Screen Content Image Quality Assessment Using Multi-Scale Difference of Gaussian

Ying Fu, Huanqiang Zeng, *Senior Member, IEEE*, Lin Ma, Zhangkai Ni, Jianqing Zhu, *Member, IEEE*, and Kai-Kuang Ma, *Fellow, IEEE*

***Abstract*— In this paper, a novel *image quality assessment* (IQA) model for the *screen content image*s (SCIs) is proposed by using *multi-scale difference of Gaussian* (MDOG). Motivated by the observation that the *human visual system* (HVS) is sensitive to the edges while the image details can be better explored in different scales, the proposed model exploits MDOG to effectively characterize the edge information of the reference and distorted SCIs at two different scales, respectively. Then, the degree of edge similarity is measured in terms of the smaller-scale edge map. Finally, the edge strength computed based on the larger-scale edge map is used as the weighting factor to generate the final SCI quality score. Experimental results have shown that the proposed IQA model for the SCIs produces high consistency with human perception of the SCI quality and outperforms the state-of-the-art quality models.**

***Index Terms*— Human visual system (HVS), image quality assessment (IQA), screen content image (SCI), multi-scale difference of Gaussian.**

I. Introduction

With rapid development of Internet and wireless communication, the *screen content image*s (SCIs) are increasingly indispensable to abundant multimedia applications, such as, screen sharing, online education, cloud gaming, remote computing, and so on [1]. Similar to the natural images, the SCIs could also suffer from various distortions during the acquisition, processing, compression, transmission, and display stages [2]. For example, the noise is often appeared in the SCI transmission stage, the contrast change and color distortions are often encountered on SCIs in the screen sharing among different electronic devices, and the artifacts are unavoidable when the SCIs are compressed. Therefore, it is very essential to develop the *image quality assessment* (IQA) models for the SCIs.

The IQA can provide guidelines for the performance optimization and measurement during each stage of image processing and communication systems [3]. Since the human eyes are the ultimate receivers of the processed images, a good objective IQA model is to evaluate the image quality in a way that it can automatically predict image quality consistently with the *human visual system* (HVS) perception. Traditionally, the *mean square error* (MSE) and *peak signal-to-noise ratio* (PSNR) are widely used to measure visual quality, because they have low complexity and present clear physical meaning. However, MSE/PSNR are widely criticized for a lack of good correlation with human subject judgements. In the past decades, many researchers have made great progresses in the design of better IQA algorithms by considering various HVS characteristics, such as image structure [4], [5], multi-scale structure representation [6], edge information [7]–[9], visual saliency [10], information fidelity [11], [12], biologically inspired features [13], [14], and so on. For example, the well-known *structure similarity* (SSIM) [4] and its variants (e.g., *multi-scale SSIM* (MSSIM) [5]) bring IQA from pixel-based stage to structure-based stage according to the observation that the human eyes are sensitive to the structure inherited in images. Liu *et al.* [8] use the contrast between pristine and distorted oriented gradient maps to measure image quality.

However, most of the existing IQA algorithms are designed for the natural images. Different from the natural images, the SCIs are generally composed of different proportions of continuous-tone content (e.g., the natural scene image) and discontinuous-tone content (e.g., the texts, graphics, charts, etc.), as shown in Fig. 1. One can further see that in each SCI, the discontinuous-tone content usually features extremely sharp edges, texts and limited color variations, while the continuous-tone content contains relatively smooth edges, complex textures, and fruitful colors [15]. Obviously, the characteristics of the SCIs are quite different from that of the natural images. Therefore, most of the existing IQA metrics cannot effectively depict the SCI perceptual quality. To conduct IQA for the SCIs, Yang *et al.* [2] construct an IQA database specifically for SCI (i.e., SIQAD) and develop a SCI *perceptual quality assessment* (SPQA) metric, which considers the difference between the textual and pictorial regions. After that, Gu *et al.* [15] design an IQA model for the SCIs by weighting the SSIM via the structural degradation measurement. Ni *et al.* [16] makes the first attempt to explore the gradient direction to effectively evaluate the SCI perceptual quality. Wang *et al.* [17] propose to incorporate visual field adaptation and information content weighting. Ni *et al.* [18] propose an IQA metric to make good use of edge information via edge modeling on SCI quality assessment.

