consists of three stages as follows. In the first stage, an edge modeling process is applied to each input image to extract its salient

This work was supported in part by the National Natural Science Foundation of China under the Grants 61372107 and 61401167, in part by the Natural Science Foundation of Fujian Province under the Grant 2016J01308, in part by the Opening Project of State Key Laboratory of Digital Publishing Technology under the grant FZDP2015-B-001, in part by the High-Level Talent Project Foundation of Huaqiao University under the Grants 14BS201 and 14BS204, and in part by the Graduate Student Scientific Research Innovation Project Foundation of Huaqiao University.

*Corresponding author (Email: zeng0043@hqu.edu.cn).

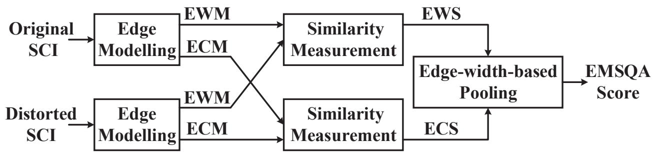

Fig. 1. The proposed IQA metric for screen content images.

edge attributes; i.e., edge contrast and edge width. This information is expressed in terms of maps, since the modelling is performed at each pixel location; thus, *edge contrast map* (ECM) and *edge width map* (EWM), respectively. Such edge modeling process will be conducted to the distorted SCI under evaluation and its original (or reference) SCI, respectively. In the second stage, the similarity between two ECMs will be measured and generate the *edge contrast similarity* (ECS). Likewise, the *edge width similarity* (EWS) will be computed based on the two EWMs. Lastly, a weighted pooling strategy based on the already-computed edge width information is exploited to produce the final IQA score, which is denoted as the *edge model-based SCI quality assessment* (EMSQA) score. The details of each step are described in the following subsections, respectively.

2.1. Edge Model

Since the HVS is highly sensitive to edges, an existing parametric edge model [11] is employed to model each input SCI for extracting its edge information. Considering that the edge model of an *ideal* step edge, centered at the location $x = x_0$, can be mathematically expressed as [11]

$$u(x; b, c, x_0) = c \cdot U(x - x_0) + b, \tag{1}$$

where $U(\cdot)$ denotes the unit step function, c represents the edge contrast, and b is the luminance intensity at the lower side of edge contrast. However, a typical edge encountered in natural images or in SCIs usually have a smooth transition, rather than showing a sharp edge as an ideal unit-step signal. Therefore, an edge in SCI incurred at $x = x_0$ can be approximately expressed as the Gaussian-kernel-smoothed step edge as illustrated in Fig. 2; that is,

$$\begin{aligned} s(x; b, c, w, x_0) &= u(x; b, c, x_0) \otimes g(x; w) \\ &= u(x; b, c, x_0) \otimes \frac{1}{\sqrt{2\pi w^2}} exp\left(\frac{-x^2}{2w^2}\right), \\ &= b + \frac{c}{2}\left(1 + \text{erf}\left(\frac{x - x_0}{\sqrt{2}w}\right)\right) \end{aligned} \tag{2}$$

where the symbol "$\otimes$" denotes the convolution operation, $\text{erf}(\cdot)$ is the error function, w is the standard deviation of the Gaussian kernel smoothing function (i.e., $g(x; w)$), which is used to model the edge width. Note that any edge can be modelled by this parametric edge model (i.e., Fig. 2) via luminance b, contrast c, and width w, respectively. Now, the key issue is how to determine these three parameters, b, c, and w, for each edge model. This can be derived by fitting (2) based on the local pixel intensity. For that, the edge signal $s(x; b, c, w, x_0)$ is further convolved with the derivative of the Gaussian filter $g'_d = (x, \sigma_d)$ and arrive at

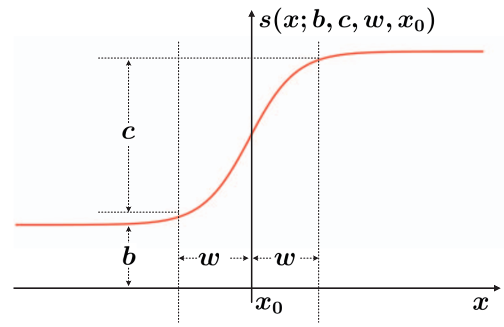

Fig. 2. An illustration of an edge model with an edge incurred at the location $x = x_0$, where c is the contrast parameter, w is the width parameter, and b is the local luminance at $x = x_0$ [11].

