RESEARCH ARTICLE   OPEN ACCESS

Local and Global Feature Learning for Blind Quality Evaluation of Screen Content and Natural Scene Images

Dr K Sailaja, MCA, M.Tech, M.Phil, Ph.D [1], P Aswani [2]

[1] Professor & HOD, Department of Computer Application
[2] Student, Department of Computer Application
[1], [2] Chadalawada Ramanamma Engineering College (Autonomous)

ABSTRACT

An increasingly crucial yet difficult problem is the blind quality assessment of screen content images (SCIs) and natural scene pictures (NSIs). In this research, we provide a dictionary of learnt local and global quality attributes that forms the basis of a blind quality assessment approach for SCIs and NSIs. To begin, we use local normalised picture patches and traditional K -means clustering to build a local lexicon. Using a locality-constrained linear code with maximum pooling, the learnt local quality characteristics may be derived with the help of this specialised dictionary. The histogram representations of binary patterns are joined to produce a global dictionary, which is then used to extract the learnt global quality characteristics. Using this dictionary, the collaborative representation technique effectively codes the learnt global quality attributes of the distorted pictures. Finally, a quality score is derived by combining all of these factors using a kernel-based support vector regression model. Extensive studies with the suggested assessment method show that the blind measure produces considerably greater consistency in line with subjective fidelity evaluations than do other comparable metrics.

Keywords: - Screen content images (SCIs), natural scene images (NSIs), image quality assessment, local and global features, locality-constrained linear coding, collaborative representation.

I. INTRODUCTION

Screen content images (SCIs) and natural scene pictures (NSIs) have been more prevalent and intimately used in everyday life as the Internet has become more widely accessible in recent years. In addition, SCIs and NSIs are becoming more popular in a wide range of multimedia applications for computers and other electronic devices, such as those used for visual screen sharing, distance learning, cloud computing, online games, instant messaging, and snapping photos. [1]–[3]. Blurring, noising, compression artefacts, contrast shift, quantization, and transmission loss are only some of the ways in which SCIs and NSIs are degraded during the capture, compression, storage, and transmission processes in multiclient communication systems. [4]–[7]. Therefore, in order to create, monitor, and enhance the performance of each processing step, precise approaches for measuring the perceptual quality of SCIs and NSIs are necessary. What this means is that a performance index based on a quality assessment model for SCIs and NSIs may be utilised to enhance compression effectiveness. Sender may also adjust SCI and NSI quality to meet specific needs. As a result, there is a lot of curiosity in this issue among academics [8]-[16]. There are two main categories of image quality assessment (IQA) techniques: subjective and objective techniques. [8]–[10]. Objective approaches give a quantifiable objective criterion for the perceived quality of distorted pictures, whereas subjective

methods rely on the subjective opinions of humans. Subjective IQA using humans is the most accurate and natural way to measure perceptual quality, but it also has a lot of drawbacks, such as being time consuming, difficult, and inconvenient. Specifically, real-time or automated systems are not suitable for implementing subjective IQA [8, 10]. Consequently, it is important to have objective IQA measures that can predict picture quality both automatically and reliably.

There have been recent efforts to develop objective IQA indicators for SCIs and NSIs [11–16], [26], [27]. There are three main types of objective IQA metrics that may be broken down by the amount of reference data available: full-reference (FR), reduced-reference (RR), and blind/no-reference (NR). To evaluate the quality of a SCI or NSI, most researchers employ FR measures, which presume full knowledge of the reference data. Wang et al[11] .'s popular structural similarity index (SSIM) is a watershed moment in the development of FR-IQA measures. You might find some more related research in [12] - [17]. In recent years, numerous blind IQA metrics for NSIs have been studied extensively [19]–[28]. For instance, Ye and Doermann [19] proposed a visual codebook-based blind IQA metric to measure NSIs quality using histograms of codeword occurrences, but the codebook size is very large. Xue *et al*. [20] proposed a quality clustering (QAC) metric that learns a set of quality-predictive centroids. These centroids are then used as a codebook to calculate the quality of an NSI patch; hence, the final quality value of the overall NSI is inferred. Further,

