Subjective and Objective Quality Assessment of Compressed Screen Content Images

Shiqi Wang, *Member, IEEE*, Ke Gu, Xiang Zhang, *Member, IEEE*, Weisi Lin, *Fellow IEEE*, Li Zhang, *Member, IEEE*, Siwei Ma, *Member, IEEE*, and Wen Gao, *Fellow, IEEE*

***Abstract*— Objectively accessing the quality of screen content images (SCIs) is a challenging problem, as SCIs may not always have identical properties as natural scenes. Here we conduct comprehensive studies on the subjective and objective quality assessment of the compressed SCIs. Firstly, we build a database that contains the distorted SCIs generated by the high efficiency video coding standard as well as its extension on screen content compression. Subsequently, subjective experiments are conducted to evaluate the perceived quality of these SCIs with compression artifacts. To automatically predict the subjective quality, a reduced-reference quality assessment model is further learnt by a set of wavelet domain features concerning the generalized spectral behavior, the fluctuations of the energy, and the information content with relatively large scale training samples. Our experimental results show that the learnt model is able to achieve better prediction on the SCI quality with a few extracted meaningful features.**

***Index Terms*— Compressed screen content images, objective quality assessment, reduced-reference, subjective quality assessment.**

I. Introduction

RECENTLY, there has been a growing interest in virtual desktops and wireless displays, the purpose of which is to access the computational resources and remote data via the network. In these applications, the constantly updated screen at the sender side is usually recorded, compressed and transmitted to the display side [1]–[3]. Therefore, the quality of screen content image (SCI) plays an important role in the quality-of-experience of the interactive screen remoting systems. Quality assessment of SCIs has emerged as an important problem because of the exponential increase in the demand for these graphically rich services. However, it has been widely acknowledged that the properties of SCIs are usually distinguished from the natural images [4], [5], making the image quality assessment (IQA) task of SCIs a non-trivial problem, as traditional IQA algorithms are usually devised relying on the statistics of natural images. In [6], a database of distorted SCIs with subjective quality rankings was created. The correlations between subjective and objective scores also suggest that directly employing the IQA models developed for natural images may not suffice for the purpose of trusted SCI IQA. In view of the distinctive features of SCIs and their application scenarios, it is highly desirable to study the subjective and objective quality of the SCIs with compression artifacts, which can further provide essential guidance in devising, monitoring and optimizing the advanced screen communication systems [7]–[9].

The distinguished properties of SCIs have also motivated many advanced coding techniques, many of which have been adopted into the screen content coding (SCC) extension to the High Efficiency Video Coding (HEVC) standard [10], [11][1]. For instance, it is observed that the prediction residuals of the textual content in SCIs are usually sparse and may contain sharp edges in various directions, such that the traditional transform based coding techniques may not be that efficient. In [12]–[14], transform skipping schemes were proposed and adopted into the HEVC standard. Furthermore, considering the fact that the screen content typically contains a finite number of distinct colors within a local region in contrast to the continuous color representation of natural content, the palette coding mode based on "base color and index map" representation was studied in [4], [15]. Motivated by the observation that the repeated patterns may often occur within one frame [16], the intra motion compensation approach was applied to improve the screen content coding efficiency.

These advanced coding tools have shown promising coding performance in both HEVC and HEVC SCC extension standardization process, revealing the great potentials in deploying these standards in the interactive screen remoting system. Accordingly, there is an urgent need to further study the perceptual distortion originated from the compression artifacts induced by HEVC and HEVC SCC extension. This motivates us to firstly build a quality assessment database of compressed SCIs. In particular, we establish a new SCI quality assessment database, including 24 source and 492 compressed SCIs. To obtain the perceived quality scores of these compressed

Manuscript received February 13, 2016; revised May 24, 2016; accepted July 10, 2016. Date of publication August 31, 2016; date of current version December 9, 2016. This work was supported by the National Natural Science Foundation of China under Grant 61322106 and Grant 61571017, the National Basic Research Program of China (973 Program) under Grant 2015CB351800, and the Singapore MoE Tier 1 Project under Grant M4011379 and Grant RG141/14. This paper was recommended by Guest Editor W.-H. Peng.

S. Wang is with Rapid-Rich Object Search Laboratory, Nanyang Technological University, Singapore 639798 (e-mail: wangshiqi@ntu.edu.sg).

K. Gu and W. Lin are with the School of Computer Engineering, Nanyang Technological University, Singapore 639798 (e-mail: guke@ntu.edu.sg; wslin@ntu.edu.sg).

X. Zhang, L. Zhang, S. Ma, and W. Gao are with the School of Electronic Engineering and Computer Science, Institute of Digital Media, Peking University, Beijing 100871, China (e-mail: x_zhang@pku.edu.cn; li.zhang@pku.edu.cn; swma@pku.edu.cn; wgao@pku.edu.cn).

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/JETCAS.2016.2598756

[1] In this paper, in order to distinguish with HEVC SCC extension, "HEVC" only specifies the first edition of HEVC, also known as HEVC version 1.

SCIs, a comprehensive subjective test is conducted. Such database allows us to directly investigate the perceptual quality of the compressed SCIs generated by the state-of-the-art codecs specified by HEVC and its SCC extension [17], and guides the generation of more test materials and the design of more advanced quality measures in the future.

In addition to the subjective testing, it is also highly desirable to study the objective IQA measures of SCIs, targeting at automatically predicting the perceived SCI quality by human subjects. Depending on the availability of the reference image that is directly rendered by the computer, SCI IQA measures can also be classified into the full-reference (FR), reduced-reference (RR) and no-reference (NR) methods. In particular, FR measures require full access to the reference SCI, and most of the popular FR methods such as the structural similarity (SSIM) index [18] and its variants [19]–[21], visual signal-to-noise ratio (VSNR) [22], gradient similarity (GSIM) [23] and visual saliency-induced index (VSI) [24] are designed based on the validations using natural images. Recently, in [25] a novel FR-IQA algorithm termed as saliency-guided quality measure of SCIs (SQMS) has been proposed, which effectively captures the structural distortion of the screen content. Regarding the NR methods, they do not allow any access to the reference SCI. Generally, NR-IQA methods are important for natural images [26]–[33]. One reason lies in that the acquisition process of natural images using physical camera sensors may inevitably involve multiple types of distortions, making the perfect quality pristine image unavailable. By contrast, as SCIs are directly generated from the computer, at the sender side the SCIs are presumed to have perfect quality. Though the reference images are not available at the client side, such neat properties of SCIs make the RR-IQA that depends on much less information from the reference image more meaningful for SCI communication applications. Clearly, the RR-IQA is providing a compromise in-between the FR and NR IQA methods, where only sparsely sampled features that can adequately capture the properties of the reference image are efficient in predicting the quality.

The key to the successful RR-IQA methods lies in how to select and combine a set of meaningful RR features that can reflect the visual quality. In view of its importance in IQA research, numerous methods have been proposed in the literature. In [34], the wavelet domain natural image statistic model (WNISM) was proposed. Subsequently, the RR-IQA methods were developed in the divisive normalization domain, leading to the DNT-RR method [35], [36]. In [37], the entropy of primitives was used to reflect the visual quality. The authors in [38], [39] observed that the reorganization of the DCT coefficients may lead to a better RR-IQA method based on their city-block distance (ROCB). In [40], the image quality was reflected in the Fourier transform domain (FTD), in which the phase and magnitude are employed to access the image quality. In [41], a novel structural degradation model (SDM) was developed using the distance between the structural degradation information of the reference and distorted images. In [42], the histograms of the edge pattern map (EPM) were utilized to infer the quality of the distorted image. In [43], the reduced-reference image quality metric for contrast change (RIQMC) was proposed by measuring the differences between the statistical information (first $\sim$ fourth-order statistics, entropy, etc.) of the reference and distorted image histograms. Again, most of these methods are designed relying on the natural scene statistics obtained from natural images, which may not always have the same characteristics of screen content [44].

Fig. 1. All 20 source SCIs in the SIQAD database.

In this paper, we learn a $RR\ Wavelet-domain\ Quality\ Measure$ of $SCIs$ ($RWQMS$) by exploring the statistics of SCIs based on relatively large training samples. In analogy to the natural scene statistics (NSS) models that lay a perceptually meaningful groundwork to capture the low-level statistical properties of natural images, we use the statistics of SCIs in wavelet domain and map the feature space to the quality scores with the help of support vector regressor (SVR) [45]. Specifically, we employ the magnitude, variance and entropy of the wavelet coefficients [46] to characterize the spectral, energy and information content of the SCIs. At the receiver side, identical features extraction process as at the sender side is performed. In particular, to compare the extracted features, the mapping strategy from the destruction of the features to the quality score is learnt from the training phase using 149 original SCIs [47] and the corresponding corrupted versions. To avoid over-fitting, the content of training and testing SCIs are totally different. Such a methodology ensures that we can effectively learn the statistics of the computer generated images, which govern the perceptual characteristics of the distorted SCIs.