In this paper, an effective IQA model for the SCIs is proposed by using *multi-scale difference of Gaussian* (MDOG). The key novelty of the proposed approach lies in the use of MDOG to effectively explore the edge information from different scales for SCI quality assessment based on the following observations: 1) a typical SCI contains fruitful edge information and the HVS is highly sensitive to such kind of information; 2) the image details can be better acquired in multiple scales. For that, the edge maps at two different scales of the reference and distorted SCIs are firstly extracted by using MDOG, respectively. The edge similarity is then measured together with edge strength-based pooling strategy to generate the final SCI

Manuscript received December 24, 2017; accepted June 24, 2018. Date of publication July 9, 2018; date of current version September 13, 2018. This work was supported in part by the National Natural Science Foundation of China under Grant 61401167, Grant 61372107, and Grant 61602191, in part by the Natural Science Foundation of Fujian Province under Grant 2016J01308 and Grant 2017J05103, in part by the Fujian-100 Talented People Program, in part by the High-level Talent Innovation Program of Quanzhou City under Grant 2017G027, in part by the Promotion Program for Young and Middle-aged Teacher in Science and Technology Research of Huaqiao University under Grant ZQN-YX403 and Grant ZQN-PY418, and in part by the High-Level Talent Project Foundation of Huaqiao University under Grant 14BS201, Grant 14BS204, and Grant 16BS108. This paper was recommended by Associate Editor A. Pizurica. *(Corresponding author: Huanqiang Zeng.)*

Y. Fu, H. Zeng, and Z. Ni are with the School of Information Science and Engineering, Huaqiao University, Xiamen 361021, China (e-mail: fuying1511@163.com; zeng0043@hqu.edu.cn; ddkklove@gmail.com).

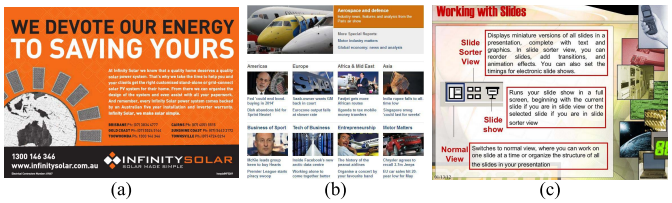
L. Ma is with Tencent AI Lab, Shenzhen 518057, China (e-mail: forest.linma@gmail.com).

J. Zhu is with the School of Engineering, Huaqiao University, Quanzhou 362021, China (e-mail: jqzhu@hqu.edu.cn).

K.-K. Ma is with the School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore 639798 (e-mail: ekkma@ntu.edu.sg).

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/TCSVT.2018.2854176



(a) (b) (c)

Fig. 1. Three SCI examples selected from the SIQAD database for demonstration.

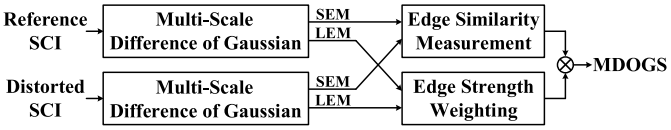


Fig. 2. Block diagram of the proposed SCI quality assessment model.

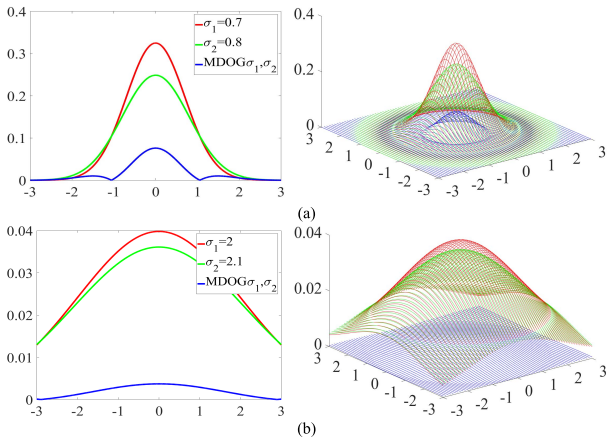


Fig. 3. An illustration of Gaussian and MDOG curves of 1D and 2D at two different scales: (a) Small Scale; (b) Large Scale.

quality score. Experimental results have shown that the proposed IQA metric produces higher consistency with the HVS perception on the evaluation of the SCI quality than the state-of-the-art IQA metrics.