$$d(x; c, w, \sigma_d, x_0) = \frac{c}{\sqrt{2\pi(w^2 + \sigma_d^2)}} exp\left(\frac{-(x - x_0)^2}{2(w^2 + \sigma_d^2)}\right). \tag{3}$$

These three parameters b, c, and w can be computed by sampling (3) at three locations $x = 0, -a$, and a. That is, given $d_1 = d(0; c, w, \delta_d, x_0)$, $d_2 = d(a; c, w, \delta_d, x_0)$, and $d_3 = d(-a; c, w, \delta_d, x_0)$, the above-mentioned three parameters can be estimated as follows:

$$c = d_1 \cdot \sqrt{2\pi a^2 / ln(l_1)} \cdot l_2^{\frac{1}{4a}}, \tag{4}$$

$$w = \sqrt{a^2 / ln(l_1) - \sigma_d^2}, \tag{5}$$

$$b = s(x_0) - \frac{c}{2}, \tag{6}$$

where $l_1 = d_1^2 / d_2 d_3$ and $l_2 = d_2 / d_3$. The sampling distance a can be chosen freely, and $a = 1$ is used in this paper.

Resulted from the edge modeling process, all the values of the parameters c and w can be separately grouped and formed an ECM and an EWM, respectively. These maps are viewed as the extracted salient edge information and served as the inputs to the subsequent processing stage (refer to Fig. 1).

To demonstrate the effectiveness of edge model for SCI, Fig. 3 shows the ECM and EWM of a typical SCI and that of its distorted version after applying certain amount of motion blur. This test SCI was selected from the SIQAD database [2]. Comparing Figs. 3 (b) and (e) (or, Figs. 3 (c) versus (f) for that regard), it can be easily observed from the pictorial region that significant amount of edge information and texture information got lost due to motion blur as expected. This shows that the edge modeling process effectively extracted salient edge attributes, and they are considered equally important to the quality assessment for SCIs.

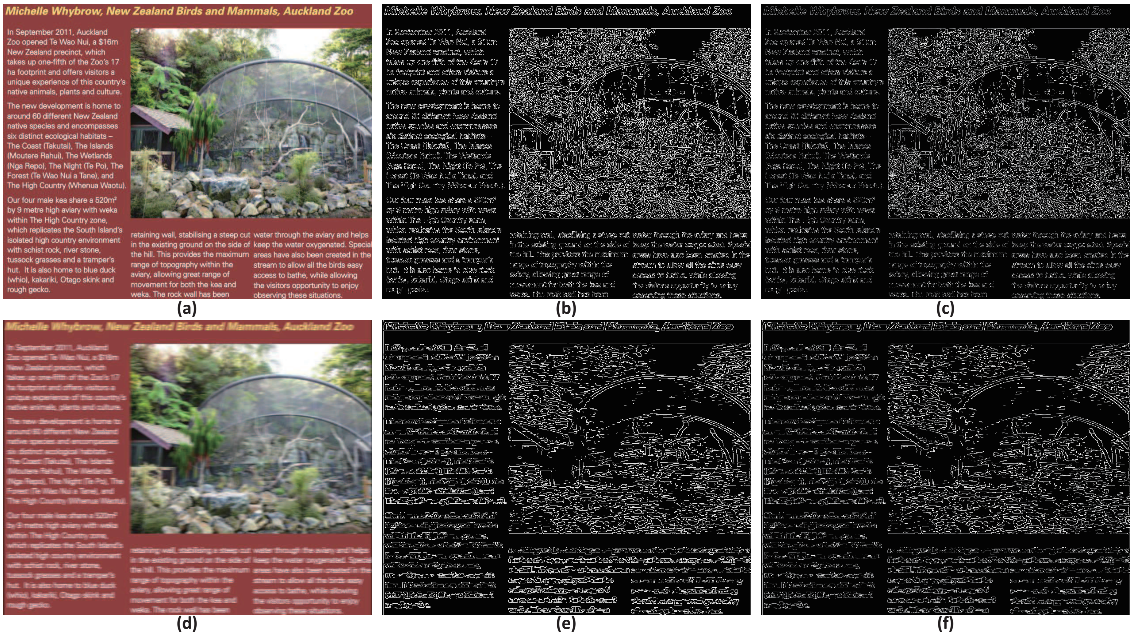

Fig. 3. An example of *edge contrast map* (ECM) and *edge width map* (EWM) obtained from applying an edge modeling process to an SCI and its distorted version with motion blur, respectively: (a) an original or reference SCI; (b) ECM of (a); (c) EWM of (a); (d) distorted SCI resulted from motion blur; (e) ECM of (d); and (f) EWM of (d).