II. Subjective Quality Assessment of Compressed SCIs

A. Previous Subjective SCI IQA Studies

To our best knowledge, in addition to the proposed database, there is only one SCI IQA database available for study, which is called screen image quality assessment database (SIQAD) [6]. It includes 980 distorted SCIs which are generated by corrupting 20 source SCIs with seven distortion types at seven distortion levels. These distortion types include Gaussian blur (GB), Gaussian noise (GN),

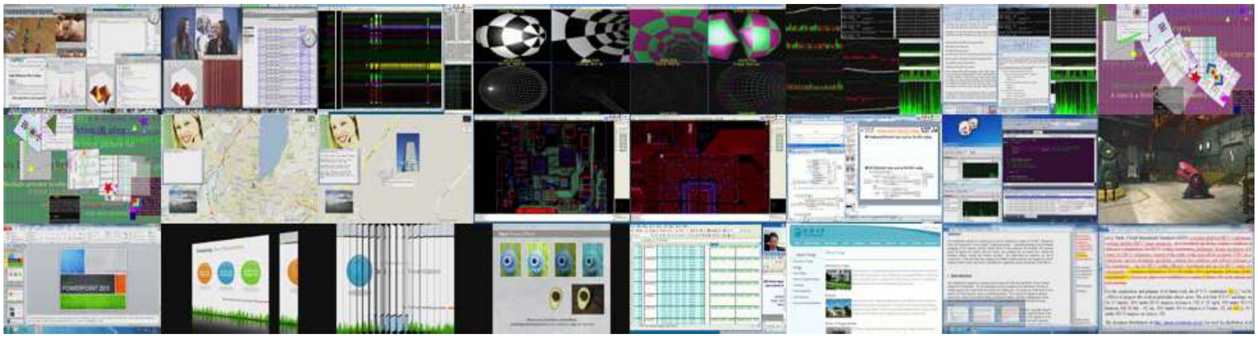


Fig. 2. All 24 source SCIs in the QACS database.

contrast change (CC), motion blur (MB), JPEG compression, JPEG2000 compression (JP2K) and Layer Segmentation based Coding (LSC).

In SIQAD, the 20 reference SCIs are obtained from webpages and digital magazines by screen recording, as illustrated in Fig. 1. The resolutions are from 626×612 to 832×728. Guided by the suggestions from ITU-R BT.500-13, the testing strategy is single stimulus (SS) method with 11 point discrete scales ranging from the worst 0 to the best 10. The subjects are asked to provide his/her visual opinion score for each distorted SCI. The specific viewing distance is 2–2.5 times of the screen height. Before the subjective testing process, the subjects are firstly trained with various distortion types and levels. This database lays a groundwork for SCI quality assessment, and subsequently advanced SCI FR-IQA algorithms have been developed to predict the SCI quality [9], [25].

B. Subjective Study of the Compressed SCIs

Here we propose a new database called quality assessment of compressed SCI (QACS), targeting at compensating for the features missing from SIQAD. As such, the distortion patterns generated by the state-of-the-art codecs including both HEVC and HEVC-SCC are involved. Following the common test conditions specified by HEVC SCC extension in [48], 24 pristine SCIs are collected, covering wide application scenarios including web pages, office (words and excels) and cloud CAD, as shown in Fig. 2. To create the distorted versions of these SCIs, each SCI is compressed with 11 QP values ranging from 30 to 50 by Intra coding of HEVC and HEVC SCC extension, respectively. In total, 528 distorted SCIs are generated. Compared with SQIAD, there are three unique features of the proposed database: the inclusion of HEVC and HEVC SCC compression artifacts, the inclusion of higher resolution SCIs, and the inclusion of more SCI content types and application scenarios [17].

The methodology of the subjective testing is consistent with that in developing the SIQAD database. Specifically, the single-stimulus is used in the subjective testing. Twenty subjects were invited to take participate in the test, in which they were asked to view the distorted SCIs with a specified viewing distance (around 2–2.5 screen heights). Again, 10-category discrete scale is employed to record the subjective opinions offered by the subjects, which can make the subjective values more distinguishable. After collecting the raw scores, we further process these scores in order to generate the final mean opinion score (MOS) for each SCI.

It is widely acknowledged that each subject has his/her own understanding on the quality of the perceived SCIs. However, considering all the 20 subjects, most of them should have similar agreements on the visual quality. This implies that we allow the dissimilarities of the subjective score to some extent. Yet, if the dissimilarities exceed the tolerance, the SCI should be rejected to ensure that only trusted MOS values are adopted in the database. Specifically, following the methodology in [49], the 25% and 75% of subjective ratings of the sorted subjective scores can be obtained for each SCI. In other words, the central 50% of subject ratings lie within the range from 25% to 75%, and if these ratings are constrained within an interval of 2, we can draw the conclusion that the subjects have arrived at an agreement of the perceived SCI quality. In this manner, in total 36 out of 528 outlier SCIs are rejected, leading to the final 492 SCIs used in the database. Furthermore, we perform the subject rejection process following the ITU-R BT 500-11 [50]. Specifically, the subjective ratings are firstly converted into Z-scores and for each subject the reliability is examined to determine whether his/her scores should be discarded. After performing this procedure, no subject gets rejected eventually.

The final MOS score is then computed by averaging the subjective ratings for each distorted SCI, which is given by

$$Q_j = \frac{1}{S} \sum_{i=1}^{S} Q_{i,j} \tag{1}$$

where S denotes the number of subjects and $Q_{i,j}$ denotes the rating by the subject i when viewing the jth distorted SCI. The final MOS scores of these distorted SCIs provide us further opportunities to analyze and evaluate the IQA performance.

In Fig. 3, we demonstrate the histogram of the MOS values, from which we can observe that the perceptual quality of the SCIs distributes over a wide MOS value range. The separation of the visual quality can also facilitate the future investigation on the perceptual characteristics of SCI and the design of trusted SCI IQA algorithms. Moreover, for each distorted SCI, we compute the 95% confidence interval to evaluate the consistency among subjects. The distribution of the confidence intervals is shown in Fig. 4. It is observed that the confidence intervals primarily locate in the range

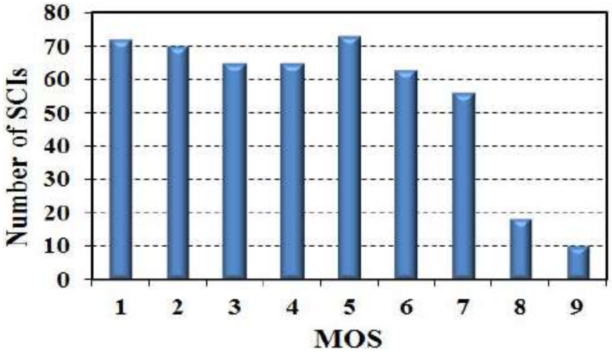


Fig. 3. Histogram of the MOS values in the QACS database.

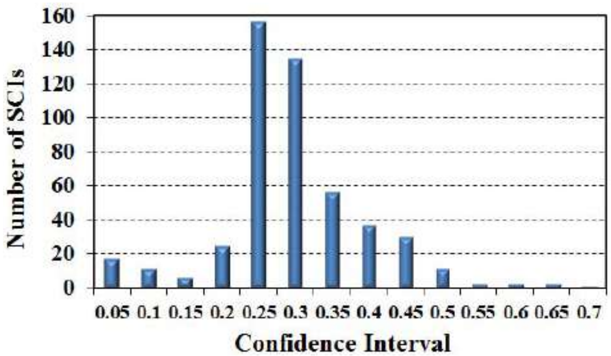


Fig. 4. Histogram of confidence intervals in the QACS database.

from 0.2 to 0.3. This further provides useful evidence of the high agreement across subjects in judging the quality of the SCIs. To facilitate future study and comparisons, we have put the database online [51].

III. Reduced-Reference SCI Quality Assessment

In this section, we will describe the learning based RR-IQA method that is specifically developed for the SCI quality prediction. In the interactive screen remoting system, usually the server and the client communicate with each other over a network [1]. In this process, to feasibly monitor the screen quality, the features from the sender and receiver SCIs are extracted and compared to monitor the perceptual degradations of the screen quality. In essence, the selected features should be sensitive to specific distortions and effectively summarize the perceptual characteristics of SCIs [44], such that the perceptual quality can be reliably learnt from the feature divergence.