The remaining of this paper is organized as follows. Section II introduces the proposed IQA model for the SCIs in detail. Section III presents the experimental results and discussions. Finally, Section IV provides the conclusions.

II. Proposed IQA Model for SCIs Using Multi-Scale Difference of Gaussian

The proposed IQA model for evaluating the quality of a distorted SCI with respective to its reference SCI is illustrated in Fig. 2. In the first stage, the *multi-scale difference of gaussian* (MDOG) operator is applied to the reference and distorted SCIs to extract the edge maps at two different scales, i.e., *smaller-scale edge map* (SEM) and *larger-scale edge map* (LEM). In the second stage, the *edge similarity map* (ESM) is generated by measuring the similarity between two SEMs. In the last stage, the edge strength is computed based on two LEMs and then used as the weighting factor to generate the final SCI quality score. The details of each stage will be introduced in the following subsections, respectively.

A. Multi-Scale Difference of Gaussian

As discussed in [7]–[9], the HVS presents high sensitivity to the edge information during the image perception. Thus, the edge information, as the most important component in SCI, needs to be deeply investigated for SCI quality assessment. Furthermore, the image details can be better depicted in a multi-scale space, since it can have multiple interpretations of the image content from the coarse level to the fine level [19]. Hence, the *multi-scale difference of gaussian* (MDOG) is designed to extract the edge information of each input SCI at different scales. The extracted edge information at each scale can be expressed in terms of map, called as *edge map* (EM):

$$EM_{\sigma_1,\sigma_2}(x,y) = \left| DOG_{\sigma_1,\sigma_2}(x,y) \otimes I(x,y) \right| \tag{1}$$

where $I(x,y)$ denotes the luminance component at the pixel location (x,y) of the input SCI, the symbol $\otimes$ denotes the convolution operation, and $DOG_{\sigma_1,\sigma_2}(x,y)$ is the difference of two Gaussian kernel with nearby scales σ_1, σ_2:

$$DOG_{\sigma_1,\sigma_2}(x,y) = G(x,y;\sigma_1) - G(x,y;\sigma_2) \tag{2}$$

where $G(x,y;\sigma)$ is the Gaussian kernel function with scale σ:

$$G(x,y;\sigma) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}} \tag{3}$$

It can be seen from equation (1) that the EM of each input SCI is obtained by subtracting its one gaussian blurred version from another less blurred version, which are achieved by convolving the Gaussian kernel with the input SCI using a scale set (σ_1, σ_2). From the spatial viewpoint, the statistical characteristic of the image varies from region to region and thus the image quality degradation is often spatially variant [4]. Hence, the edge information of the input SCI is computed locally. Specifically, the above-described MDOG process is performed pixel by pixel horizontally and vertically over all the rows and columns of the input SCI by using a 7×7 Gaussian square window. By employing different sets of (σ_1, σ_2) in equation (1), the EMs at different scales can be effectively extracted, which will be served as the inputs to the subsequent processing stage for conducting the edge similarity measurement and edge strength weighting.

Fig. 3 shows the Gaussian and MDOG curves of *1 dimension* (1D) and 2D at two scales as examples. One can see that the 2D Gaussian and MDOG kernels are rotational symmetry and thus can effectively consider the frequency responses of the input image in all directions. Moreover, the MDOG using different scales is able to keep a handful of spatial frequencies of the input SCI in different frequency bands, which are corresponding to the edge information of the input SCI in different scales. In other words, this MDOG process is to mimic how neural processing in the retina of the eye extracts details from images destined for transmission to the brain [20]. Hence, the MDOG naturally becomes an efficient solution to explore the edge information that the HVS is sensitive. To demonstrate the ability of MDOG on the extraction of edge information from the SCIs, Fig. 4 shows the *smaller-scale edge map*s (SEMs) and *larger-scale edge map*s (LEMs) of a typical SCI and its distorted version with motion blur, respectively. It can be observed that in the EMs of the distorted SCI extracted by the MDOG, the edge in pictorial region has a significant lose while that in textual region is more messy, compared with that of the reference SCI. Similar observation can be found in other SCIs with various distortions. In other words, the EMs resulted from MDOG can effectively capture the edge degradations from different scales, which is expected to effectively depict the SCI perceptual quality. Therefore, the MDOG is exploited as effective edge description in the proposed SCI quality assessment.