2.2. Edge Contrast Similarity and Edge Width Similarity

Given a reference SCI r and its distorted version d, let C_r and C_d be the ECMs of r and d, respectively. By following the same practice as proposed in [5] on the computation of the degree of similarity, our *edge contrast similarity* (ECS) measurement based on the two input data fields, C_r and C_d, can be hence computed according to

$$ECS(x,y) = \frac{2C_r(x,y)C_d(x,y) + T_c}{C_r^2(x,y) + C_d^2(x,y) + T_c}, \tag{7}$$

where $ECS(x,y) \in [0,1]$ and $T_c = 330$ is a small positive constant that is used here to avoid any numerical instability.

Similarly, let W_r and W_d denote the EWMs of r and d, respectively. Our proposed *edge width similarity* (EWS) can be computed according to

$$EWS(x,y) = \frac{2W_r(x,y)W_d(x,y) + T_w}{W_r^2(x,y) + W_d^2(x,y) + T_w}, \tag{8}$$

where $EWS(x,y) \in [0,1]$ and $T_w = 10$ is used in our work for preventing from any numerical instability.

Considering both $ECS(x,y)$ and $EWS(x,y)$ can consistently reflect the perceptual similarity, denoted as $S(x,y)$, between r and d, they are combined together in a similar way as practiced in [5]; that is,

$$S(x,y) = [ECS(x,y)]^{\alpha} \cdot [EWS(x,y)]^{\beta}, \tag{9}$$

where α and β are two positive constants that can be used to adjust the relative importance of $ECS(x,y)$ and $EWS(x,y)$. By further considering that these two measurements are equally important to SCI quality assessment, $\alpha = \beta = 1$ is set in our work.

2.3. Edge-width-based Pooling

Since each pixel makes a different contribution to the whole image quality via the computed $S(x,y)$, the final image quality assessment score of the distorted SCI can be obtained by using a pooling strategy. Further note that the edge width w in the edge model (refer to Fig. 2) reflects the edge structure, and the smaller the value of w, the shaper the edge. Considering that the HVS perception is very sensitive to edge structure, intuitively it is meaningful to exploit the edge width information, $w(x,y)$, obtained at each pixel location as its weighting factor to weight the importance of $S(x,y)$ on the computation of the final image quality score of the distorted SCI d, with respect to the original SCI r. The computed IQA score is called the *edge model-based SCI quality assessment* (EMSQA); i.e.,

$$EMSQA = \frac{\sum_{(x,y)\in\Omega} W_m(x,y) \cdot S(x,y)}{\sum_{(x,y)\in\Omega} W_m(x,y)}, \tag{10}$$

where Ω denotes the set of pixel locations of the entire image under evaluation.

3. EXPERIMENTAL RESULTS

In this section, we compare the figure of merit computed by the proposed EMSQA metric with that of other state-of-the-

Table 1. Performance comparisons of different IQA metrics on the *SIQAD* database [2].