Regarding the visual perception of SCIs and natural images, there are both similarities and differences. The similar aspect lies in that the ultimate receiver of both SCIs and natural images is the human visual system (HVS). Therefore, the characteristics of HVS should be naturally considered in the selection of perceptually meaningful features, such that the feature selection process turns out to be "perceptual quality aware". However, it is generally believed that the adaptation of HVS follows the statistics of natural scene, which is based on the hypothesis that the natural images only occupy a small subspace of all possible images [52]. By contrast, SCIs may not always belong to this subspace and the statistics of SCIs may not usually obey the NSS [53], rendering them unnatural appearance. In view of these considerations, the distinct features of SCIs should also be taken into consideration in the quality prediction process.

In analogy to the traditional learning based IQA methods [26], [32], [46], [54], the proposed RR-IQA scheme is also composed of the two modules: quality-aware feature extraction and quality prediction. As aforementioned, since the ultimate receiver of SCIs is HVS, the features that provide a "reduced description" should also coincide with visual perception. One distinct property of visual perception is that the spatial receptive field of simple cells can be characterized as being localized, oriented and bandpass [55], [56], comparable to the basis of wavelet. This implies that the wavelet transform exhibits close relationships with the visual cortex property. It is also widely acknowledged that the images are naturally multiscale [57], [58] and low level vision involves the decomposition of the visual signal into multiscale representation [31]. Therefore, from the perspective of HVS perception, the wavelet domain features are appropriate in the quality evaluation. Moreover, the textual regions in typical SCIs contain abundant high contrast edges [9], which deliver important information of the SCI content. As such, from the perspective of SCI perception, the wavelet transform, which is able to provide the multiscale representation of the edge information [59], is important for SCI perception due to the discriminative ability of different types of edges.

Inspired by these observations, the wavelet domain features such as coefficient magnitude, variance and entropy [46] are adopted in the RR-IQA method to characterize the SCI information. Given the extracted features from the reference and distorted SCIs, the next procedure is to map the feature divergence to the quality score. In the literature, the learning based regression models are widely employed in this procedure. In this work, we adopt the SVR for the purpose of regression, which has been shown to achieve good performance on the high-dimensional regression problems. The aim of the training process is learning to map the departures from reference SCI features to the quality score based on the statistics captured from SCIs.

The framework of the proposed learning based SCI RR-IQA method is illustrated in Fig. 5. Specifically, the RR features at the transmitter side are obtained by transforming the rendered SCIs into the wavelet domain to capture the statistical properties with multiscale representation. The features are subsequently transmitted to the receiver through an ancillary channel. After receiving the distorted SCI, the feature extraction unit at the receiver side calculates the features in a similar fashion. In the final stage, the divergence between the features that are extracted from the transmitter and receiver SCIs is mapped into one quality score with the learning based quality prediction mechanism.

A. Wavelet Domain Feature Extraction

The wavelet domain feature extraction is achieved by a scale-space orientation decomposition, which is implemented

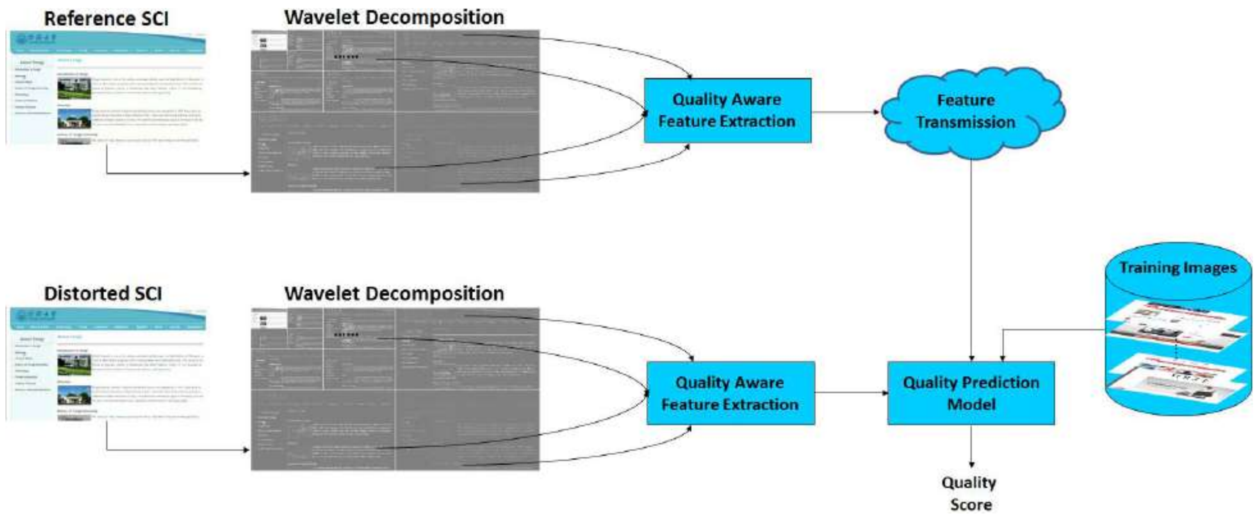


Fig. 5. Framework of the proposed learning based SCI RR-IQA method.

with the multiscale steerable pyramid structure along horizontal, vertical and diagonal directions (HL, LH, and HH). The high frequency subbands carry important edge information of SCIs to convey the semantic information with multiscale representation. In this work, the input SCI is decomposed into four scales and in each scale the HL and LH subbands are merged together, as they usually share similar statistics. In this manner, we can get eight subbands in total. Within each subband, the features representing the coefficient magnitude, variance and information [46] are extracted based on the wavelet coefficients.

Firstly, we compute the mean of the log-domain wavelet coefficients to measure the amplitude spectrum of the SCI in each subband. Specifically, assume the coefficient at position (i, j) of the kth subband is $\chi_k(i, j)$, then the feature that represents the coefficient magnitude is given by

$$\mu_k = \frac{1}{N_H * N_W} \sum_{i=1}^{N_H} \sum_{j=1}^{N_W} \log |\chi_k(i, j)| \tag{2}$$

where N_H and N_W are the height and width of the kth subband.

Subsequently, the variance of the coefficient is computed to reflect the fluctuation and distribution of the energy

$$\sigma_k = \frac{1}{N_H * N_W} \sum_{i=1}^{N_H} \sum_{j=1}^{N_W} \log \left| \chi_k(i, j) - \frac{\sum_{i=1}^{N_H} \sum_{j=1}^{N_W} \chi_k(i, j)}{N_H * N_W} \right|. \tag{3}$$

It is also worth noting that in accordance with the coefficient magnitude, the log domain variance is computed.

Finally, the entropy is calculated to reflect the information content in wavelet domain

$$I_k = \sum_{i=1}^{N_H} \sum_{j=1}^{N_W} p(\chi_k(i, j)) \log p(\chi_k(i, j)) \tag{4}$$

where $p(\chi_k(i, j))$ indicates the probability distribution of the coefficient in the subband.

After computing the features of the reference SCI in each subbands, they are stacked into one vector to efficiently summarize the perceptually relevant content of the screen visual signal [46]

$$\begin{aligned} \mathbf{f} &= \{\mathbf{f}_{\boldsymbol{\mu}}; \mathbf{f}_{\boldsymbol{\sigma}}; \mathbf{f}_{\boldsymbol{I}}\} \\ &= \{\mu_1, \mu_2, \ldots \mu_8; \sigma_1, \sigma_2, \ldots, \sigma_8; I_1, I_2, \ldots, I_8\}. \end{aligned} \tag{5}$$

Analogous to the reference SCI, the feature vector of the distorted SCI $\widetilde{\mathbf{f}} = \{\widetilde{\mathbf{f}}_{\boldsymbol{\mu}}; \widetilde{\mathbf{f}}_{\boldsymbol{\sigma}}; \widetilde{\mathbf{f}}_{\boldsymbol{I}}\}$ can be obtained in a similar fashion. Therefore, it is natural to compare the features of the original and distorted SCIs and relate them with the perceptual quality. Fig. 6 shows the impact of compression distortions on the feature divergence with different quantization parameter (QP) settings, where

$$\begin{aligned} \Delta\mathbf{f}_{\boldsymbol{\mu}} &= \mathbf{f}_{\boldsymbol{\mu}} - \widetilde{\mathbf{f}}_{\boldsymbol{\mu}} = \{\Delta\mu_1, \Delta\mu_2, \ldots, \Delta\mu_8\} \\ \Delta\mathbf{f}_{\boldsymbol{\sigma}} &= \mathbf{f}_{\boldsymbol{\sigma}} - \widetilde{\mathbf{f}}_{\boldsymbol{\sigma}} = \{\Delta\sigma_1, \Delta\sigma_2, \ldots, \Delta\sigma_8\} \\ \Delta\mathbf{f}_{\mathbf{I}} &= \mathbf{f}_{\mathbf{I}} - \widetilde{\mathbf{f}}_{\mathbf{I}} = \{\Delta I_1, \Delta I_2, \ldots, \Delta I_8\}. \end{aligned} \tag{6}$$

