

No Reference Quality Assessment for Screen Content Images With Both Local and Global Feature Representation

Yuming Fang[✉], Senior Member, IEEE, Jiebin Yan, Leida Li, Member, IEEE,
Jinjian Wu[✉], Member, IEEE, and Weisi Lin, Fellow, IEEE

Abstract—In this paper, we propose a novel no reference quality assessment method by incorporating statistical luminance and texture features (NRLT) for screen content images (SCIs) with both local and global feature representation. The proposed method is designed inspired by the perceptual property of the human visual system (HVS) that the HVS is sensitive to luminance change and texture information for image perception. In the proposed method, we first calculate the luminance map through the local normalization, which is further used to extract the statistical luminance features in global scope. Second, inspired by existing studies from neuroscience that high-order derivatives can capture image texture, we adopt four filters with different directions to compute gradient maps from the luminance map. These gradient maps are then used to extract the second-order derivatives by local binary pattern. We further extract the texture feature by the histogram of high-order derivatives in global scope. Finally, support vector regression is applied to train the mapping function from quality-aware features to subjective ratings. Experimental results on the public large-scale SCI database show that the proposed NRLT can achieve better performance in predicting the visual quality of SCIs than relevant existing methods, even including some full reference visual quality assessment methods.

Index Terms—Screen content image, visual quality assessment, no reference quality assessment, local feature, global feature.

I. INTRODUCTION

RECENTLY, more and more screen content images (SCIs) emerge over Internet, and they are widely used as the

medium of signal transmission in many multimedia applications. Currently, many relevant algorithms for SCI have been proposed, including SCI segmentation [4], SCI compression [5], SCI quality assessment [16], [53], [54], etc. Undoubtedly, visual quality of SCIs has a great influence on users' viewing experience. Generally, visual quality of SCIs might be degraded with various distortion types during image processing including acquisition, transmission and coding [1], [2]. Therefore, it is highly desired to design effective visual quality assessment (VQA) metrics for SCIs in various SCI processing applications.

There have been a number of visual quality assessment (VQA) methods devised in recent years. Traditional VQA methods such as mean square error (MSE) and peak signal to noise (PSNR) have been widely used in industry due to their simple implementation. However, the predicted quality results from these methods are not consistent with subjective ratings, since these methods predict visual quality of images without considering the characteristics of the human visual system (HVS). In order to overcome the shortcomings of these methods, many advanced full reference (FR) methods which require complete reference information for image quality prediction with consideration of the characteristics of the HVS have been proposed, including structure similarity (SSIM) [6] and its variants: multi-scale SSIM (MSSIM) [7], information content weighted SSIM (IWSSIM) [8], feature similarity (FSIM) [9], internal generative mechanism (IGM) [10], visual saliency-based index (VSI) [11], gradient magnitude similarity deviation (GMSD) [12].

There have been also many reduced reference (RR) [22], [24], [48], [49], and no reference (NR) image quality estimation methods [13], [14], [17], [42], [43], [46], [47] proposed in the past decades. For RR methods, only part of reference information is required for image quality prediction, and global features are always extracted as the indicator of image quality. For NR methods, they do not require any reference information for VQA. Fang *et al.* [43] built nature scene statistics (NSS) models for NR VQA of contrast distorted images based on moment features. There have been several NR metrics [44], [45], [57] proposed based on one type of local structure descriptor: the local binary pattern (LBP) [38]. Zhang *et al.* designed a NR metric based on the joint generalized local binary pattern (GLBP) statistics in [44], and Freitas *et al.* proposed a NR model based on texture information extracted

Manuscript received June 17, 2017; revised October 7, 2017 and November 24, 2017; accepted November 26, 2017. Date of publication December 8, 2017; date of current version January 5, 2018. This work was supported in part by the National Natural Science Foundation of China under Grant 61571212 and Grant 61772388, and in part by the Natural Science Foundation of Jiangxi, China, under Grant 20071BBE50068, Grant 20171BCB23048, and Grant GJJ160420. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Stefan Winkler. (Corresponding author: Leida Li.)

Y. Fang and J. Yan are with the School of Information Technology, Jiangxi University of Finance and Economics, Nanchang 330032, China (e-mail: fa0001ng@e.ntu.edu.sg).

L. Li is with the School of Information and Control Engineering, China University of Mining and Technology, Xuzhou 221116, China (e-mail: reader1104@hotmail.com).

J. Wu is with the School of Electronic Engineering, Xidian University, Xi'an 710126, China.

W. Lin is with the School of Computer Science and Engineering, Nanyang Technological University, Singapore 639798.

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TIP.2017.2781307

from multiple local ternary pattern (LTP) channels in the form of histograms [45]. Li *et al.* designed a NR VQA model to calculate structure degradation based on the perceptual property in the HVS that there are separate mechanisms in human visual cortex to process the first- and second-order patterns [57]. Zhang *et al.* [58] proposed NSS models with the consideration of fitting parameters of NSS models and the fitting errors between real distributions and fitted NSS models as well as the likelihood of fitting error. In natural image quality evaluator (NIQE) [17], the authors designed a NR VQA metric by using the deviations from the nature scene statistic model in natural images. Zhang *et al.* recently proposed the integrated-local NIQE (IL-NIQE) based on multivariate Gaussian model with three additional statistical features based on NIQE [46]. Xue *et al.* [56] designed a quality-aware clustering approach named QAC to learn a series of centroids, which are then used to calculate the visual quality scores of image patches. The overall perceptual quality of distorted image is obtained by average pooling of the quality scores of image patches.

Most existing VQA methods are designed for natural images and they are not effective in visual quality prediction of SCIs. Unlike natural images, SCIs include diverse forms of visual content, such as pictorial and textual regions. As shown in our previous study [16], the characteristics of SCIs and those of natural images are different greatly. Generally, the mean subtracted contrast normalized (MSCN) coefficients [39] extracted from natural images follow generalized Gaussian distribution, while the distribution curve of MSCN coefficients from SCIs fluctuates greatly and a sharp pimpling appears [16]. Thus, the VQA metrics designed for quality prediction of natural images are not effective for VQA of SCIs.

Recently, there have been some metrics designed for visual quality prediction of SCIs. Yang *et al.* [16] conducted an user study for visual quality prediction of SCIs and presented a detailed analysis of the subjective experiment. In addition, they built an objective quality assessment metric for SCIs. Wang *et al.* [19] designed a FR VQA metric for SCIs by incorporating local information content weighting and viewing field adaption. Fang *et al.* proposed a FR objective VQA model for SCIs based on uncertainty weighting with the consideration that the HVS is more sensitive to high-frequency information than other smooth regions in the visual scene [20]. Gu *et al.* designed a structure-induced quality method (SIQM) by weighting the quality map with the proposed structural degradation strategy [21]. Ni *et al.* proposed a VQA model for SCIs based on gradient direction denoted by GSS [53] and edge information denoted by EMSQA [54] and ESIM [55]. In the study Wang *et al.* [22] designed a RR image quality assessment (RR-IQA) metric for SCIs by considering visual perception of SCIs. Wang *et al.* [59], conducted studies on subjective quality evaluation of the compressed SCIs and designed a RR quality prediction model in wavelet domain. These methods mentioned above require the complete information of reference image. However, the original images are not available in most practical applications. Gu *et al.* proposed a NR quality estimation method for SCIs (BQMS) with structural degradation model and the free energy based perceptual theory, and established a large scale database to

train NSS models [23]. Shao *et al.* [25] proposed a blind quality predictor for SCI (BLIQUP-SCI) from the perspective of sparse representation. These two NR VQA metrics designed for SCIs cannot obtain consistent visual quality estimation results with human perception. It is still challenging to design effective NR VQA metrics for SCIs.