	Distortions	PSNR	SSIM [5]	MSSIM [12]	IWSSIM [13]	VIF [14]	IFC [15]	VSNR [16]	MAD [17]	FSIM [18]	GSIM [9]	GMSD [10]	SPQA [2]	Q_s [1]	EMSQA
PLCC	GN	0.9053	0.8806	0.8783	0.8804	0.9011	0.8791	0.8840	0.8852	0.7428	0.8448	0.8956	0.8921	-	0.8889
	GB	0.8603	0.9014	0.8984	0.9079	0.9102	0.9061	0.8890	0.9120	0.7206	0.8831	0.9094	0.9058	-	0.9156
	MB	0.7044	0.8060	0.8240	0.8414	0.8490	0.6782	0.7829	0.8361	0.6874	0.7711	0.8436	0.8315	-	0.8753
	CC	0.7401	0.7435	0.8371	0.8404	0.7076	0.6870	0.7667	0.3933	0.7507	0.8077	0.7827	0.7992	-	0.7688
	JPEG	0.7545	0.7487	0.7756	0.7998	0.7986	0.7606	0.7972	0.7662	0.5566	0.6778	0.7746	0.7696	-	0.7904
	J2K	0.7893	0.7749	0.7951	0.8040	0.8205	0.7963	0.8170	0.8344	0.6675	0.7242	0.8509	0.8252	-	0.7850
	LSC	0.7805	0.7307	0.7729	0.8155	0.8385	0.7679	0.7982	0.8184	0.5964	0.7218	0.8559	0.7958	-	0.7747
	Overall	0.5869	0.7561	0.6161	0.6527	0.8189	0.6395	0.5982	0.6191	0.5389	0.5663	0.7387	**0.8584**	**0.8573**	**0.8648**
SROCC	GN	0.8790	0.8694	0.8679	0.8743	0.8888	0.8717	0.8662	0.8721	0.7373	0.8404	0.8856	0.8823	-	0.8745
	GB	0.8573	0.8921	0.8883	0.9060	0.9059	0.9106	0.8827	0.9087	0.7286	0.8796	0.9119	0.9017	-	0.9154
	MB	0.7130	0.8041	0.8238	0.8421	0.8492	0.6737	0.7799	0.8357	0.6641	0.7753	0.8441	0.8255	-	0.8788
	CC	0.6828	0.6405	0.7506	0.7563	0.6433	0.6396	0.6694	0.3907	0.7175	0.7148	0.6378	0.6154	-	0.6306
	JPEG	0.7569	0.7576	0.7787	0.7978	0.7924	0.7636	0.8084	0.7674	0.5879	0.6796	0.7712	0.7673	-	0.7871
	J2K	0.7746	0.7603	0.7855	0.7998	0.8131	0.7980	0.8112	0.8382	0.6363	0.7125	0.8436	0.8152	-	0.7762
	LSC	0.7930	0.7371	0.7711	0.8214	0.8463	0.7713	0.8088	0.8154	0.5979	0.7145	0.8592	0.8003	-	0.7798
	Overall	0.5604	0.7566	0.6115	0.6545	0.8065	0.6011	0.5743	0.6067	0.5279	0.5551	0.7305	**0.8416**	**0.8456**	**0.8504**
RMSE	GN	6.3372	7.0679	7.1309	7.7044	6.4673	7.1096	6.9721	6.9391	9.9860	7.9811	6.6354	6.7394	-	6.8320
	GB	7.7376	6.5701	6.6638	6.3619	6.2859	6.4193	6.9506	6.2269	10.5230	7.1210	6.3111	6.4301	-	6.0710
	MB	9.2287	7.6967	7.3675	7.0600	6.8704	9.5544	8.0897	7.1322	9.4432	8.2788	6.9816	7.2223	-	6.2880
	CC	8.4591	8.4116	6.8818	6.8184	8.8876	9.1407	8.0760	11.5652	8.3190	7.4160	7.8297	7.6184	-	8.0440
	JPEG	6.1665	6.2295	5.9311	5.6406	5.6551	6.1004	5.6726	6.0380	7.8072	6.9085	5.9427	6.0000	-	5.7560
	J2K	6.3819	6.5691	6.3040	6.1804	5.9412	6.2875	5.9929	5.7276	7.7404	7.1675	5.4591	5.8706	-	6.4390
	LSC	5.3336	5.8253	5.4141	4.9379	4.6497	5.4657	5.1429	4.9025	6.8486	5.9046	4.4121	5.1664	-	5.3950
	Overall	11.5898	9.3676	11.2744	10.8444	8.1969	11.0048	11.4706	11.2409	12.0583	11.798	9.6484	**7.3421**	**7.3030**	**7.1860**

art metrics using the SCI images chosen from the SIQAD database [2]. This database is specifically established for evaluating the perceptual quality of SCI, and it contains 20 reference images and 980 distorted SCIs, including 7 types of distortions with 7 different levels of degradations generated for each type of distortion. The distortion types under our investigations include Gaussian noise (GN), Gaussian blur (GB), motion blur (MB), contrast change (CC), JPEG compression, JPEG2000 (JP2K) compression, and layer-segmentation-based coding (LSC). Three standard performance evaluation procedures and criteria are used [19], namely, the *Pearson linear correlation coefficient* (PLCC) for prediction accuracy, the *Spearman rank order correlations coefficient* (SROCC) for prediction monotonicity, and the *root mean square prediction error* (RMSE) for prediction consistency. Note that a higher value of PLCC and SROCC means a better accuracy and prediction monotonicity, while a smaller RMSE indicates a better performance.

To demonstrate the superiority, the proposed EMSQA metric is compared with the classical and state-of-the-art quality assessment metrics, including PSNR, SSIM [5], MSSIM [12], IWSSIM [13], VIF [14], IFC [15], VSNR [16], MAD [17], FSIM [18], GSIM [9], GMSD [10], SPQA [2] and Q_s [1], where the latter two IQA methods are specifically designed for the evaluation of SCIs. Table 1 documents the computed quality assessment measurements of these methods for each concerned distortion type using the test SCIs chosen from the SIQAD database. The top three performance figures of each measurement scheme are highlighted in boldface. Moreover, the performance comparisons on each distortion type are also provided, except Q_s, which does not provide the results on each distortion.