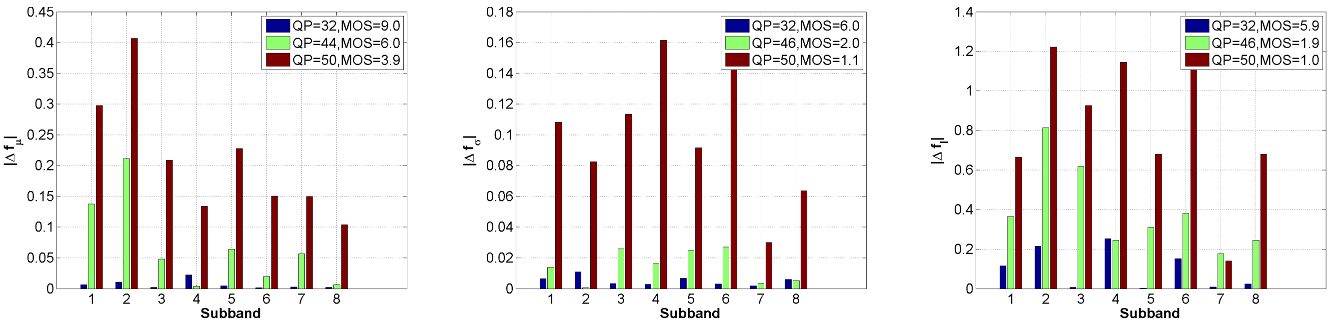
It is observed that the absolute values of $\Delta\mathbf{f}_{\boldsymbol{\mu}}$, $\Delta\mathbf{f}_{\boldsymbol{\sigma}}$ and $\Delta\mathbf{f}_{\mathbf{I}}$ exhibit strong trend associated with different QP settings and MOS values. This further demonstrates that the extracted features are sensitive to distortions and able to provide a good indication of the perceptual quality.

Finally the feature divergence is calculated to infer the quality of the distorted SCI

$$\begin{aligned} \Delta\mathbf{f} &= \{\Delta\mathbf{f}_{\boldsymbol{\mu}}; \Delta\mathbf{f}_{\boldsymbol{\sigma}}; \Delta\mathbf{f}_{\boldsymbol{I}}\} \\ &= \{\Delta\mu_1, \ldots \Delta\mu_8; \Delta\sigma_1, \ldots, \Delta\sigma_8; \Delta I_1, \ldots, \Delta I_8\}. \end{aligned} \tag{7}$$

B. Quality Prediction Model

The quality prediction model serves to predict the perceptual quality of the distorted SCI given the feature divergence vector $\Delta\mathbf{f}$. To avoid over-fitting, it is required to be learnt with the distorted SCIs of different content together with the associated MOS values. However, currently as the SCI IQA database is still largely lacking, it is difficult to collect enough



(a) (b) (c)

Fig. 6. The absolute feature divergence under different QP settings. (a) "SC_PPT_DOC_XLS"; (b) "SC_Desktop"; (c) "SC_Map".

Fig. 7. Illustration of the 149 training SCIs from the saliency database in [47].

learning samples. This prompts us to adopt an alternative strategy that generates the groundtruth data with an FR IQA model. Specifically, the 149 SCIs in the saliency database [47] and the associated around 12 000 distorted SCIs are used for training. As illustrated in Fig. 7, the contents of these SCIs are different with those in QACS and SIQAD. The corresponding distorted SCIs are generated by corrupting these images with eight distortion types (Gaussian blur, white noise, motion blur, contrast change, JPEG, JPEG2000, HEVC and HEVC-SCC extension) and 10 distortion levels. Subsequently, the specifically developed SCI FR-IQA algorithm SQMS [25] that has been proven to achieve state-of-the-art performance is used to generate the training scores. The correlations between SQMS and the MOS scores in both SIQAD and QACS show that this method is able to provide higher prediction accuracy across different distortion types and SCI contents compared with state-of-the-art IQA algorithms. Finally, these SCIs and SQMS scores are used to learn the quality prediction model.

The training and testing procedures are illustrated in Fig. 8. Given the training original and distorted SCIs, the divergence features and the SQMS scores are acquired. Accordingly, the quality prediction model is trained with the SVR, which is given by

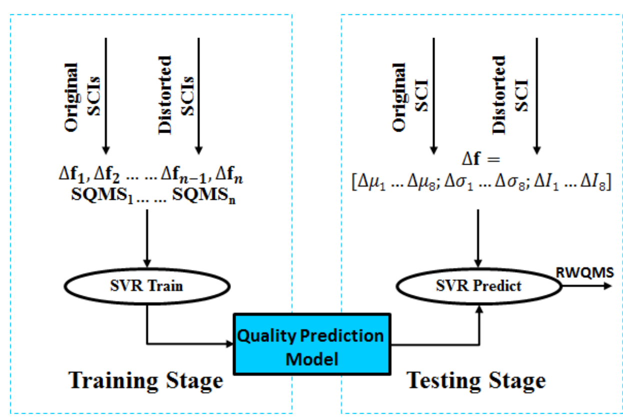


Fig. 8. Illustration of the training and testing procedures.

$$\text{model} = \text{SVR\_Train}\{\boldsymbol{\Delta}\mathbf{f_1}, \ldots \boldsymbol{\Delta}\mathbf{f_n}, \text{SQMS}_\mathbf{1}, \ldots, \text{SQMS}_\mathbf{n}\} \tag{8}$$

where n denotes the number of SCIs used in training.

In the testing procedure, given the divergence feature vector $\boldsymbol{\Delta}\mathbf{f}$, the RR SCI quality of the proposed RWQMS method is

TABLE I

PERFORMANCE COMPARISON WITH STATE-OF-THE-ART FR AND RR IQA ALGORITHMS BASED ON THE SIQAD DATABASE

IQA Model	SSIM	PSNR	VSI	GSIM	VSNR	SQMS	DNT-RR	EPM	WNISM	FTB	SDM	RIQMC	RWQMS
PLCC	0.5912	0.5869	0.5568	0.5686	0.5966	0.8872	0.5291	0.6711	0.5857	0.4691	0.6034	0.2732	0.8103
MAE	9.0934	9.0393	9.2875	9.1663	8.8284	5.2926	9.7913	8.5726	9.4566	10.132	9.0139	11.316	6.8200
RMSE	11.545	11.589	11.890	11.775	11.487	6.6039	12.147	10.612	11.602	12.641	11.414	13.769	8.3892
SRCC	0.5836	0.5608	0.5381	0.5483	0.5703	0.8803	0.5054	0.6529	0.5188	0.4575	0.6020	0.2395	0.7815
KRCC	0.4235	0.4226	0.3874	0.4054	0.4381	0.6936	0.3615	0.4582	0.3540	0.3268	0.4322	0.1627	0.5835

TABLE II

PERFORMANCE COMPARISON WITH STATE-OF-THE-ART FR AND RR IQA ALGORITHMS BASED ON THE QACS DATABASE

IQA Model	SSIM	PSNR	VSI	GSIM	VSNR	SQMS	DNT-RR	EPM	WNISM	FTB	SDM	RIQMC	RWQMS
PLCC	0.8764	0.8669	0.8715	0.8921	0.7050	0.9059	0.8083	0.6658	0.6326	0.6864	0.6590	0.4241	0.8489
MAE	0.8392	0.8759	0.8337	0.7752	1.1869	0.6949	1.0292	1.3207	1.3978	1.2768	1.3974	1.6776	0.9161
RMSE	1.0684	1.1059	1.0879	1.0025	1.5733	0.9396	1.3062	1.6552	1.7182	1.6134	1.6686	2.0091	1.1727
SRCC	0.8829	0.8656	0.8719	0.8947	0.7172	0.9096	0.8094	0.6552	0.6154	0.6887	0.7463	0.3489	0.8504
KRCC	0.7072	0.6768	0.6941	0.7215	0.5383	0.7470	0.6198	0.4697	0.4352	0.5048	0.5467	0.2502	0.6606

TABLE III

PERFORMANCE COMPARISON FOR INDIVIDUAL DISTORTION TYPE IN SIQAD (THE TOP TWO METHODS ARE HIGHLIGHTED)

SIQAD	Gaussian Noise		Gaussian Blur		Motion Blur		Contrast Change		Compression	
	PLCC	SRCC	PLCC	SRCC	PLCC	SRCC	PLCC	SRCC	PLCC	SRCC
DNT-RR	0.8189	0.8211	**0.8946**	**0.8875**	**0.7928**	**0.7903**	**0.7846**	**0.6719**	0.5987	0.5867
EPM	0.8420	0.8151	0.8318	0.8293	0.7143	0.7112	0.7151	0.5090	0.5468	0.5168
WNISM	**0.8570**	**0.8442**	0.8524	0.8370	0.6618	0.6606	0.7402	0.6142	0.3170	0.2991
FTB	0.7185	0.7165	0.7358	0.7400	0.5984	0.5866	0.5207	0.1112	0.5798	0.5649
SDM	**0.8694**	**0.8635**	0.7836	0.8199	0.5434	0.5307	0.7831	0.6617	**0.6837**	**0.6777**
RIQMC	0.4260	0.4144	0.3235	0.3169	0.2772	0.3034	0.5479	0.4506	0.3024	0.2869
RWQMS	0.8171	0.8080	**0.9088**	**0.9077**	**0.8738**	**0.8728**	**0.8049**	**0.6873**	**0.6386**	**0.6095**

predicted by the quality prediction model

$$\text{RWQMS} = \text{SVR\_Predict}(\boldsymbol{\Delta}\mathbf{f}, \text{model}). \tag{9}$$

Here the LIBSVM package [60] is used in the SVR process [45].