It is well known that the HVS is sensitive to edge information [26]–[28], which is the basic component in textual regions and high-frequency parts of the pictorial regions in SCIs. Thus, we use the edge information to measure the visual distortion in SCIs. Meanwhile, the variation of luminance information is also highly correlated with the visual quality of visual scenes [29], [30]. In this study, we propose a novel no reference visual quality estimation method for SCIs by statistical luminance features and texture information (NRLT) with both local and global feature representation. First, we extract luminance map of SCIs through the local normalization to the original intensity map to calculate the statistical luminance features. Then four filters with different directions are adopted to compute the gradient maps based on the luminance map, which are denoted as the first order derivative information. The texture features are calculated by local binary patterns (LBPs) on the gradient maps which are denoted as the second order derivative information. We further utilize the histogram to represent statistical luminance and texture features in global scope. Support vector regression (SVR) is employed to train the visual quality estimation model from the quality-aware features to subjective ratings. Experimental results on a large-scale SCI database show that the proposed NRLT can obtain better performance in estimating the perceptual quality of SCIs than other relevant ones, even including some FR methods.

There is much difference between the proposed method and other existing ones. In RR-IQA [22], the authors first computed the significance maps of original SCI and distorted versions by incorporating the primary visual information and the uncertainty measure, and utilized the histogram to represent the combined significance maps [22]. The final quality score of the distorted SCI is obtained by comparing these two histograms. Compared with that method [22], the proposed method extract statistical luminance and texture features without any reference information to perceive the quality degradation of distorted SCIs. In BLIQUP-SCI [25], the authors used four FR methods including PSNR, SSIM, FSIM and VIF [31] to generate the labels for SCIs. However, these four FR methods are with poor performance in predicting the visual quality of SCIs [16]. In the proposed method, we use the subjective quality scores of SCIs as the labels to train the visual quality prediction model. Meanwhile, BQMS [23] needs to fit the relationship of the structure degradation features and the free energy feature before extracting the quality-aware features, while we use the histogram to represent the statistical luminance and texture features directly.

II. PROPOSED METHOD

In this section, we describe the proposed NRLT for SCIs in detail. As shown in the study [16], the luminance change and texture variation influence the visual quality of SCIs greatly

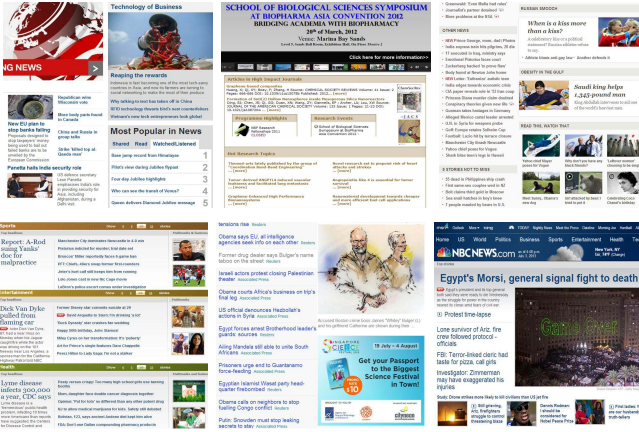


Fig. 1. The visual examples of SCIs [16].

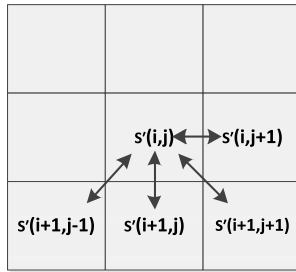


Fig. 2. Pairwise products computed along different directions: horizontal, main-diagonal, vertical, and secondary diagonal.

as well as that for natural images [6]. Meanwhile, the main difference between SCIs and natural images is that there are textual regions in SCIs as shown in Fig. 1, and the human attention is easily attracted to understand the meanings of the words. In this study, the luminance and texture features are employed to capture the perceptual distortion in SCIs. First, we apply the local contrast normalization to SCIs to eliminate the redundant information in visual scenes [39]. Then the statistical luminance features are represented in the form of histogram. The four filters with different directions as defined in Fig. 3 are adopted to compute gradient maps, which are then used to extract the second order derivatives by employing LBP descriptor. Furthermore, we extract the texture information by the second order derivatives. The feature vector is composed of statistical luminance and texture features calculated in three scales. Finally, SVR is adopted to train the quality estimation model from the quality-aware feature vector to subjective ratings.

A. Feature Extraction

As introduced above, we extract luminance and texture features to capture the variations of SCIs. Here, we extract the statistical luminance features from the normalized luminance map. The normalization operation is conducted as [17]:

$$S'(i, j) = \frac{S(i, j) - \mu_s}{\sigma_s + C} \quad (1)$$

where $S(i, j)$ and $S'(i, j)$ denote the original and normalized values at location (i, j) ; $i \in \{1, 2, \dots, I\}$ and

0	0	0	0	0	0	0	1	0	0	0	0	1	0	0	0	1	0	-1	0		
1	3	8	3	1	0	0	8	3	0	0	0	0	3	8	0	0	3	0	-3	0	
0	0	0	0	0	0	1	3	0	-3	-1	-1	-3	0	3	1	0	8	0	-8	0	
-1	-3	-8	-3	-1	0	0	0	-3	-8	0	0	0	-8	-3	0	0	0	3	0	-3	0
0	0	0	0	0	0	0	0	-1	0	0	0	0	0	-1	0	0	0	1	0	-1	0
g_1					g_2					g_3					g_4						

Fig. 3. Filters for calculating the gradient maps.

$j \in \{1, 2, \dots, J\}$ represent the spatial indices; μ_s and σ_s represent the mean and standard variance values of local region; C is a constant. μ_s and σ_s can be computed as:

$$\mu_s = \sum_{m=-M}^M \sum_{h=-H}^H \omega_{(m,h)} S(i+m, j+h) \quad (2)$$

$$\sigma_s = \sqrt{\sum_{m=-M}^M \sum_{h=-H}^H \omega_{(m,h)} [S(i+m, j+h) - \mu_s]^2} \quad (3)$$

where $\{\omega_{(m,h)} | m = -M, \dots, M; h = -H, \dots, H\}$ defines a unit-volume Gaussian window; M and H are set to 3.

As shown in [39], the transformed luminance $S'(i, j)$ follows the generalized Gaussian distribution (GGD) for natural images, while it is inconsistent for SCIs [16]. Thus, after local normalization, we directly employ absolute value of the transformed luminance $S'(i, j)$ to represent the global luminance information in the form of histogram. Here, the number of the histogram bins is set to 10 and we can obtain 10 elements $\{f_1, f_2, \dots, f_{10}\}$ to represent statistical luminance features. The histogram is calculated as below:

$$f_n = \frac{1}{IJ} \sum_{i=1}^I \sum_{j=1}^J \Omega(|S'(i, j)|, Q(n)) \quad (4)$$

$$\Omega(a, b) = \begin{cases} 1, & a \in b \\ 0, & \text{otherwise} \end{cases} \quad (5)$$

where n is the bin index of the histogram in the range of $[1, 10]$; $Q(n)$ denotes the interval of each bin; I and J are the image height and width, respectively.