B. Edge Similarity Measurement

Based on the above-mentioned analysis, as shown in Fig. 4, the EMs resulted from MDOG at two scales (i.e., SEM and LEM) are used to conduct the SCI quality assessment in this work. One can further observe from Fig. 4 that the SEM features more shape edges and preserves more image details, while the LEM tends

Fig. 4. An example of *smaller-scale edge map* (SEM) and *larger-scale edge map* (LEM) of a typical SCI and its distorted version. (a) Original SCI; (d) the corresponding distorted SCI with motion blur; (b) and (e) are the corresponding SEMs of (a) and (d); (c) and (f) are the corresponding LEMs of (a) and (d).

to be more smooth and mainly contains the contour information. Consequently, the SEM is more suitable for measuring the degree of edge information preservation (ie., edge similarity) between the reference and distorted SCIs. On the contrary, the LEM is more proper to be a weighting factor to control the importance of each pixel in the evaluation of the overall SCI quality assessment, since the HVS is sensitive to the image contour [16].

For ease of notation, let r and d denote the reference SCI and its distorted version. The degree of *edge similarity* (ES) can be computed by following the commonly-used practice adopted in many IQA models (e.g., [4], [8], [9]), as follows:

$$ES(x, y) = \frac{2SEM_r(x, y) \cdot SEM_d(x, y) + T}{SEM_r^2(x, y) + SEM_d^2(x, y) + T} \tag{4}$$

where the SEM_r and SEM_d are the EMs resulted from performing the operation in equation (1) on the reference and distorted SCIs with a smaller scale set (σ_1, σ_2) and T is a small positive constant to prevent numerical instability.

C. Edge Strength Weighting

After obtaining the $ES(x, y)$ that reflects the edge similarity at each pixel location (x, y), the overall quality score of the SCI can be obtained via a weighting strategy to consider the perceptual importance of different image pixels, since all image pixels are not equally perceived by the HVS. From the viewpoint of HVS, the visual resolution decays spatially from a spot in the human retina [21], and the fixation point in the image tends to be significantly different from its neighbours [22]. In other words, the HVS is more sensitive to those pixels in the image contour, which would be effectively captured by the LEM, as shown in Fig. 4. Therefore, it is meaningful to design a simple and yet efficient weighting factor for each pixel location based on the LEMs. If the pixel location (x, y) of the LEMs of either the reference or distorted SCIs yields a larger value, it means that this position has stronger *edge strength* and the HVS will thus be more sensitive to this position. Therefore, the *edge strength weighting* is developed by selecting the maximum value in two LEMs of reference and distorted SCIs as the weighting factor:

$$W(x, y) = \max(LEM_r, LEM_d) \tag{5}$$

where the LEM_r and LEM_d are the EMs produced by performing the operation in equation (1) on reference and distorted SCIs with a larger scale set (σ_1, σ_2).

Finally, the proposed IQA score for SCI, called as the *multi-scale difference of Gaussian score* (MDOGS), can be pooled on the edge similarity map by using the edge strength weighting, as follows:

$$MDOGS = \frac{\sum_{(x,y)\epsilon\Omega} ES(x, y) \cdot W(x, y)}{\sum_{(x,y)\epsilon\Omega} W(x, y)} \tag{6}$$

where Ω denotes all the pixel locations of the SCI.