IV. EXPERIMENTAL RESULTS

A. Protocol

To validate the proposed RWQMS algorithm, both the SIQAD and QACS datasets are used for testing. Again, the QACS database contains 492 distorted SCIs with two distortion types and SIQAD contains 980 distorted SCIs with seven distortion types. The proposed method is compared with state-of-the-art RR and FR IQA algorithms. The RR-IQA methods include WNISM [34], DNT-RR [35], EPM [42], FTB [40], SDM [41] and RIQMC [43]. In addition, the popular FR algorithms including PSNR, SSIM [61], VSNR [22], GSIM [23], VSI [24] and SQMS [25] are compared as well.

To evaluate the performance of different IQA measures, five evaluation metrics are reported, including Spearman rank correlation coefficient (SRCC), Pearson linear correlation coefficient (PLCC), mean absolute error (MAE), root mean-squared error (RMSE), and Kendall's rank correlation coefficient (KRCC). Specifically, as specified in Video Quality Experts Group (VQEG) Phase I FR-TV test [62], PLCC, MAE and RMSE evaluate the prediction accuracy by performing a nonlinear mapping between the subjective and objective scores. The SRCC and KRCC evaluate the prediction monotonicity, which are independent of any monotonic score mapping. A better objective quality measure is expected to achieve higher values in PLCC, SRCC, and KRCC, and lower values in RMSE and MAE.

TABLE IV

PERFORMANCE COMPARISON FOR INDIVIDUAL DISTORTION TYPE IN QACS (THE TOP TWO METHODS ARE HIGHLIGHTED)

QACS	HEVC		HEVC_SCCEXT	
	PLCC	SRCC	PLCC	SRCC
DNT-RR	**0.8200**	**0.8199**	**0.7804**	**0.7813**
EPM	0.6540	0.6371	0.6504	0.6302
WNISM	0.6627	0.6423	0.5948	0.5479
FTB	0.6267	0.6266	0.7245	0.7376
SDM	0.6499	0.7395	0.6740	0.7205
RIQMC	0.5033	0.3382	0.4959	0.3354
RWQMS	**0.8478**	**0.8437**	**0.8407**	**0.8452**

B. Performance Evaluation

The validation results of the two screen content databases are shown in Tables I and II, where it can be observed that the proposed RWQMS significantly outperforms the state-of-the-art RR algorithms with respect to both prediction accuracy and monotonicity. Although it is usually unfair to directly compare the RR-IQA with the FR-IQA approaches, the performance of the FR-IQA approaches can still provide a good indication on the status of the RR-IQA methods. Regarding SIQAD, it can be observed that the proposed algorithm achieves better prediction performance compared with the FR-IQA methods except SQMS, which was specifically developed for SCIs. For QACS database, the RWQMS performs slightly inferior to the state-of-the-art FR algorithms. Moreover, we further evaluate the breakdown prediction performance for the individual distortion type, and the results are shown in Tables III and IV. It is worth

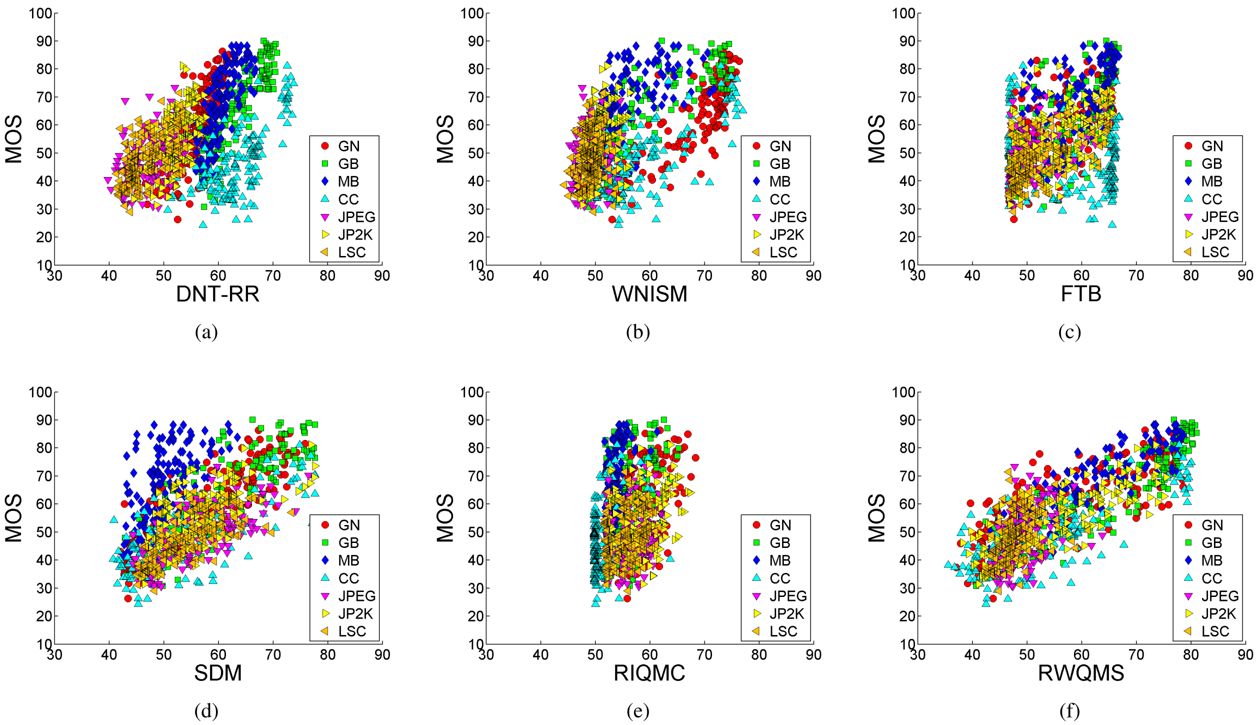


Fig. 9. Scatter plots of the MOS vs objectively predicted scores on SIQAD.