We have provided some luminance map samples in Fig. 4. As shown in Fig. 4, the changes of luminance maps of distorted SCIs degraded by different distortion types are different from each other compared with the luminance map of the original SCI. For Gaussian noise (GN) distortion, many white points appear in the associated luminance map randomly due to random distribution of GN. For Gaussian blur (GB) distorted SCIs, the edge information is preserved and the detailed structure information is destroyed, which can be clearly reflected by the luminance map and the high peak appearing in the corresponding distribution. Different from GB distorted SCIs, part of detailed structure information is preserved in motion blur (MB) distorted SCI. For contrast change (CC) distorted SCIs, the variation in the luminance map is smaller than those with other distortion types. And the corresponding luminance distribution of CC distorted SCI is similar to that of the original SCI. The reason might be that the distortion of contrast change always occurs in global scope and the local normalization is not sensitive to this type of distortion relatively. Meanwhile, the block artifacts are

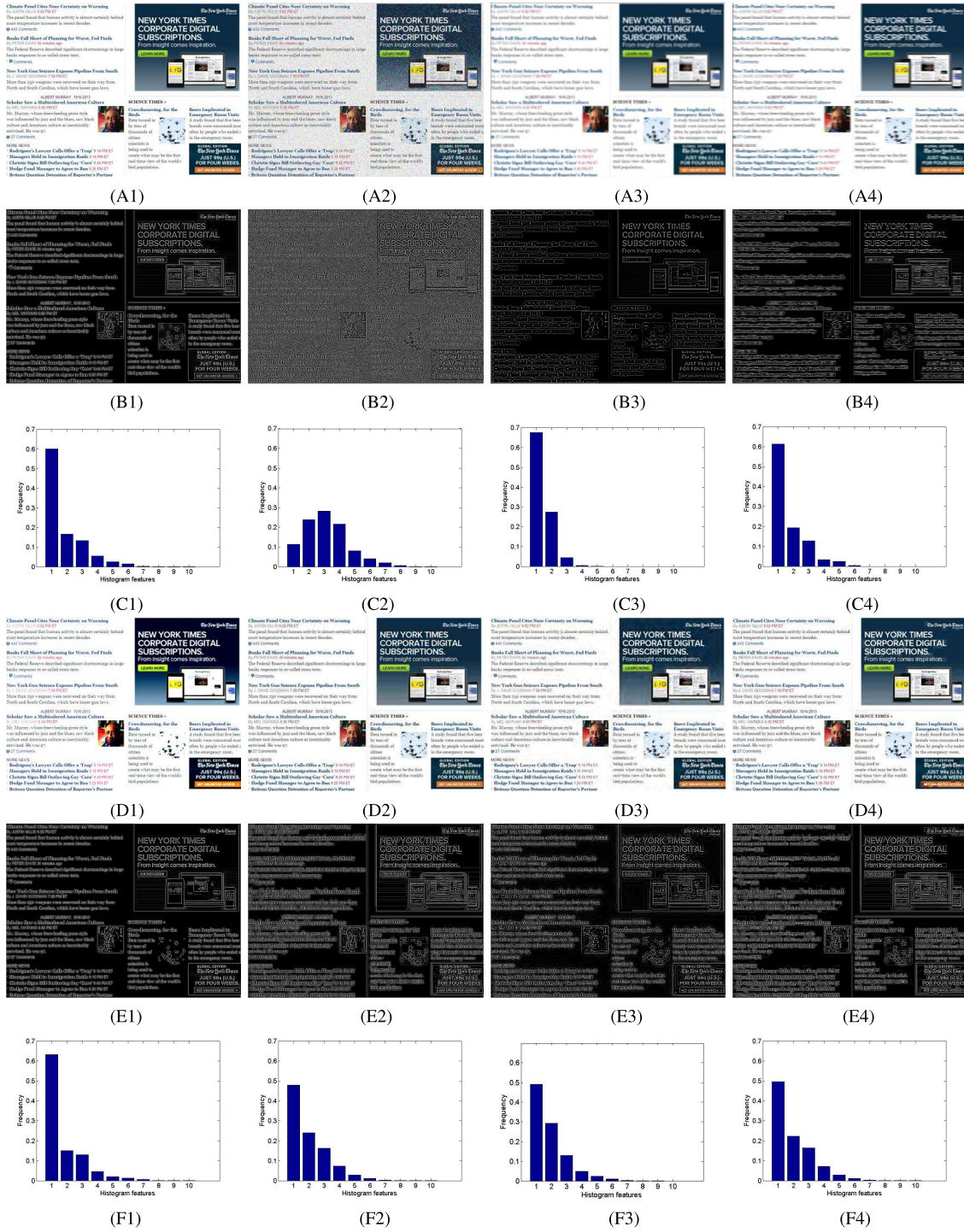


Fig. 4. The visual samples of feature maps. The first and fourth row show SCIs; the second and fifth rows show the corresponding luminance maps; the third and sixth rows show the corresponding luminance histograms. (A1) is the reference SCI. (A2), (A3), (A4), (D1), (D2), (D3), and (D4) represent distorted SCIs by GN, GB, MB, CC, JPEG, JPEG2000, and Layer Segmentation Based Coding.

clear in the luminance maps of SCIs distorted by compression distortion. Thus, luminance features are effective in capturing the quality variation of distorted SCIs with different distortion types.

Considering that the visual content of visual scenes could be well represented by orientation information [32], [33], and local orientation information in visual scenes is important to visual perception [34], we also extract the luminance information by using the histogram features of pairwise

products of neighboring MSCN coefficients along different directions: horizontal (M_H), main-diagonal (M_{D1}), vertical (M_V), and secondary-diagonal (M_{D2}), as illustrated in Fig. 2. The pairwise products are calculated as follows:

$$M_H(i, j) = S(i, j)'S(i, j + 1)' \quad (6)$$

$$M_{D1}(i, j) = S(i, j)'S(i + 1, j + 1)' \quad (7)$$

$$M_V(i, j) = S(i, j)'S(i + 1, j)' \quad (8)$$

$$M_{D2}(i, j) = S(i, j)'S(i + 1, j - 1)' \quad (9)$$

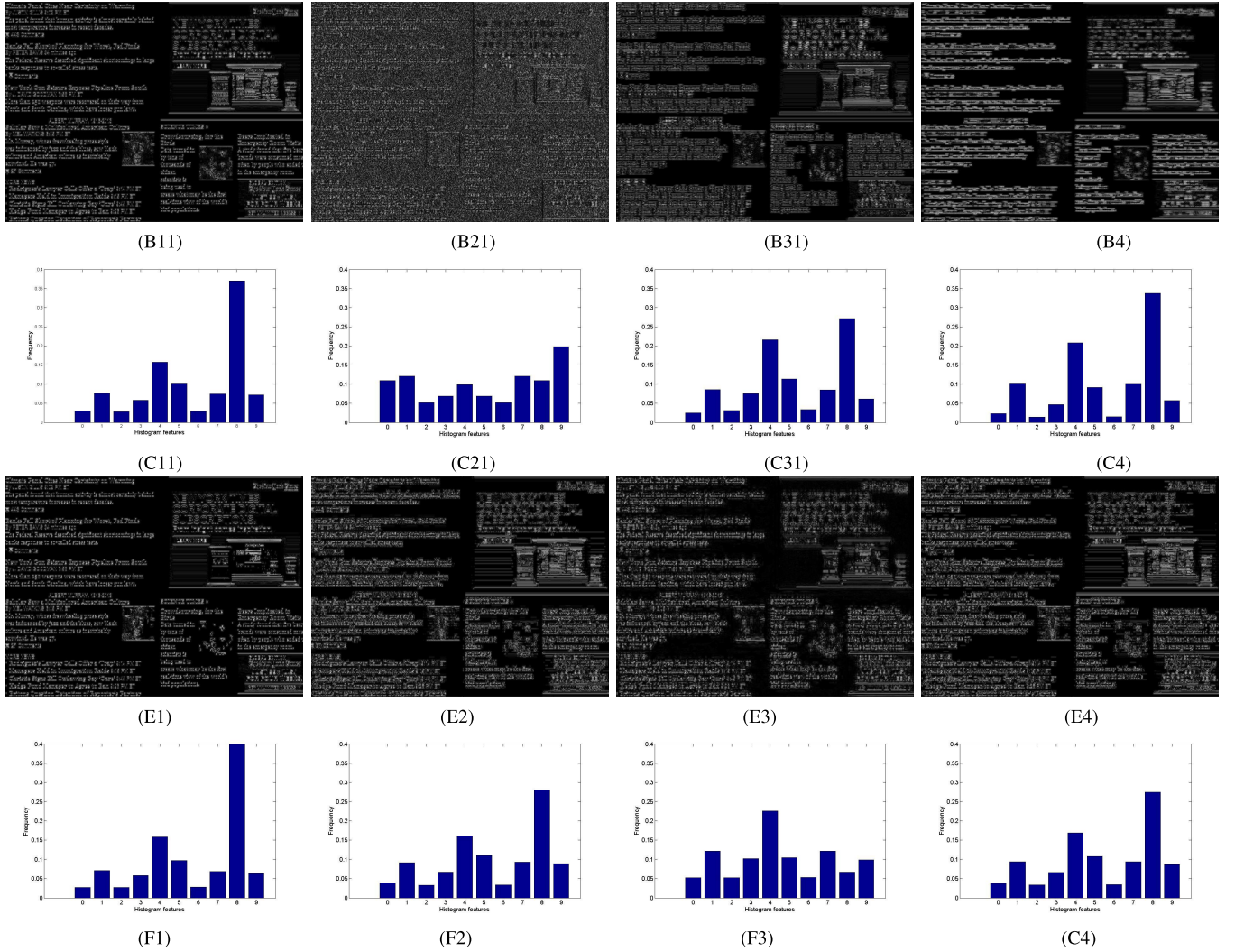


Fig. 5. The visual samples of feature maps. The first and second rows show the corresponding gradient maps and distributions of (A1-A4) illustrated in Fig. 4, the third and fourth rows show the the corresponding gradient maps and distributions of (D1-D4) illustrated in Fig. 4.

The similar operations as described in Eqs. (4) and (5) are conducted based on M_H , M_{D1} , M_V and M_{D2} , for $i \in \{1, 2, \dots, I-1\}$ and $j \in \{1, \dots, J-1\}$. Thus, we can obtain other 40 luminance features in the form of histogram. In total, there are 50 quality-aware features for luminance feature representation extracted for each distorted SCI with one scale.

Existing studies show that derivative information at different orders is highly correlated with different texture features of natural scenes [35]–[37]. The first-order derivative information is generally related to the slope and elasticity of a surface which characterize the geometric properties partially, while the second order derivatives can be used to capture the curvature related geometric properties [37]. As we know, there are many words in textual regions of SCIs, where the distortion can be well captured by texture information as well as other high-frequency pictorial regions. The second order derivatives can effectively capture the variation of the tiny structure which has influence on visual distortion of SCIs. Here, we first calculate gradient maps denoted by first-order derivative information based on four filters with different

directions, as defined in Fig. 3. The gradient maps are calculated based on the luminance map as follows:

$$G_1 = g_1 \otimes S', \quad (10)$$

$$G_2 = g_2 \otimes S', \quad (11)$$

$$G_3 = g_3 \otimes S', \quad (12)$$

$$G_4 = g_4 \otimes S', \quad (13)$$

where S' denotes luminance map of the corresponding SCI; g_1 , g_2 , g_3 and g_4 are the filters with four different directions as defined in Fig. 3; G_1 , G_2 , G_3 and G_4 are the corresponding gradient maps extracted by convolution operation $\otimes$. We provide some gradient map samples in Fig. 5. The GB distortion will affect the visual quality of SCI by which the non-connected regions in the SCI may connect and the high-frequency region may become smooth. The orientation information of texture regions in distorted SCI might be also influenced, thus we adopt four filters with different directions to capture the distortion.

In order to extract the second-order derivative information, we employ the rotation invariant uniform LBP [38] based on these four gradient maps. The texture features are extracted from the second-order derivative information in the form of histogram. The ordinary LBP formula can be expressed as follows:

$$LBP_{E,L} = \sum_{i=0}^{E-1} \theta(I_i - I_c) 2^i, \quad (14)$$

$$\theta(I_i - I_c) = \begin{cases} 1, & (I_i - I_c) \geq 0 \\ 0, & (I_i - I_c) < 0 \end{cases} \quad (15)$$

where E and L represent the number of neighbors and the radius of the neighborhood; $(I_0, I_1, \dots, I_{E-1})$ denote pixel values of E circularly symmetric neighboring pixels; I_c represents intensity of the center pixel in the local region. Based on the study [38], the local rotation invariant uniform LBP operator can be defined as follows:

$$RLBP'_{E,L} = \begin{cases} \sum_{i=0}^{E-1} \theta(I_i - I_c), & \psi(LBP_{E,L}) \leq 2 \\ E + 1, & \text{Otherwise} \end{cases} \quad (16)$$

$$\psi(LBP_{E,L}) = \|\theta(I_{E-1} - I_c) - \theta(I_0 - I_c)\| + \sum_{i=0}^{E-1} \|\theta(I_i - I_c) - \theta(I_{i-1} - I_c)\| \quad (17)$$

where ψ is used to calculate the number of bitwise transitions. The rotation invariant uniform LBP would have $E + 2$ distinct patterns and each pattern can be mapped to one bin of a histogram. In our implementation, E is set to 8, and thus there would be 10 bins for each LBP histogram. We calculate texture features by the second-order derivative information based on four gradient maps with the same operation. Totally, there are 40 quality-aware features for texture information extracted for each distorted SCI with one scale. The histograms of texture information are shown in Fig. 5, where we can observe that the histograms of texture information are different for different distortion types.

B. Regression Model for Quality Prediction

Given a distorted SCI, we can obtain 90 features with the above operations for each scale, including 50 features for luminance information and the others for texture information. For a better feature representation, we resize images by downsampling with factors of 2 and 4. Thus, we can obtain three images with three different scales (original image, Image2 by downsampling with factor 2, and Image3 by downsampling with factor 4), we find that there is no improvement in predicting the visual quality of SCIs when the number of scales is more than 3. There are 270 features for each distorted SCI in total. SVR with radial basis function (RBF) kernel is employed as the mapping function from the quality-aware features to subjective ratings [39], [50] by utilizing the LIBSVM package [51]. In the implementation, we divide the database

into training and testing subsets randomly for 1000 times, with 80% as the training dataset and the rest as the testing dataset associated with 4 reference SCIs, and the median performance is reported.

III. EXPERIMENTAL RESULTS

A. Screen Content Image Database

To test the effectiveness of the proposed NRLT, we conduct comparison experiments on SCI database SIQAD [15], [16]. There are 1000 SCIs contained in this database in total, including 20 reference SCIs with diverse visual content and 980 distorted versions degraded by seven distortion types (GN, GB, MB, CC, JPEG2000, Layer Segmentation Based Coding (LSC), and JPEG), with seven degradation levels for each distortion type. Although the visual content of images in this database is diverse, there is no SCI for graphics. The textual content mainly includes English content. Since the subjective test for this database was conducted by using computer, image resolutions are adjusted with a certain size for displaying on computer. DMOS values are provided as subjective scores in this database.

B. Evaluation Methodology

Here, three commonly used approaches are adopted to calculate the correlation between objective scores (output by an algorithm) and subjective scores (user ratings for SCIs): Pearson Linear Correlation Coefficient (PLCC), Spearman Rank-order Correlation Coefficient (SRCC), and Root Mean Squared Error (RMSE). PLCC can be used to estimate the prediction accuracy, while SRCC can be adopted to evaluate the prediction monotonicity. RMSE is a measure of deviation between the subjective and objective quality scores. In general, a better perceptual quality prediction metric has higher PLCC and SRCC values, and lower RMSE value. Given the n -th image in the database (with N images totally), its subjective and objective quality scores are s_n and o_n . PLCC can be calculated as follows.

$$PLCC = \frac{\sum_{n=1}^N (s_n - \hat{s})(o_n - \hat{o})}{\sqrt{\sum_{n=1}^N (s_i - \hat{s}) * \sum_{n=1}^N (o_i - \hat{o})}}, \quad (18)$$

where $\hat{s}$ and $\hat{o}$ denote the mean values of s_n and o_n , respectively.

We can estimate SRCC as follows.

$$SRCC = 1 - \frac{6 \sum_{n=1}^N d_n^2}{N(N^2 - 1)}, \quad (19)$$

where d_n is the difference between the n -th image's ranks in subjective and objective results.