TABLE I

PERFORMANCE COMPARISON BETWEEN THE PROPOSED MDOGS MODEL AND THE STATE-OF-THE-ART IQA METRICS IN TERMS OF PLCC, SROCC, AND RMSE ON THE SIQAD DATABASE. THE TOP THREE BEST PERFORMANCES ARE HIGHLIGHTED WITH THE BOLD TYPE

Criteria	Distortions	PSNR	SSIM [4]	MSSIM [5]	IWSSIM [23]	VIF [11]	IFC [12]	MAD [24]	FSIM [25]	ES [7]	GMSD [9]	SPQA [2]	GSS [16]	EMSQA [18]	SIQM [15]	SQI [17]	$MDOGS$
PLCC	GN	**0.9053**	0.8806	0.8783	0.8804	**0.9011**	0.8791	0.8852	0.7428	0.8167	0.8956	0.8921	0.8645	0.8889	0.892	0.8829	**0.8982**
	GB	0.8603	0.9014	0.8984	0.9079	0.9102	0.9061	0.9120	0.7206	0.8653	0.9094	0.9058	0.9073	**0.9156**	0.912	**0.9202**	**0.9195**
	MB	0.7044	0.8060	0.8240	0.8414	**0.8490**	0.6782	0.8361	0.6874	0.7873	0.8436	0.8315	0.8314	**0.8753**	0.845	**0.8789**	0.8421
	CC	0.7401	0.7435	**0.8371**	**0.8404**	0.7076	0.6870	0.3933	0.7507	0.8005	0.7827	0.7992	0.6091	0.7688	0.790	0.7724	**0.8011**
	JPEG	0.7545	0.7487	0.7756	**0.7998**	**0.7986**	0.7606	0.7662	0.5566	0.6065	0.7746	0.7696	0.7948	0.7904	0.771	**0.8218**	0.7885
	J2K	0.7893	0.7749	0.7951	0.8040	0.8205	0.7963	**0.8344**	0.6675	0.6717	**0.8509**	0.8252	0.8130	0.7850	0.794	0.8271	**0.8606**
	LSC	0.7805	0.7307	0.7729	0.8155	**0.8385**	0.7679	0.8184	0.5964	0.6569	**0.8559**	0.7958	0.8084	0.7747	0.720	0.8310	**0.8316**
	Overall	0.5869	0.7561	0.6161	0.6527	0.8198	0.6395	0.6191	0.5389	0.5273	0.7387	0.8584	0.8461	**0.8648**	0.852	**0.8644**	**0.8839**
SROCC	GN	0.8790	0.8694	0.8679	0.8743	**0.8888**	0.8717	0.8721	0.7373	0.8295	**0.8856**	0.8823	0.8521	0.8745	0.871	0.8602	**0.8882**
	GB	0.8573	0.8921	0.8883	0.9060	0.9059	0.9106	0.9087	0.7286	0.8922	0.9119	0.9017	0.9053	**0.9154**	0.910	**0.9244**	**0.9192**
	MB	0.7130	0.8041	0.8238	0.8421	**0.8492**	0.6737	0.8357	0.6641	0.8083	0.8441	0.8255	0.8397	**0.8788**	0.840	**0.8810**	0.8345
	CC	0.6828	0.6405	**0.7506**	**0.7563**	0.6433	0.6396	0.3907	**0.7175**	0.6833	0.6378	0.6154	0.5974	0.6306	0.705	0.6677	0.6644
	JPEG	0.7569	0.7576	0.7787	**0.7978**	0.7924	0.7636	0.7674	0.5879	0.6543	0.7712	0.7673	0.7969	**0.7871**	0.775	**0.8189**	0.7856
	J2K	0.7746	0.7603	0.7855	0.7998	0.8131	0.7980	**0.8382**	0.6363	0.6824	**0.8436**	0.8152	0.8141	0.7762	0.777	0.8169	**0.8622**
	LSC	0.7930	0.7371	0.7711	0.8214	**0.8463**	0.7713	0.8154	0.5979	0.6899	**0.8592**	0.8003	0.8164	0.7798	0.725	0.8432	**0.8513**
	Overall	0.5604	0.7566	0.6115	0.6545	0.8065	0.6011	0.6067	0.5279	0.5637	0.7305	0.8416	0.8359	**0.8504**	0.845	**0.8548**	**0.8822**
RMSE	GN	**6.3372**	7.0679	7.1309	7.0744	**6.4673**	7.1096	6.9391	9.9860	8.6084	6.6354	6.7394	7.497	6.8320	7.0165	-	**6.5576**
	GB	7.7376	6.5701	6.6638	6.3619	6.2859	6.4193	6.2269	10.5230	7.6073	6.3111	6.4301	6.383	**6.0710**	**5.8367**	-	**5.9639**
	MB	9.2287	7.6967	7.3675	7.0260	**6.8704**	9.5544	7.1322	9.4432	8.0165	6.9816	7.2223	7.225	**6.2880**	**6.0869**	-	7.0121
	CC	8.4591	8.4116	**6.8818**	**6.8184**	8.8876	9.1407	11.5652	8.3190	7.5395	7.8297	7.6184	9.976	8.0440	8.1079	-	**7.5284**
	JPEG	6.1665	6.2295	5.9311	**5.6406**	**5.6551**	6.1004	6.0380	7.8072	7.4712	5.9427	6.0000	5.702	5.7560	**5.6548**	-	5.7787
	J2K	6.3819	6.5691	6.3040	6.1804	5.9412	6.2875	5.7276	7.7404	7.6998	**5.4591**	**5.8706**	6.051	6.4390	6.0820	-	**5.2930**
	LSC	5.3336	5.8253	5.4141	4.9379	**4.6497**	5.4657	4.9025	6.8486	6.4327	**4.4121**	5.1664	5.022	5.3950	5.3576	-	**4.7382**
	Overall	11.5898	9.3676	11.2744	10.8444	8.1969	11.0048	11.2409	12.0583	12.1634	9.6484	7.3421	7.631	**7.1860**	7.4936	**7.1982**	**6.6951**