Some observations are provided as follows. First, it can be observed that the proposed EMSQA metric is able to achieve the highest correlation or consistency with the subjective quality ratings and outperforms all the state-of-the-art metrics under comparison. Second, it can be observed that the proposed EMSQA metric is able to more accurately assess and reflect the degradations caused by Gaussian blur (GB) and motion blur (MB) particularly. In fact, this is expected since blurring will easily degrade edges and make significant changes on the extracted edge information. Refer to Fig. 3 for a demonstration.

4. CONCLUSION

This paper presents a novel screen content image quality assessment metric, which is able to deliver more objective evaluation for the screen content images from the perceptual point of view. That is, the figure of merit computed by the proposed metric is more consistent with what is to be judged by the human visual system. Given a distorted SCI under assessment, the measurement is performed with reference to its original SCI. The novelty of this work lies in the use of an existing edge model to extract salient edge information for conducting quality assessment. Extensive experiments conducted over an SCI benchmark database have demonstrated that the proposed IQA metric clearly outperforms all other state-of-the-art IQA metrics on delivering more consistent assessment in accordance with that perceived by the human visual system.

5. REFERENCES

[1] S. Wang, K. Gu, K. Zeng, Z. Wang, and W. Lin, "Perceptual screen content image quality assessment and compression," *IEEE International Conference on Image Processing*, pp. 1434–1438, September 2015.

[2] H. Yang, Y. Fang, and W. Lin, "Perceptual quality assessment of screen content images," *IEEE Transactions on Image Processing*, vol. 24, no. 11, pp. 4408–4421, August 2015.

[3] S. Wang, K. Ma, W. Wang, H. Yeganeh, Z. Wang, and W. Lin, "A patch-structure representation method for quality assessment of contrast changed images," *IEEE Signal Processing Letters*, vol. 22, no. 11, pp. 2387–2390, October 2015.

[4] W. Lin and C. C. Jay Kuo, "Perceptual visual quality metrics: A survey," *Journal of Visual Communication and Image Representation*, vol. 22, no. 2, pp. 297–312, May 2011.

[5] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *IEEE Transactions on Image Processing*, vol. 13, no. 4, pp. 600–612, April 2004.

[6] G. Chen, C. Yang, and S. Xie, "Edge-based structural similarity for image quality assessment," *IEEE International Conference on Acoustics, Speech and Signal Processing*, pp. 14–19, May 2006.

[7] W. Xue and X. Mou, "An image quality assessment metric based on non-shift edge," *IEEE International Conference on Image Processing*, pp. 3309–3312, September 2011.

[8] X. Zhang, X. Feng, W. Wang, and W. Xue, "Edge strength similarity for image quality assessment," *IEEE Signal Processing Letters*, vol. 20, no. 4, pp. 319–322, April 2013.

[9] A. Liu, W. Lin, and M. Narwaria, "Image quality assessment based on gradient similarity," *IEEE Transactions on Image Processing*, vol. 21, no. 4, pp. 1500–1512, April 2012.

[10] W. Xue, L. Zhang, X. Mou, and A. C. Bovik, "Gradient magnitude similarity deviation: A highly efficient perceptual image quality index," *IEEE Transactions on Image Processing*, vol. 23, no. 5, pp. 684–695, February 2014.

[11] P. J. L. van Beek, *Edge-based image representation and coding*, Ph.D. thesis, Delft University of Technology, Delft, The Netherlands, 1995.

[12] Z. Wang, E. P. Simoncelli, and A. C. Bovil, "Multiscale structural similarity for image quality assessment," *IEEE Conference on Signals, Systems and Computers*, vol. 2, pp. 1398–1402, November 2003.

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

[14] H. R. Sheikh and A. C. Bovik, "Image information and visual quality," *IEEE Transactions on Image Processing*, vol. 15, no. 2, pp. 430–444, February 2006.

[15] H. R. Sheikh, A. C. Bovik, and G. de Veciana, "An information fidelity creation for image quality assessment using natural scene statistics," *IEEE Transactions on Image Processing*, vol. 14, no. 12, pp. 2117–2128, December 2005.

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

[17] E. C. Larson and D. M. Chandler, "Most apparent distortion: Full-reference image quality assessment and the role of strategy," *Journal of Electronic Imaging*, vol. 19, no. 1, pp. 011006–1–011006–21, 2010.

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

[19] VQEG, "Final report from the video quality experts group on the validation of objective models of video quality assessment," *http://www.its.bldrdoc.gov/vqeg/vqeg-home.aspx*, August 2015.