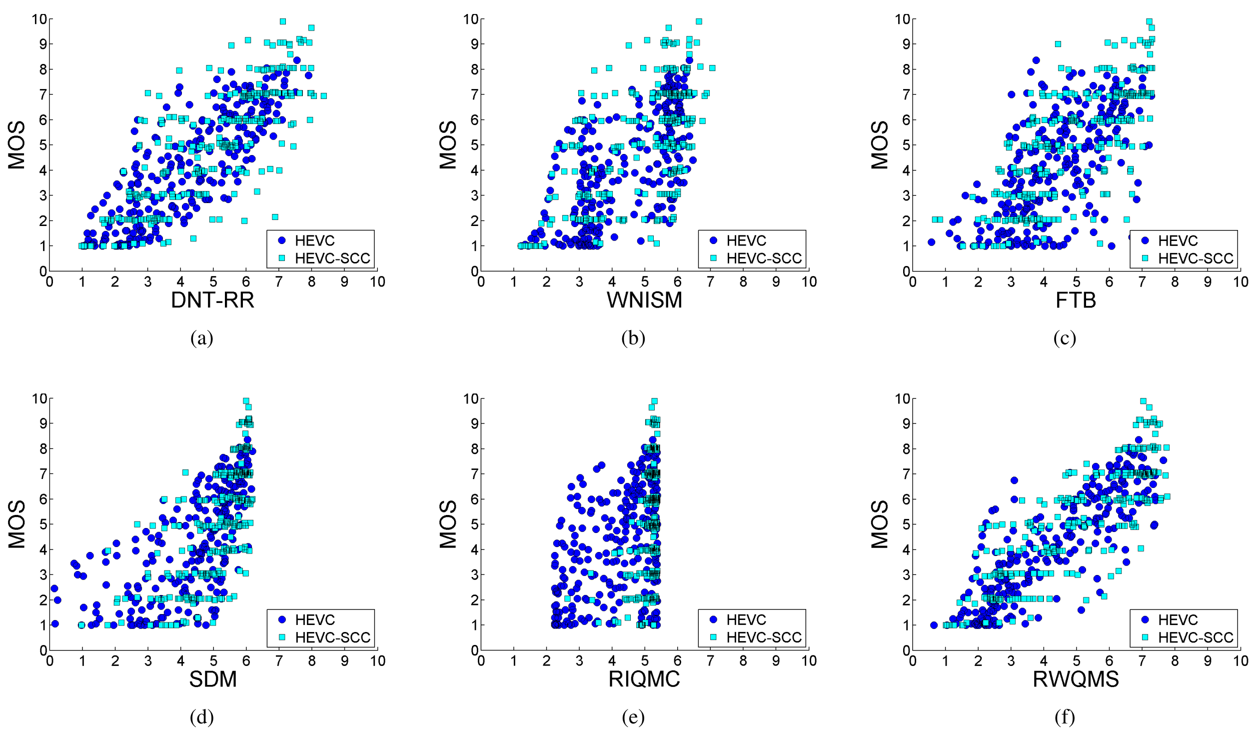


Fig. 10. Scatter plots of the MOS vs objectively predicted scores on QACS.

noting that for SIQAD, the compression distortion includes JPEG, JPEG2000 and Layered compression. It can be seen that, in general, for most cases the proposed method is among the best. Scatter plots of objective prediction after regression versus human ratings are shown in Figs. 9 and 10. It is also observed that the proposed method can accurately predict the human rating scores across different distortion types.

To further evaluate the performance of the proposed method, following [63] we carry out a statistical significant analysis. Specifically, the analysis is based on the variance-based hypothesis testing, in which the residuals difference between the DMOS and the predicted scores from the objective IQA methods are assumed to follow the Gaussian distribution. The results of SIQAD and QACS are provided in Table V and VI,

TABLE V
STATISTICAL SIGNIFICANCE EVALUATION OF THE PROPOSED METHOD OVER THE STATE-OF-THE-ART METHODS BASED ON SIQAD

IQA Model	SSIM	PSNR	VSI	GSIM	VSNR	SQMS	DNT-RR	EPM	WNISM	FTB	SDM	RIQMC
Proposed	1	1	1	1	1	0	1	1	1	1	1	1

TABLE VI
STATISTICAL SIGNIFICANCE EVALUATION OF THE PROPOSED METHOD OVER THE STATE-OF-THE-ART METHODS BASED ON QACS

IQA Model	SSIM	PSNR	VSI	GSIM	VSNR	SQMS	DNT-RR	EPM	WNISM	FTB	SDM	RIQMC
Proposed	0	-	-	0	1	0	1	1	1	1	1	1

TABLE VII
COMPARISONS BETWEEN THE DIRECT SUMMATIONS OVER RHE FEATURE DIVERGENCE AND THE PROPOSED METHODS

	SIQAD		QACS	
	Direct	SVR	Direct	SVR
PLCC	0.6812	0.8103	0.5784	0.8489
MAE	8.4918	6.8200	1.5346	0.9161
RMSE	10.4793	8.3892	1.8098	1.1727
SRCC	0.3404	0.7815	0.6942	0.8504
KRCC	0.2411	0.5835	0.5014	0.6606

TABLE VIII
COMPARISONS OF DIFFERENT FEATURE DIVERGENCE COMPUTATION STRATEGIES

	SIQAD			QACS		
	Proposed	Divide	All	Proposed	Divide	All
PLCC	0.810	0.767	0.492	0.849	0.725	0.543
MAE	6.820	7.497	10.164	0.916	1.211	1.577
RMSE	8.389	9.186	12.461	1.173	1.529	1.863
SRCC	0.782	0.692	0.478	0.850	0.710	0.596
KRCC	0.584	0.502	0.340	0.661	0.523	0.431

where the symbol "1" denotes that the proposed IQA method is statistically better than that of the column, "0" denotes that the IQA method of the column is better than the proposed method, and "-" means that the two methods are statistically indistinguishable. It can be observed that the proposed model is statistically superior to all the compared RR IQA algorithms. Moreover, for SIQAD database, the proposed method is only inferior to SQMS and performs better than the other FR methods. For QACS database, the proposed method achieves statistically similar performance compared with the VSI algorithm, better than VSNR, and inferior to the SQMS, SSIM and GSIM.

To further verify the proposed learning based RR-IQA algorithm, we compare the proposed IQA measure (with SVR) with the method of direct quality prediction (without SVR). Specifically, we use the direct summations over the feature divergence vector $\mathbf{\Delta f}$ to predict the SCI quality. The comparison results are shown in Table VII, where "Direct" indicates the direct summation approach. It is not surprising to see that the proposed method significantly outperforms the "Direct" approach, which further provides useful evidence of the advantages of the learning strategy in RWQMS.

The proposed method employs the linear difference in feature space to compute the feature divergence, which serves as the input to the learning model. To provide more evidence regarding the effectiveness of this method, an experiment is conducted by comparing the performance with two alternative approaches. Specifically, the first approach computes the feature divergence by dividing the feature of the distorted SCI by that of the original SCI, and the second method treats all 48 features as the input to SVR. The comparison results are shown in Table VIII, where "Divide" denotes the first and "All" denotes the second approach, respectively. It can be observed that the proposed method is superior to the two alternative approaches. This implies that computing the linear difference in feature space provides useful information about the quality variations and leads to significant performance improvement in terms of SCI quality prediction accuracy.

TABLE IX
PERFORMANCE COMPARISONS USING 80%/20% TRAINING/TESTING SPLITS WITH 1000 ITERATIONS IN SIQAD

	PLCC	SRCC	KRCC
RWQMS	0.8421	0.8339	0.6436
DIIVINE	0.6890	0.6591	0.4798
BLIINDS-II	0.7264	0.6841	0.5022
BRISQUE	0.7855	0.7561	0.5683

C. Comparisons With Learning based IQA Algorithms

In this subsection, we compare our approach with three learning based algorithms, including BLIINDS-II [31], DIIVINE [32] and BRISQUE [33]. Specifically, we split the database according to the content (80% for training, 20% for testing) and retrain these algorithms. The resulting median PLCC, SRCC and KRCC are obtained over 1000 iterations of random train/test splits. The results are demonstrated in Table IX, from which we can observe that the proposed algorithm is superior to these learning based algorithms.

D. Complexity Comparison

The complexity of the IQA algorithms is evaluated by the actual running time. Specifically, we run these algorithms on the SIQAD database with a computer of Intel I7-4790 CPU@3.60GHz and 8GB RAM. The software platform is MATLAB R2014. In Table X, we tabulate the running time of different IQA methods. The average running time in terms of seconds/pic is computed. In particular, regarding the proposed method only the computational complexity of the feature extraction is taken into account. It can be observed that the computational complexity of the proposed method is much less than the other RR-IQA algorithms and comparable to the state-of-the-art FR methods such as VSI. The low complexity computation can also enable its applications in SCI communication scenarios.

TABLE X
COMPUTATIONAL COST OF THE PROPOSED AND 12 IQA METHODS

IQA Model	SSIM	PSNR	VSI	GSIM	VSNR	SQMS	DNT-RR	EPM	WNISM	FTB	SDM	RIQMC	RWQMS
Time (seconds/image)	0.015	0.002	0.169	0.019	0.199	0.071	4.061	0.453	1.838	0.735	0.356	0.414	0.135

E. Discussions

Our RWQMS method has been shown to achieve better prediction in both SIQAD and QACS databases based on a set of wavelet domain features. The superior performance originates from using large amount of training data labeled by the state-of-the-art FR-IQA algorithm SQMS to automatically learn the quality prediction model. Since these features are fast and easy to calculate, it is practical to deploy such algorithm in the interactive screen-remoting system to real-time monitor the SCI quality. In particular, in the scenario of SCI communication, by embedding the extracted features in the bitstream, at the receiver side the perceptual quality predicted by the proposed model can be feasibly acquired to monitor the perceptual degradations after compression and transmission. As one of the first attempts that apply machine learning in addressing the SCI RR-IQA problem, the proposed algorithm can also be improved from the perspectives of effectiveness (in terms of prediction accuracy) and efficiency (in terms of the computational complexity and transmission overhead) in the future. Firstly, we will continue to develop more accurate SCI FR-IQA measure that can provide us the trusted quality score to learn the quality prediction model. As such, the RR-IQA performance can accordingly get improved. Secondly, in addition to the distortion created by compression, more distortion types will be considered. Thirdly, how are these extracted statistical features that can reflect the quality of SCI related and what is the best way to explore minimal number of features to predict the SCI quality will also be studied. Finally, the extension of the wavelet domain features to the NR SCI IQA needs further exploration, which is also very meaningful from a practical point of view.