III. EXPERIMENTAL RESULTS AND DISCUSSIONS

A. SCI Database and Evaluation Protocols

In this section, we validate the performance of the proposed model on the public available *SIQAD* database provided by [2]. This SIQAD database is specifically designed for comprehensively evaluating the perceptual quality of the SCIs, which contains 20 reference images and 980 distorted SCIs with 7 types of distortions at 7 different levels. Those distortion types include Gaussian noise (GN), Gaussian blur (GB), motion blur (MB), contrast change (CC), JPEG compression, JPEG2000 (J2K) compression, and layer segmentation based coding (LSC).

To meaningfully compare the consistency between the subjective evaluation scores (i.e., MOS/DMOS) and the objective scores s resulted from various IQA models, a mapping process suggested in VQEG HDTV test [26] is performed to map the dynamic range of the scores produced from each model into a common scale, as follows.

$$Q_i = \beta_1 \left(\frac{1}{2} - \frac{1}{1 + e^{\beta_2 (s_i - \beta_3)}} \right) + \beta_4 s_i + \beta_5 \tag{7}$$

where s_i is the computed quality score of the i-th distorted SCI resulted from an IQA model, Q_i is the corresponding mapped score, and β_1, β_2, β_3, β_4 and β_5 are the regression model parameters to be fitted by minimizing the sum of squared differences between the mapped objective scores $Q(s)$ and the subjective evaluation scores (i.e., MOS/DMOS) (for details, refer to [26]). After mapping, the performance of various IQA models are measured in terms of three commonly-used performance indexes as suggested in [26], 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 error* (RMSE) for prediction consistency. Note that the higher values of PLCC and SROCC provide better accuracy and prediction monotonicity, while a smaller RMSE value indicates a better performance.