RMSE can be calculated as follows.

$$RMSE = \sqrt{\frac{\sum_{n=1}^N (o_n - s_n)^2}{N}} \quad (20)$$

The estimated quality scores by different IQA metrics might have different ranges, we use a five-parameter mapping function to nonlinearly regress the quality scores into a common

TABLE I
EXPERIMENTAL RESULTS OF NRLT AND OTHER EXISTING FR METHODS ON SIQAD DATABASE

Components	PSNR	SSIM	GMSD	MAD	IFC	SPQA	SQI	GSS	EMSQA	NRLT
PLCC	0.5869	0.5912	0.7259	0.6191	0.6395	0.8584	0.8644	0.8461	0.8648	0.8442
SRCC	0.5608	0.5836	0.7305	0.6067	0.6011	0.8416	0.8548	0.8359	0.8504	0.8202
RMSE	11.5859	11.5450	9.4684	11.2409	10.9920	7.3421	7.1782	7.6310	7.1860	7.5957

TABLE II
EXPERIMENTAL RESULTS OF NRLT AND OTHER EXISTING RR AND NR METHODS ON SIQAD DATABASE

Components	NIQE	ILNIQE	BRISQUE	GMLOG	GWH-GLBP	RR-IQA	BLIQUP-SCI	BQMS	NRLT
PLCC	0.3749	0.3854	0.8113	0.7608	0.7903	0.8014	0.7705	0.8115	0.8442
SRCC	0.3568	0.3212	0.7749	0.7035	0.7233	0.7655	0.7990	0.8005	0.8202
RMSE	13.152	13.2085	8.2565	9.2530	8.7480	8.5620	10.0200	9.3042	7.5957

space as follows:

$$f(x) = \beta_1 \left(\frac{1}{2} - \frac{1}{1 + e^{(\beta_2(x - \beta_3))}} \right) + \beta_4 x + \beta_5 \quad (21)$$

where $(\beta_1, \dots, \beta_5)$ are parameters to be fitted.

C. Comparison Experiments

Here, we compare the proposed NRLT with the following state-of-the-art perceptual quality estimation methods: PSNR, SSIM [6], GMSD [12], MAD [41], IFC [40], SPQA [16], SQI [19], NIQE [17], ILNIQE [46], BRISQUE [39], GMLOG [50], GWH-GLBP [52], GSS [53], EMSQA [54], RR-IQA [22], BLIQUP-SCI [25] and BQMS [23]. Among these metrics, PSNR, IFC, MAD, GMSD, SPQA, SQI, GSS and EMSQA are FR VQA metrics. Specifically, SPQA, SQI, GSS and EMSQA are designed for VQA of SCIs. RR-IQA is a RR VQA method and it is designed for quality prediction of SCIs. NIQE, ILNIQE, BRISQUE, GMLOG, GWH-GLBP, BLIQUP-SCI and BQMS are NR VQA metrics, among which BLIQUP-SCI and BQMS are built specifically for blind quality prediction of SCIs. The performance of VQA metrics is evaluated by three different criteria: PLCC, SRCC and RMSE.

For FR methods, the results of SPQA, SQI, GSS and EMSQA are taken from the originally published papers. For other FR methods, we use the source code from their public websites to conduct the experiment. For RR and NR methods, the results of RR-IQA, BLIQUP-SCI and BQMS are taken from the originally published papers, and for NIQE and ILNIQE, the pre-trained models are used to calculate the visual quality scores of distorted SCIs directly. For other NR methods and the proposed NRLT, we randomly utilize 80% images in the database for training to learn a mapping function, which is subsequently used to calculate the perceptual quality scores of the rest distorted SCIs. There are 20 reference SCIs and each reference SCI is associated with 49 distorted versions. Thus, there are 784 distorted SCIs in training set associated with 16 reference SCIs, and the rest samples are used for testing. We repeat this operation with 1000 times, and the media performance is reported. The experimental results are shown in Tables I and II. From Table I, we can observe that GMSD can obtain higher accuracy in estimating the perceptual quality

TABLE III
EXPERIMENTAL RESULTS OF NRLT AND OTHER EXISTING METHODS BY DISTORTION-BASED SEPARATION

Components	BRISQUE	GMLOG	GWH-GLBP	NRLT
PLCC	0.6468	0.4023	0.6347	0.7469
SRCC	0.6010	0.3691	0.6136	0.7515
RMSE	8.7210	8.8043	6.5934	7.2507

of SCIs than other methods designed for quality prediction of natural images, including SSIM, IFC and MAD. Among all the compared metrics, NRLT can obtain competitive performance on perceptual quality evaluation of SCIs compared with FR methods designed for VQA of SCIs such as SPQA, SQI, GSS and EMSQA, and it can get better performance compared with these metrics designed for VQA of natural images.

From Table II, we can observe that NIQE and ILNIQE get worse performance on visual quality estimation of SCIs than BRISQUE, RR-IQA, BQMS and NRLT. GMLOG and GWH-GLBP designed based on gradient information can obtain high performance in quality evaluation of SCIs and even better than some FR metrics, which demonstrates that gradient information can capture the distortion of SCIs well. Among the compared methods, NRLT can obtain the best performance of visual quality estimation of SCIs, as demonstrated by the highest PLCC and SRCC values, and the lowest RMSE value of NRLT in Table I.

To demonstrate the robustness of the proposed NRLT, we conduct one more experiment by using distortion-based separation. There are 7 distortion types in this database. We use the images with 6 distortion types as the training set and the rest are used as testing set. The median performance is shown in Table III. From this Table, we can observe that the proposed NRLT can also obtain the consistent results with subjective ratings in estimating the visual quality of SCIs. And it can obtain better performance than other existing related metrics.

D. Performance Analysis

As shown in Table I, compared with the following VQA metrics designed for natural images such as SSIM and GMSD, the VQA metrics including SPQA and SQI designed

specifically for VQA of SCIs can achieve better performance in quality estimation for SCIs. The main reason is that these VQA methods designed for SCIs consider the difference between visual perception of pictorial and textual regions in SCIs. There are many words contained in SCIs which tend to attract human attention as well as other high-frequency pictorial regions. In SPQA, different visual quality calculation methods for pictorial and textual regions are designed under the consideration that the HVS perception to pictorial and textual regions are different, and an activity weighting strategy is proposed to predict the effect of pictorial and textual regions to the overall visual quality of SCIs. SQI is a FR VQA metric designed for VQA of SCIs by incorporating viewing field adaption based on the perceptual property that the extent of the visual field used to perceive useful information is much smaller in textual region than that in pictorial region. Also, information content weighting is used to combine the quality scores of pictorial and textual regions to obtain the final quality result of SCIs. Thus, SPQA and SQI can obtain better quality prediction performance for SCIs than these metrics designed for natural images.

In SPQA and SQI, SCI is divided into pictorial and textual regions for visual quality prediction. Different from these methods, SCI is regarded as a whole when its quality is predicted in GSS, EMSQA and NRLT. SPQA and SQI consider the differences between visual properties of pictorial and textual regions, and a simple linear combination of these quality scores of pictorial and textual regions is used to capture complicated interaction of visual perception for SCIs. As we can observe in Table I, these methods which do not segment SCI into pictorial and textual regions including GSS, EMSQA and NRLT can also obtain good objective quality prediction results with high correlation with human subjective rating.

As shown in Table II, NIQE and ILNIQE deliver lower performance than other methods for the reason that the used features meet a certain distribution of natural images and it is inconsistent with feature distributions of SCIs. Thus, multivariate Gaussian models learnt from natural images are not suitable for VQA of SCIs. Meanwhile, NIQE and ILNIQE measure the deviation of the statistics of distorted image patches from the statistics of reference image patches pre-learned from high-quality images. On the contrary, we extract quality-aware features based on the characteristics of the HVS. The normalization operation is used to mimic the nonlinear masking process in early human vision and eliminate the redundant information of visual scenes. The luminance information is extracted in the form of histogram based on the normalized intensity map. Furthermore, we calculate the gradient maps with different directions as the first order derivatives, and LBP is used to compute the second order derivatives. The statistical texture features are obtained in the form of histogram based on the second order derivatives. With the statistical luminance and texture features, the proposed method can obtain the best quality prediction performance for SCIs among the compared NR methods, as shown by experimental results in Table II.