V. CONCLUSION

The major contributions of this paper are twofold. We create a new subjective SCI database and learn a reduced reference quality measure that can be feasibly applied in monitoring the end-to-end quality of SCIs. In particular, the database has the unique feature that includes the distortions created by both HEVC and HEVC-SCC extension. Extensive subjective experiments have been conducted to evaluate the quality of the compressed SCIs. Moreover, the specifically learnt reduced reference objective model automatically predicts the quality of the SCIs with a few extracted features, and experimental results show that the proposed method can accurately predict the SCI quality across different types of distortions.

REFERENCES

[1] Y. Lu, S. Li, and H. Shen, "Virtualized screen: A third element for cloud-mobile convergence," *IEEE Multimedia Mag.*, vol. 18, no. 2, pp. 4–11, Feb. 2011.

[2] H. Shen, Z. Pan, H. Sun, Y. Lu, and S. Li, "A proxy-based mobile Web browser," in *Proc. Int. Conf. Multimedia*, 2010, pp. 763–766.

[3] H. Shen and Y. Lu, "A high-performanance remote computing platform," in *Proc. IEEE Int. Conf. Pervas. Comput. Commun.*, Mar. 2009, pp. 1–6.

[4] C. Lan, G. Shi, and F. Wu, "Compress compound images in H.264/MPGE-4 AVC by exploiting spatial correlation," *IEEE Trans. Image Process.*, vol. 19, no. 4, pp. 946–957, Apr. 2010.

[5] T. Lin, P. Zhang, S. Wang, K. Zhou, and X. Chen, "Mixed chroma sampling-rate high efficiency video coding for full-chroma screen content," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 23, no. 1, pp. 173–185, Jan. 2013.

[6] 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.

[7] S. Wang, A. Rehman, Z. Wang, S. Ma, and W. Gao, "SSIM-motivated rate-distortion optimization for video coding," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 22, no. 4, pp. 516–529, Apr. 2012.

[8] S. Wang, A. Rehman, Z. Wang, S. Ma, and W. Gao, "Perceptual video coding based on SSIM-inspired divisive normalization," *IEEE Trans. Image Process.*, vol. 22, no. 4, pp. 1418–1429, Apr. 2013.

[9] 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.

[10] G. J. Sullivan, J.-R. Ohm, W.-J. Han, and T. Wiegand, "Overview of the high efficiency video coding (HEVC) standard," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 22, no. 12, pp. 1649–1668, Dec. 2012.

[11] 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.

[12] A. Gabriellini, M. Naccari, M. Mrak, and D. Flynn, "Spatial transform skip in the emerging high efficiency video coding standard," in *Proc. IEEE Int. Conf. Image Process.*, Sep./Oct. 2012, pp. 185–188.

[13] A. Gabriellini, M. Naccari, M. Mrak, D. Flynn, and G. Van Wallendael, "Adaptive transform skipping for improved coding of motion compensated residuals," *Signal Process., Image Commun.*, vol. 28, no. 3, pp. 197–208, Mar. 2013.

[14] M. Mrak and J.-Z. Xu, "Improving screen content coding in HEVC by transform skipping," in *Proc. 20th Eur. Signal Process. Conf. (EUSIPCO)*, Aug. 2012, pp. 1209–1213.

[15] W. Zhu, W. Ding, J. Xu, Y. Shi, and B. Yin, "Screen content coding based on HEVC framework," *IEEE Trans. Multimedia*, vol. 16, no. 5, pp. 1316–1326, Aug. 2013.

[16] D. K. Kwon and M. Budagavi, *RCE3: Results of Test 3.3 on Intra Motion Compensation*, document JCTVC-n0205, 14th Meeting, Vienna, Austria, 2013, p. 2013.

[17] S. Shi, X. Zhang, S. Wang, R. Xiong, and S. Ma, "Study on subjective quality assessment of screen content images," in *Proc. Picture Coding Symp. (PCS)*, May/Jun. 2015, pp. 75–79.

[18] 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.

[19] Y. Fang *et al.*, "Objective quality assessment for image retargeting based on structural similarity," *IEEE J. Emerg. Sel. Topics Circuits Syst.*, vol. 4, no. 1, pp. 95–105, Mar. 2014.

[20] Z. Wang, E. P. Simoncelli, and A. C. Bovik, "Multiscale structural similarity for image quality assessment," in *Proc. 37th Asilomar Conf. Signals, Syst. Comput.*, vol. 2. Nov. 2003, pp. 1398–1402.

[21] S. Wang, K. Ma, H. Yeganeh, Z. Wang, and W. Lin, "A patch-structure representation method for quality assessment of contrast changed images," *IEEE Signal Process. Lett.*, vol. 22, no. 12, pp. 2387–2390, Dec. 2015.

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

[23] 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.

[24] 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.

[25] K. Gu *et al.*, "Saliency-guided quality assessment of screen content images," *IEEE Trans. Multimedia*, vol. 18, no. 6, pp. 1098–1110, Jun. 2016.
[26] Q. Wu *et al.*, "Blind image quality assessment based on multichannel feature fusion and label transfer," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 26, no. 3, pp. 425–440, Mar. 2016.
[27] M. Saad, P. Corriveau, and R. Jaladi, "Consumer content framework for blind photo quality evaluation," in *Proc. 9th Int. Workshop Video Process. Quality Metrics Consum. Electron.*, 2015, pp. 1–6.
[28] W. Hou, X. Gao, D. Tao, and X. Li, "Blind image quality assessment via deep learning," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 26, no. 6, pp. 1275–1286, Jun. 2015.
[29] Q. Wang, J. Chu, L. Xu, and Q. Chen, "A new blind image quality framework based on natural color statistic," *Neurocomputing*, vol. 173, pp. 1798–1810, Jan. 2016.
[30] Q. Wu, Z. Wang, and H. Li, "A highly efficient method for blind image quality assessment," in *Proc. IEEE Int. Conf. Image Process.*, Sep. 2015, pp. 339–343.
[31] M. A. Saad, A. C. Bovik, and C. Charrier, "Blind image quality assessment: A natural scene statistics approach in the DCT domain," *IEEE Trans. Image Process.*, vol. 21, no. 8, pp. 3339–3352, Aug. 2012.
[32] A. K. Moorthy and A. C. Bovik, "Blind image quality assessment: From natural scene statistics to perceptual quality," *IEEE Trans. Image Process.*, vol. 20, no. 12, pp. 3350–3364, Dec. 2011.
[33] 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.
[34] Z. Wang and E. P. Simoncelli, "Reduced-reference image quality assessment using a wavelet-domain natural image statistic model," *Proc. SPIE*, vol. 5666, pp. 149–159, Mar. 2005.
[35] Q. Li and Z. Wang, "Reduced-reference image quality assessment using divisive normalization-based image representation," *IEEE J. Sel. Topics Signal Process.*, vol. 3, no. 2, pp. 202–211, Apr. 2009.
[36] 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.
[37] S. Wang, X. Zhang, S. Ma, and W. Gao, "Reduced reference image quality assessment using entropy of primitives," in *Proc. Picture Coding Symp. (PCS)*, Dec. 2013, pp. 193–196.
[38] 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.
[39] L. Ma, S. Li, F. Zhang, and K. N. Ngan, "Reduced-reference image quality assessment using reorganized DCT-based image representation," *IEEE Trans. Multimedia*, vol. 13, no. 4, pp. 824–829, Aug. 2011.
[40] M. Narwaria, W. Lin, I. V. McLoughlin, S. Emmanuel, and L. T. Chia, "Fourier transform-based scalable image quality measure," *IEEE Trans. Image Process.*, vol. 21, no. 8, pp. 3364–3377, Aug. 2012.
[41] K. Gu, G. Zhai, X. Yang, and W. Zhang, "A new reduced-reference image quality assessment using structural degradation model," in *Proc. IEEE Int. Symp. Circuits Syst.*, May 2013, pp. 1095–1098.
[42] M. Zhang, W. Xue, and X. Mou, "Reduced reference image quality assessment based on statistics of edge," *Proc. SPIE*, vol. 7876, pp. 787611-1–787611, Jan. 2011.
[43] K. Gu, G. T. Zhai, and M. Lin, "The analysis of image contrast: From quality assessment to automatic enhancement," *IEEE Trans. Cybern.*, vol. 46, no. 1, pp. 284–297, Jan. 2015.
[44] Z. Wang and A. C. Bovik, "Reduced- and no-reference image quality assessment," *IEEE Signal Process. Mag.*, vol. 28, no. 6, pp. 29–40, Nov. 2011.
[45] B. Schölkopf, A. J. Smola, R. C. Williamson, and P. L. Bartlett, "New support vector algorithms," *Neural Comput.*, vol. 12, no. 5, pp. 1207–1245, May 2000.
[46] L. He, D. Tao, X. Li, and X. Gao, "Sparse representation for blind image quality assessment," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2012, pp. 1146–1153.
[47] C. Shen and Q. Zhao, "Webpage saliency," in *Proc. ECCV*, 2014, pp. 33–46.
[48] H. Yu, R. Cohen, K. Rapaka, and J. Xu, *Common Test Conditions for Screen Content Coding*, document JCTVC-S1015, Strasbourg, France, 2014.
[49] L. Ma, W. Lin, C. Deng, and K. N. Ngan, "Image retargeting quality assessment: A study of subjective scores and objective metrics," *IEEE J. Sel. Topics Signal Process.*, vol. 6, no. 6, pp. 626–639, Oct. 2012.
[50] Methodology for the Subjective Assessment of the Quality of Television Pictures, document ITU-R Rec. BT.500-11, ITU, 2000.
[51] QACS Database, accessed on 2016-1-17. [Online]. Available: https://www.dropbox.com/s/kstmdpiqtn8ekje/QACS.rar?dl=0
[52] T. Brandão and M. P. Queluz, "No-reference image quality assessment based on DCT domain statistics," *Signal Process.*, vol. 88, no. 4, pp. 822–833, Apr. 2008.
[53] S. Wang, K. Gu, K. Zeng, Z. Wang, and W. Lin, "Perceptual screen content image quality assessment and compression," in *Proc. IEEE Int. Conf. Image Process.*, Sep. 2015, pp. 1434–1438.
[54] K. Gu, G. Zhai, X. Yang, and W. Zhang, "Using free energy principle for blind image quality assessment," *IEEE Trans. Multimedia*, vol. 17, no. 1, pp. 50–63, Jan. 2015.
[55] B. A. Olshausen, P. Sallee, and M. S. Lewicki, "Learning sparse image codes using a wavelet pyramid architecture," in *Proc. Adv. Neural Inf. Process. Syst.*, 2001, pp. 887–893.
[56] B. A. Olshausen and D. J. Field, "Wavelet-like receptive elds emerge from a network that learns sparse codes for natural images," *Nature*, vol. 381, pp. 607–609, Apr. 1996.
[57] 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.
[58] W. S. Geisler, "Visual perception and the statistical properties of natural scenes," *Annu. Rev. Psychol.*, vol. 59, pp. 167–192, Jan. 2008.
[59] S. Mallat and S. Zhong, "Characterization of signals from multiscale edges," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 14, no. 7, pp. 710–732, Jul. 1992.
[60] C.-C. Chang and C.-J. Lin, "LIBSVM: A library for support vector machines," *ACM Trans. Intell. Syst. Technol.*, vol. 2, no. 3, pp. 27:1–27:27, 2011.
[61] 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.
[62] A. M. Rohaly *et al.*, "Final report from the video quality experts group on the validation of objective models of video quality assessment," *ITU-T Standards Contrib. COM*, vol. 1, pp. 9–80, 2000.
[63] H. R. Sheikh, M. F. Sabir, and A. C. Bovik, "A statistical evaluation of recent full reference image quality assessment algorithms," *IEEE Trans. Image Process.*, vol. 15, no. 11, pp. 3440–3451, Nov. 2006.