B. Performance Comparison

To evaluate the performance, we compare the proposed MDOGS model with the classical and state-of-the-art IQA models, including PSNR, SSIM [4], MSSIM [5], IWSSIM [23], VIF [11], IFC [12], MAD [24], FSIM [25], ES [7], GMSD [9], SPQA [2], GSS [16], EMSQA [18], SIQM [15] and SQI [17]. Note that the latter five methods are particularly designed for the IQA of the SCIs. The corresponding results are shown in Table I, where the best three performances are highlighted in bold type. Moreover, the PLCC, SROCC and RMSE values of all the above-mentioned metrics are reported except the RMSE values of SQI, since its program code is not available. It is worth to noting that the parameters of the proposed IQA model, including different scale sets (σ_1, σ_2) for the SEM and LEM, T, the Gaussian window size, are determined based on a subset of SIQAD database, which contains 4 reference SCIs and their corresponding 196 distorted SCIs. Following the same practices as suggested in [25] and [27], those parameter values, leading to higher SROCC, will be selected. Consequently, (σ_1, σ_2) = (0.7, 0.8) for the SEM, (σ_1, σ_2) = (2.0, 2.1) for the LEM, T = 0.04, and the Gaussian window size = 7×7 are empirically determined through extensive experiments.

Firstly, it can be clearly seen from the results that the proposed model is able to achieve the highest *overall* PLCC and SROCC values and lowest *overall* RMSE values. It indicates that the proposed model is mostly consistent with the subjective judgment made by the HVS, compared with the state-of-the-art IQA models. Moreover, it is interesting to find that all the IQA models particularly developed for the SCIs (i.e., SPQA, GSS, EMSQA, SIQM, SQI, and the proposed model) outperform those IQA methods designed for the natural images. This is because they design the SCI IQA models by taking the special characteristics of the SCIs into account. Furthermore, compared with other state-of-the-art edge-based IQA models (i.e., ES, GMSD, GSS, and EMSQA), the proposed model also can obtain superior performance. This might be due to that the proposed model

using multi-scale difference of Gaussian is much more effective to explore the edge information in different scales for SCI quality assessment.

In addition, Table I also shows the performances on individual distortion to comprehensively evaluate each IQA model's ability on assessing image quality's degradations caused by individual distortion. It can be found that the proposed model obtains the most top-three performances on 7 distortions compared with other IQA models. Specifically, in the comparisons in terms of PLCC, the proposed model is among the top-three models 5 times, followed by VIF (4 times) and SQI (3 times). In the comparisons in terms of SROCC, the proposed model is among the top-three models 4 times, followed by VIF, GMSD, EMSQA, and SQI (3 times). In the comparisons in terms of RMSE, the proposed model is among the top-three models 5 times, followed by VIF (4 times) and SIQM (3 times). Moreover, it can be observed that the proposed model performs competitively on the distortions of GN, GB, and compression. This is because the proposed IQA model can accurately reflect the image structure change resulted from the distortions or artifacts caused by the noise, blurring and compression. Meanwhile, the proposed model does not perform well on CC. This is because the CC mainly influences the image intensity rather than the image structure. Future work will further consider the image intensity changes on the design of SCI quality assessment model.

IV. Conclusion

In this paper, we develop a novel IQA model for the SCIs by using multi-scale difference of Gaussian. The novelty of the proposed model lies in the use of MDOG to effectively explore the edge information from two scales for conducting the SCI quality assessment. Specifically, the smaller-scale edge maps of the reference and distorted SCIs are used to measure their edge similarity, which is further fused using edge strengthen weighting designed based on their larger-scale edge maps to generate the final SCI quality score. Experimental results on the benchmark database have demonstrated that the proposed IQA model is superior to multiple state-of-the-art IQA models.

References

[1] S. Wang, K. Gu, S. Ma, and W. Gao, "Joint chroma downsampling and upsampling for screen content image," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 26, no. 9, pp. 1595–1609, Sep. 2016.

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

[3] W. Lin and C.-C. J. Kuo, "Perceptual visual quality metrics: A survey," *J. Vis. Commun. Image Represent.*, vol. 22, no. 4, pp. 297–312, 2011.

[4] 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.

[5] Z. Wang, E. P. Simoncelli, and A. C. Bovil, "Multi-scale structural similarity for image quality assessment," in *Proc. IEEE Conf. Signals Syst. Comput.*, vol. 2, Nov. 2003, pp. 1398–1402.

[6] J. Qian, D. Wu, L. Li, D. Cheng, and X. Wang, "Image quality assessment based on multi-scale representation of structure," *Digit. Signal Process.*, vol. 33, pp. 125–133, Oct. 2014.