Although RR-IQA [22] uses the reference information to construct the quality prediction model, it employs the average

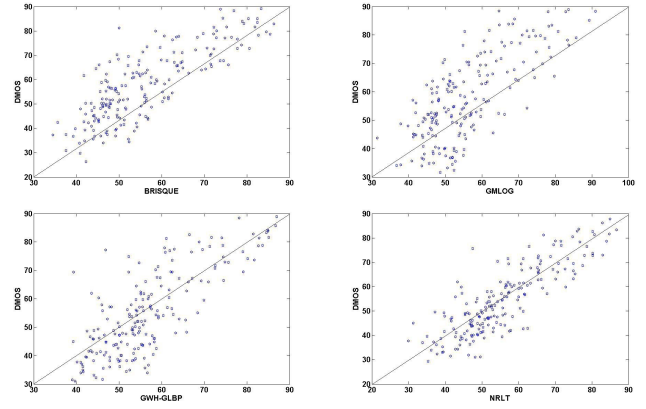


Fig. 6. The scatter plots of predicted quality scores by different methods against the DMOS values (192) on the SIQAD. The horizontal axis in each figure denotes predicted quality scores and the vertical axis in each figure represents the DMOS values. First row: BRISQUE and GMLOG; Second row: GWH-GLBP and NRLT.

pooling scheme to fuse the similarity of each corresponding bins of histograms, which leads to relatively poor performance of visual quality prediction of SCIs. BLIUP-SCI [25] belongs to opinion-unaware quality prediction method, since it does not need the subjective quality scores in the training phase. However, it uses four FR methods designed for natural images to generate the labels, and obtains the final quality score of SCI by a linear summation of local and global quality scores. It cannot obtain promising quality prediction performance of SCIs, since the used four FR methods cannot get good performance for quality prediction of SCIs [16]. BQMS uses 1000 training samples to fit the linear regression model before NSS feature extraction. Different from BQMS, we utilize histograms to represent the luminance and texture features in the proposed method. Experimental results shown in Table II demonstrate the effectiveness of the proposed NRLT.

Meanwhile, we also provide the scatter plots describing the distributions of objective quality scores against DMOS values from different NR methods including BRISQUE, GMLOG, GWH-GLBP and the proposed NRLT in Fig. 6. From this figure, we can observe that the points from NRLT are more centralized than these from other existing methods, which demonstrates that the predicted visual quality scores by NRLT are more consistent with subjective ratings. These results are consistent with the experimental results in Table II.

E. Feature Validation

The quality-aware features extracted in this study mainly include two parts: the first part is luminance features and the second part is texture features. As shown in Fig. 4, the luminance maps are altered due to the existing distortion, and different distortion types lead to the distinct changes of luminance distributions. For GN distortion, many white points appear in the associated luminance map randomly. For CC distorted SCI, the variation in luminance map is smaller than these with other distortion types, and its corresponding luminance distribution is similar to that of the original SCI. The reason is that the distortion of contrast change always occurs

TABLE IV
EXPERIMENTAL RESULTS OF DIFFERENT FEATURES

Components	Luminance	Texture	NRLT
PLCC	0.8333	0.7781	0.8442
SROCC	0.7936	0.7245	0.8202
RMSE	7.8341	8.9273	7.5957

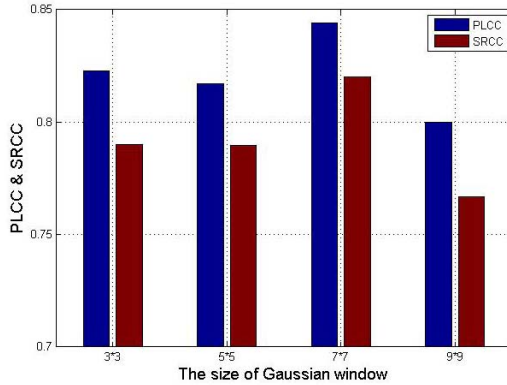


Fig. 7. The comparison results with different sizes of Gaussian window in terms of PLCC and SRCC.

in global scope and the local normalization is not sensitive to this type of distortion. For GB distorted SCI, the edge information is preserved and the detailed structure information is destroyed, which can be clearly reflected by the luminance map. And high peak appears in the corresponding distribution of GB distorted SCI, as shown in Fig. 4. On the contrast, part of detailed structure information can be preserved in MB distorted SCI. For GB and MB distorted SCIs, luminance features are more effective than texture features, since the distortions of luminance features are more distinguishable. Although the variations can be reflected by luminance distributions of JPEG, JPEG2000 and LSC distorted SCIs compared with the distribution of original SCI, the differences among distributions of JPEG, JPEG2000 and LSC distorted SCIs are slightly small, as shown in Fig. 4. In these three cases, texture features are more effective than luminance features in measuring visual distortion of SCIs, since the distributions of texture features are more distinguishable than those of luminance features, as shown in Fig. 5.

In this section, we conduct the comparison experiment to demonstrate the effect of the selected features on SIQAD. We divide the quality-aware features into two groups: luminance features and texture features. Only the luminance features or texture features are used to train the proposed model by SVR. The experimental results are shown in Table IV. In this table, the first column lists the evaluation index, while the second and third columns list the experimental results by only luminance features or texture features. From this table, we can observe that the proposed method with statistical luminance features can obtain better performance than that with statistical texture features. Overall, the proposed method with both statistical luminance and texture features can obtain better performance than that with only one type of features (statistical luminance features or statistical texture features).

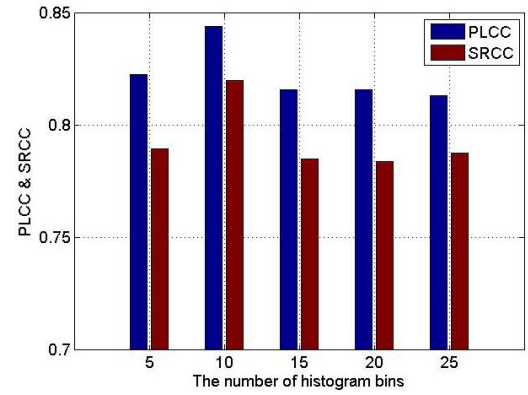


Fig. 8. The comparison results with different number of histogram bins in terms of PLCC and SRCC.

TABLE V
EXPERIMENTAL RESULTS OF DIFFERENT SCALES

Number	1	2	3	4	5
PLCC	0.8051	0.8227	0.8442	0.8368	0.8116
SROCC	0.7713	0.7937	0.8202	0.8078	0.7768
RMSE	8.3584	7.9978	7.5957	7.2228	8.2369

F. Parameter Setting

In this experiment, we explore the effect of the parameter setting of feature extraction. One parameter considered in this work is the size of Gaussian window used to calculate the luminance map, which is the basis of feature extraction. Inspired by [18] and [19], the extent of the visual field used to extract useful information in pictorial region is much larger than that in textual region. The sizes of the Gaussian window are set to 3*3, 5*5, 7*7 and 9*9 in the experiment for comparison. The experimental results are shown in Fig. 7. As we can observe, the performance by the window size 7*7 is superior to those from other cases. Thus, we set the size of Gaussian window to 7*7. During feature extraction, we employ histogram to represent statistical luminance features. In general, the feature histogram with a large number of bins is unstable, while the feature histogram with a small number of bins cannot represent the characteristics of feature distributions well. We have conducted one comparison experiment by using different number of bins to evaluate the performance of feature histograms, where the bin is set to 5, 10, 15 and 20. The experimental results are shown in Fig. 8. From this figure, we can see that the proposed method can obtain robust quality prediction performance with different histogram bins. We set the number of histogram bins to 10 for the optimal performance. For feature representation, we have extracted the quality-aware features with different number of scales. The experimental results are shown in Table V. We set the number of histogram bins to 10 for the optimal performance. For feature representation, we have extracted the quality-aware features with different number of scales. The experimental results are shown in Table V. From this table, we can see that there is no improvement in visual quality prediction of SCIs when the number of scales is

more than three. Thus, we set the number of scales to three.