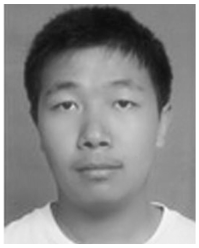
Shiqi Wang (M'15) received the B.S. degree in computer science from the Harbin Institute of Technology, Harbin, China, in 2008, and the Ph.D. degree in computer application technology from the Peking University, Beijing, China, in 2014. He was a Postdoc Fellow with the Department of Electrical and Computer Engineering, University of Waterloo, Waterloo, ON, Canada.

He is currently with the Rapid-Rich Object Search Laboratory, Nanyang Technological University, Singapore, as a Research Fellow. He has authored and co-authored over 70 technical articles in refereed journals and proceedings in the areas of image and image/video coding, processing, quality assessment and analysis. He has also proposed more than 20 technical proposals to ISO/MPEG, ITU-T and AVS video coding standards.

Ke Gu (M'13) received the B.S. and PhD degrees in electronic engineering from Shanghai Jiao Tong University, Shanghai, China, in 2009 and 2015, respectively.

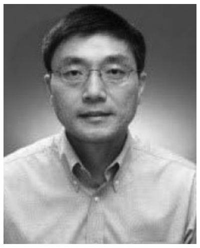
From July 2014 to November 2014, he was visiting at the Department of Electrical and Computer Engineering, University of Waterloo, Canada. From December 2014 to January 2015, he was visiting at the School of Computer Engineering, Nanyang Technological University, Singapore. From February 2015 to March 2015 and from November 2015 to December 2015, he was a visiting student at the Department of Computer Science and Technology, Peking University, Beijing, China. His research interests include quality assessment, contrast enhancement and visual saliency detection.

Dr. Gu is a Special Session Organizer for VCIP 2016.



Xiang Zhang (M'15) received the B.S. degree in computer science from the Harbin Institute of Technology, Harbin, China, in 2013. He is currently working toward the Ph.D. degree at Peking University.

His research interests include image/video quality assessment, video compression and visual retrieval.



Weisi Lin (SM'98–F'16) received the Ph.D. degree from Kings College, London University, London, U.K., in 1993.

He is currently an Associate Professor with the School of Computer Engineering, Nanyang Technological University, and served as a Lab Head of Visual Processing, Institute for Infocomm Research. He authors over 300 scholarly publications, holds seven patents, and receives over $ 4 million in research grant funding. He has maintained active long-term working relationship with a number of companies. His research interests include image processing, video compression, perceptual visual and audio modeling, computer vision, and multimedia communication. He served as an Associate Editor of Journal of Visual Communication and Image Representation.

Dr. Lin served as an Associate Editor of IEEE Transactions on Multimedia, IEEE Signal Processing Letters. He is also on six IEEE Technical Committees and Technical Program Committees of a number of international conferences. He was the Lead Guest Editor for a special issue on perceptual signal processing of the IEEE Journal of Selected Topics in Signal Processing in 2012. He is a Chartered Engineer in the U.K., a Fellow of the Institution of Engineering Technology, and an Honorary Fellow of the Singapore Institute of Engineering Technologists. He Co-Chaired the IEEE MMTC special interest group on quality of experience. He was an Elected Distinguished Lecturer of APSIPA in 2012/3.

Li Zhang (M'07) received the Ph.D. degree in computer science from the Institute of Computing Technology, Chinese Academy of Sciences, Beijing, China, in 2009.

Her research interests include video coding, processing and communication.

Siwei Ma (S'03–M'12) received the B.S. degree from Shandong Normal University, Jinan, China, in 1999, and the Ph.D. degree in computer science from the Institute of Computing Technology, Chinese Academy of Sciences, Beijing, China, in 2005.

From 2005 to 2007, he held a post-doctorate position with the University of Southern California, Los Angeles. Then, he joined the Institute of Digital Media, School of Electronic Engineering and Computer Science, Peking University, Beijing, China, where he is currently a Professor. He has published over 100 technical articles in refereed journals and proceedings in the areas of image and video coding, video processing, video streaming, and transmission.

Wen Gao (M'92–SM'05–F'09) received the Ph.D. degree in electronics engineering from the University of Tokyo, Tokyo, Japan, in 1991.

He is a Professor of computer science at Peking University, Beijing, China. Before joining Peking University, he was a Professor of computer science at Harbin Institute of Technology from 1991 to 1995, and a Professor at the Institute of Computing Technology of Chinese Academy of Sciences. He has published extensively including five books and over 600 technical articles in refereed journals and conference proceedings in the areas of image processing, video coding and communication, pattern recognition, multimedia information retrieval, multimodal interface, and bioinformatics. He served or serves on the editorial board for several journals, such as EURASIP Journal of Image Communications and Journal of Visual Communication and Image Representation.

Dr. Gao served or serves on the editorial board for several journals, such as IEEE Transactions on Circuits and Systems for Video Technology, IEEE Transactions on Multimedia, IEEE Transactions on Image Processing, and IEEE Transactions on Autonomous Mental Development. He chaired a number of prestigious international conferences on multimedia and video signal processing, such as IEEE ICME and ACM Multimedia, and also served on the advisory and technical committees of numerous professional organizations.