[7] X. Zhang, X. Feng, W. Wang, and W. Xue, "Edge strength similarity for image quality assessment," *IEEE Signal Process. Lett.*, vol. 20, no. 4, pp. 319–322, Apr. 2013.

[8] A. Liu, W. Lin, and M. Narwaria, "Image quality assessment based on gradient similarity," *IEEE Trans. Image Process.*, vol. 21, no. 4, pp. 1500–1512, Apr. 2012.

[9] W. Xue, L. Zhang, X. Mou, and A. C. Bovik, "Gradient magnitude similarity deviation: A highly efficient perceptual image quality index," *IEEE Trans. Image Process.*, vol. 23, no. 2, pp. 684–695, Feb. 2014.

[10] L. Zhang, Y. Shen, and H. Li, "VSI: A visual saliency-induced index for perceptual image quality assessment," *IEEE Trans. Image Process.*, vol. 23, no. 10, pp. 4270–4281, Aug. 2014.

[11] 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.

[12] 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.

[13] F. Gao and J. Yu, "Biologically inspired image quality assessment," *Signal Process.*, vol. 124, pp. 210–219, Jul. 2016.

[14] C. Deng and D. Tao, "Color image quality assessment with biologically inspired feature and machine learning," *Vis. Commun. Image Process.*, vol. 124, pp. 77440Y-1–77440Y-7, Aug. 2010.

[15] K. Gu, S. Wang, G. Zhai, S. Ma, and W. Lin, "Screen image quality assessment incorporating structural degradation measurement," in *Proc. IEEE Int. Symp. Circuits Syst.*, May 2015, pp. 125–128.

[16] Z. Ni, L. Ma, H. Zeng, C. Cai, and K.-K. Ma, "Gradient direction for screen content image quality assessment," *IEEE Signal Process. Lett.*, vol. 23, no. 10, pp. 1394–1398, Oct. 2016.

[17] S. Wang, K. Ma, W. Wang, H. Yeganeh, Z. Wang, and W. Lin, "Objective quality assessment and perceptual compression of screen content images," *IEEE Comput. Graph. Appl.*, vol. 38, no. 1, pp. 47–58, Jan./Feb. 2018.

[18] Z. Ni, L. Ma, H. Zeng, C. Cai, and K.-K. Ma, "Screen content image quality assessment using edge model," in *Proc. IEEE Int. Conf. Image Process.*, Aug. 2016, pp. 81–85.

[19] T. Lindeberg, "Scale-space theory: A basic tool for analysing structures at different scales," *J. Appl. Statist.*, vol. 21, nos. 1–2, pp. 224–270, Jun. 1994.

[20] S. Nilufar, N. Ray, and H. Zhang, "Object detection with DOG scale-space: A multiple kernel learning approach," *IEEE Trans. Image Process.*, vol. 21, no. 8, pp. 3744–3756, Aug. 2012.

[21] A. L. N. T. Costa and M. N. Do, "A retina-based perceptually lossless limit and a Gaussian foveation scheme with loss control," *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 3, pp. 438–453, Jun. 2014.

[22] Y. Fang, W. Lin, B.-S. Lee, C.-T. Lau, Z. Chen, and C.-W. Lin, "Bottom-up saliency detection model based on human visual sensitivity and amplitude spectrum," *IEEE Trans. Multimedia*, vol. 14, no. 1, pp. 187–198, Feb. 2012.

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

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

[25] L. Zhang, L. Zhang, X. Mou, and D. Zhang, "FSIM: A feature similarity index for image quality assessment," *IEEE Trans. Image Process.*, vol. 20, no. 8, pp. 2378–2386, Aug. 2011.

[26] Video Quality Experts Group. (Aug. 2015). *Final Report From the Video Quality Experts Group on the Validation of Objective Models of Video Quality Assessment*. [Online]. Available: http://www.its.bldrdoc.gov/vqeg/vqeg-home.aspx

[27] S.-H. Bae and M. Kim, "A novel image quality assessment with globally and locally consilient visual quality perception," *IEEE Trans. Image Process.*, vol. 25, no. 5, pp. 2392–2406, May 2016.