IV. CONCLUSION

In this study, we have proposed a no reference quality assessment method for SCIs inspired by perceptual properties that the HVS is sensitive to luminance and texture information. The proposed NRLT metric includes two steps. In the first step, a luminance map is computed through local normalization and the histogram of the luminance map is used to represent the statistical luminance features of SCIs in the global scope. Furthermore, four filters with different orientations are used to calculate gradient maps of the luminance map, by which we extract LBP histogram features as statistical texture features for SCIs in the global scope. In the second step, SVR is adopted to train the quality prediction model from quality-aware features to visual quality of SCI. Experimental results demonstrate that the proposed NRLT can obtain superior performance against state-of-the-art approaches, even including some FR VQA metrics.

REFERENCES

- [1] J. Lei, S. Li, C. Zhu, M.-T. Sun, and C. Hou, "Depth coding based on depth-texture motion and structure similarities," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 25, no. 2, pp. 275–286, Feb. 2015.
- [2] S. Ma, X. Zhang, J. Zhang, C. Jia, S. Wang, and W. Gao, "Nonlocal in-loop filter: The way toward next-generation video coding?" *IEEE MultiMedia*, vol. 23, no. 2, pp. 16–26, Apr./Jun. 2016.
- [3] H. Hu, Y. Wen, T.-S. Chua, and X. Li, "Toward scalable systems for big data analytics: A technology tutorial," *IEEE Access*, vol. 2, pp. 652–687, 2014.
- [4] S. Minaee and Y. Wang, "Screen content image segmentation using robust regression and sparse decomposition," *IEEE J. Emerg. Sel. Topics Circuits Syst.*, vol. 6, no. 4, pp. 573–584, Dec. 2015.
- [5] J. Xu, R. Joshi, and R. A. Cohen, "Overview of the emerging HEVC screen content coding extension," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 26, no. 1, pp. 50–62, Jan. 2016.
- [6] 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.
- [7] Z. Wang, E. P. Simoncelli, and A. C. Bovik, "Multiscale structural similarity for image quality assessment," in *Proc. IEEE Asilomar Conf. Signals, Syst., Comput.*, Nov. 2003, pp. 1398–1402.
- [8] 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.
- [9] 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.
- [10] J. Wu, W. Lin, G. Shi, and A. Liu, "Perceptual quality metric with internal generative mechanism," *IEEE Trans. Image Process.*, vol. 22, no. 1, pp. 43–54, Jan. 2013.
- [11] 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.
- [12] 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.
- [13] 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.
- [14] Z. Wang and A. C. Bovik, "Mean squared error: Love it or leave it? A new look at signal fidelity measures," *IEEE Signal Process. Mag.*, vol. 26, no. 1, pp. 98–117, Jan. 2009.
- [15] H. Yang, Y. Fang, W. Lin, and Z. Wang, "Subjective quality assessment of screen content images," in *Proc. IEEE Int. Workshop QoMEX*, Sep. 2014, pp. 257–262.
- [16] 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, Nov. 2015.
- [17] A. Mittal, R. Soundararajan, and A. C. Bovik, "Making a 'completely blind' image quality analyzer," *IEEE Signal Process. Lett.*, vol. 20, no. 3, pp. 209–212, Mar. 2013.
- [18] M. S. Castelhan and K. Rayner, "Eye movements during reading, visual search, and scene perception: An overview," in *Cognitive and Cultural Influences on Eye Movements*, K. Rayner, D. Shem, X. Bai, and G. Yan, Eds. Psychology Press, 2008, pp. 3–33.
- [19] S. Wang, K. Gu, K. Zeng, Z. Wang, and W. Lin, "Objective quality assessment and perceptual compression of screen content images," *IEEE Comput. Graph. Appl.*, to be published, doi: [10.1109/MCG.2016.46](https://doi.org/10.1109/MCG.2016.46).
- [20] Y. Fang, J. Yan, J. Liu, S. Wang, Q. Li, and Z. Guo, "Objective quality assessment of screen content images by uncertainty weighting," *IEEE Trans. Image Process.*, vol. 26, no. 4, pp. 2016–2027, Apr. 2017.
- [21] 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.
- [22] S. Wang, K. Gu, X. Zhang, W. Lin, S. Ma, and W. Gao, "Reduced-reference quality assessment of screen content images," *IEEE Trans. Circuits Syst. Video Technol.*, to be published, doi: [10.1109/TCSVT.2016.2602764](https://doi.org/10.1109/TCSVT.2016.2602764).
- [23] K. Gu, G. Zhai, W. Lin, X. Yang, and W. Zhang, "Learning a blind quality evaluation engine of screen content images," *Neurocomputing*, vol. 196, pp. 140–149, Jul. 2016.
- [24] L. Ma, S. Li, and K. N. Ngan, "Reduced-reference video quality assessment of compressed video sequences," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 22, no. 10, pp. 1441–1456, Oct. 2012.
- [25] F. Shao, Y. Gao, F. Li, and G. Jiang, "Toward a blind quality predictor for screen content images," *IEEE Trans. Syst., Man, Cybern., Syst.*, to be published, doi: [10.1109/TSMC.2017.2676180](https://doi.org/10.1109/TSMC.2017.2676180).
- [26] A. J. Bell and T. J. Sejnowski, "The 'independent components' of natural scenes are edge filters," *Vis. Res.*, vol. 37, no. 23, pp. 3327–3338, 1997.
- [27] H. C. Nothdurft, "Sensitivity for structure gradient in texture discrimination tasks," *Vis. Res.*, vol. 25, no. 12, pp. 1957–1968, 1985.
- [28] A. B. Watson and J. A. Solomon, "Model of visual contrast gain control and pattern masking," *J. Opt. Soc. Amer. A, Opt. Image Sci.*, vol. 14, no. 9, pp. 2379–2391, Sep. 1997.
- [29] R. A. Frazor and W. S. Geisler, "Local luminance and contrast in natural images," *Vis. Res.*, vol. 46, no. 10, pp. 1585–1598, 2006.
- [30] V. Mante, R. A. Frazor, V. Bonin, W. S. Geisler, and M. Carandini, "Independence of luminance and contrast in natural scenes and in the early visual system," *Nature Neurosci.*, vol. 8, no. 12, pp. 1690–1697, 2005.
- [31] 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.
- [32] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2005, pp. 886–893.
- [33] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *Int. J. Comput. Vis.*, vol. 60, no. 2, pp. 91–110, 2004.
- [34] J. Wu, W. Lin, G. Shi, Y. Zhang, W. Dong, and Z. Chen, "Visual orientation selectivity based structure description," *IEEE Trans. Image Process.*, vol. 24, no. 11, pp. 4602–4613, Nov. 2015.
- [35] S. A. Wallis and M. A. Georgeson, "Mach bands and multiscale models of spatial vision: The role of first, second, and third derivative operators in encoding bars and edges," *J. Vis.*, vol. 12, no. 13, p. 18, 2012.
- [36] B. Zhang, Y. Gao, S. Zhao, and J. Liu, "Local derivative pattern versus local binary pattern: Face recognition with high-order local pattern descriptor," *IEEE Trans. Image Process.*, vol. 19, no. 2, pp. 533–544, Feb. 2010.
- [37] D. Huang, C. Zhu, Y. Wang, and L. Chen, "HSOG: A novel local image descriptor based on histograms of the second-order gradients," *IEEE Trans. Image Process.*, vol. 23, no. 11, pp. 4680–4695, Nov. 2014.
- [38] T. Ojala, M. Pietikainen, and T. Maenpää, "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 24, no. 7, pp. 971–987, Jul. 2002.
- [39] A. Mittal, A. K. Moorthy, and A. C. Bovik, "No-reference image quality assessment in the spatial domain," *IEEE Trans. Image Process.*, vol. 21, no. 12, pp. 4695–4708, Dec. 2012.
- [40] 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.
- [41] 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.

- [42] K. Gu, G. Zhai, W. Lin, X. Yang, and W. Zhang, "No-reference image sharpness assessment in autoregressive parameter space," *IEEE Trans. Image Process.*, vol. 24, no. 10, pp. 3218–3231, Oct. 2015.
- [43] Y. Fang, K. Ma, Z. Wang, W. Lin, Z. Fang, and G. Zhai, "No-reference quality assessment of contrast-distorted images based on natural scene statistics," *IEEE Signal Process. Lett.*, vol. 22, no. 7, pp. 838–842, Jul. 2015.
- [44] M. Zhang, C. Muramatsu, X. Zhou, T. Hara, and H. Fujita, "Blind image quality assessment using the joint statistics of generalized local binary pattern," *IEEE Signal Process. Lett.*, vol. 22, no. 2, pp. 207–210, Feb. 2015.
- [45] P. G. Freitas, W. Y. L. Akamine, and M. C. Q. Farias, "No-reference image quality assessment based on statistics of local ternary pattern," in *Proc. 8th IEEE Int. Conf. Quality Multimedia Exper.*, Jun. 2016, pp. 1–6.
- [46] L. Zhang, L. Zhang, and A. C. Bovik, "A feature-enriched completely blind image quality evaluator," *IEEE Trans. Image Process.*, vol. 24, no. 8, pp. 2579–2591, Aug. 2015.
- [47] L. Li, W. Lin, X. Wang, G. Yang, K. Bahrami, and A. C. Kot, "No-reference image blur assessment based on discrete orthogonal moments," *IEEE Trans. Cybern.*, vol. 46, no. 1, pp. 39–50, Jan. 2016.
- [48] A. Rehman and Z. Wang, "Reduced-reference image quality assessment by structural similarity estimation," *IEEE Trans. Image Process.*, vol. 21, no. 8, pp. 3378–3389, Aug. 2012.
- [49] J. Wu, W. Lin, G. Shi, L. Li, and Y. Fang, "Orientation selectivity based visual pattern for reduced-reference image quality assessment," *Inf. Sci.*, vol. 351, pp. 18–29, Jul. 2016.
- [50] W. Xue, X. Mou, L. Zhang, A. C. Bovik, and X. Feng, "Blind image quality assessment using joint statistics of gradient magnitude and Laplacian features," *IEEE Trans. Image Process.*, vol. 23, no. 11, pp. 4850–4862, Nov. 2014.
- [51] C.-C. Chang and C.-J. Lin, "LIBSVM: A library for support vector machines," *ACM Trans. Intell. Symp. Technol.*, vol. 2, no. 3, p. 27, 2011.
- [52] Q. Li, W. Lin, and Y. Fang, "No-reference quality assessment for multiply-distorted images in gradient domain," *IEEE Signal Process. Lett.*, vol. 23, no. 4, pp. 541–545, Apr. 2016.
- [53] 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.
- [54] 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.*, Sep. 2016, pp. 81–85.
- [55] Z. Ni, L. Ma, H. Zeng, J. Chen, C. Cai, and K.-K. Ma, "ESIM: Edge similarity for screen content image quality assessment," *IEEE Trans. Image Process.*, vol. 26, no. 10, pp. 4818–4831, Oct. 2017.
- [56] W. Xue, L. Zhang, and K.-K. Ma, "Learning without human scores for blind image quality assessment," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2013, pp. 995–1002.
- [57] Q. Li, W. Lin, and Y. Fang, "BSD: Blind image quality assessment based on structural degradation," *Neurocomputing*, vol. 236, pp. 93–103, May 2016.
- [58] Y. Zhang, J. Wu, X. Xie, L. Li, and G. Shi, "Blind image quality assessment with improved natural scene statistics," *Digit. Signal Process.*, vol. 57, pp. 56–65, Oct. 2016.
- [59] S. Wang *et al.*, "Subjective and objective quality assessment of compressed screen content images," *IEEE J. Emerg. Sel. Topics Circuits Syst.*, vol. 6, no. 4, pp. 532–543, Dec. 2016.



Yuming Fang (M'13–SM'17) received the B.E. degree from Sichuan University, Chengdu, China, the M.S. degree from the Beijing University of Technology, Beijing, China, and the Ph.D. degree from Nanyang Technological University, Singapore. He is currently a Professor with the School of Information Technology, Jiangxi University of Finance and Economics, Nanchang, China. His research interests include visual attention modeling, visual quality assessment, image retargeting, computer vision, 3D image/video processing, and so on. He serves as

an Associate Editor for the IEEE ACCESS. He is on the Editorial Board of *Signal Processing: Image Communication*.



Jiebin Yan received the B.E. degree from Beijing City University, Beijing, China, in 2015. He is currently pursuing the M.S. degree with the School of Information Technology, Jiangxi University of Finance and Economics, Nanchang, China. His research interests include image quality assessment and machine learning.



Leida Li (M'14) received the B.S. and Ph.D. degrees from Xidian University, Xi'an, China, in 2004 and 2009, respectively. Since 2008, he has been a visiting Ph.D. student with the Department of Electronic Engineering, National Kaohsiung University of Applied Sciences, Taiwan. From 2014 to 2015, he was a Visiting Research Fellow with the Rapid-Rich Object Search Laboratory, School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore, where he was a Senior Research Fellow from 2016 to 2017. He is currently

a Full Professor with the School of Information and Control Engineering, China University of Mining and Technology, China. His research interests include multimedia quality assessment, information hiding, and image forensics.



Jinjian Wu (M'14) received the B.Sc. and Ph.D. degrees from Xidian University, Xi'an, China, in 2008 and 2013, respectively. From 2011 to 2013, he was a Research Assistant at Nanyang Technological University, Singapore. From 2013 to 2014, he was a Post-Doctoral Research Fellow at Nanyang Technological University. From 2013 to 2015, he was a Lecturer at Xidian University. Since 2015, he has been an Associated Professor with the School of Electronic Engineering, Xidian University. His research interests include visual perceptual modeling, saliency estimation, quality evaluation, and just noticeable difference estimation. He served as a TPC member for ICME2014–2015, PCM2015–2016, ICIP2015, and QoMEX2016. He received the Best Student Paper of ISCAS 2013.



Weisi Lin (M'92–SM'98–F'16) received the Ph.D. degree from King's College, London University, U.K. He is currently a Professor with the School of Computer Science and Engineering, Nanyang Technological University. He has published 180+ journal papers, 230+ conference papers, filed seven patents, and authored two books. His areas of expertise include image processing, perceptual signal modeling, video compression, and multimedia communication. He is a fellow of IET and an Honorary Fellow of the Singapore Institute of Engineering

Technologists. He has been an Associate Editor of the IEEE TRANSACTIONS ON IMAGE PROCESSING, the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY, the IEEE TRANSACTIONS ON MULTIMEDIA, and the IEEE SIGNAL PROCESSING LETTERS. He has been a Technical Program Chair for the IEEE ICME 2013, PCM 2012, QoMEX 2014, and the IEEE VCIP 2017. He has been an invited/panelist/keynote/tutorial speaker in 20+ international conferences, as well as a Distinguished Lecturer of the IEEE Circuits and Systems Society 2016–2017, and Asia-Pacific Signal and Information Processing Association, 2012–2013.