Mittal *et al*. [21] presented the NSI quality evaluator (NIQE), which does not require training with human-scored distorted NSIs. Inspired by NIQE, Zhang *et al.* [22] developed an integrated-local NIQE (IL_NIQE) by integrating more NSI feature statistics. Further, Xue *et al.* [23] presented a blind natural NSI IQA metric that uses joint statistics of the gradient magnitude and Laplacian of Gaussian features (GM_LOG). In addition, Zhou *et al.* [24] developed a natural IQA metric by analyzing the usefulness and effectiveness of two complementary image components, i.e., the gradient phase and magnitude. Xu *et al.* [25] presented a blind IQA metric by aggregating soft-weighted high-order statistical differences between a small codebook and normalized NSI patches. From the image content perspective, SCIs can be regarded as a mixture of natural images, computer graphics, document images, and other components. Obviously, these conventional blind metrics may not perform well on SCIs, as the statistical properties of the pictorial and textual regions in such images are distinct from those of natural images. Furthermore, the same levels of distortion in distinct regions may yield differences in perceptual quality [10]. In other words, the application of blind IQA to SCIs is significantly more complex than it is for NSIs, because it is affected by the quality of the natural images, computer graphics, document images, and other content [12]–[17], [29], [30].

II. RELATEDWORKS

Blind quality assessment of SCIs has received little research attention; thus, only a small number of stud- ies have focused on this area [29]–[32]. For instance, Gu *et al.* [29] presented a blind quality metric (BQMS) for SCIs based on a new SCI statistical model. Further, Qian *et al.* [30] devised a blind SCI quality IQA metric utiliz- ing an edge-preserving filter-based free energy and structural degradation model. Gu *et al.* [32] proposed a blind SCI IQA metric through bag data learning. However, the performance improvement yielded by these methods is limited by insufficient consideration of the statistical properties of the SCIs. Thus, the efficacy of blind IQA metrics for SCIs can be improved significantly.

In the current state of the research described above, the existing blind IQA metrics are either developed for NSIs or designed for SCIs. Only a very few quality metrics work for both simultaneously [25], [31]. In practical multime- dia application systems, we may encounter cross-content-type images (e.g., NSIs, SCIs, and other image types). Efficient general blind IQA metrics that do not depend on image types are required in such circumstances. To further advance the development of blind IQA metrics for SCIs and NSIs, in this study, we proposed an effective blind quality metric for SCIs and NSIs that fuses learned local and global quality features to efficiently represent both the local fine details and global statistical structures of the images. The key contributions of our work are summed up as follows:

(1) The local and global dictionaries are preconstructed based on an image statistical model. With these local dic- tionaries, the learned local and global quality features can comprehensively characterize the features of the images (e.g., spots, lines, and corners, which are the basic elements of SCIs and NSIs).

(2) The learned local quality features can be obtained using a strategy of locality-constrained linear coding (LLC) with max pooling. Meanwhile, the collaborative representa- tion (CR) algorithm is used to efficiently code the learned global quality features of the distorted images using global dictionary.

(3) To the best of our knowledge, this is the first attempt to combine LLC-based quality local features and CR-based global quality features to achieve a fused representation for distorted SCIs and NSIs. The combined use of two types of features can effectively mitigate the respective shortcomings of the individual local and global features.

III. PROPOSED SYSTEM ARCHITECTURE

A flow chart of the proposed blind metric is depicted in Fig. 1. The process is composed of two stages: i) local and global feature learning and ii) perceived quality prediction. In the feature learning stage, the local and global dictionaries are preconstructed in advance based on an image statistical model. The local and global quality features are learned using these dictionaries and able to express the micro- and macro- structures of the distorted images. In the perceptual quality prediction stage, SVR is implemented to determine the overall quality score. The final human perception of visual signals is a synthesis of local and global visual perception [32]–[35]. The average distortion in an image influences the overall human visual per- ception to some degree, whereas an extremely distorted local region lowers the overall visual quality severely. Because of the fundamental difference in the computational methods for the local and global quality features, we expect the two fea- ture representations to provide different types of information. That is, most local quality features convey texture information for a given image patch. However, global quality features include contour representations, texture features, and shape descriptors. Global and local texture features provide different types of information about an image because the support over which the texture is calculated varies. Note that it is necessary to

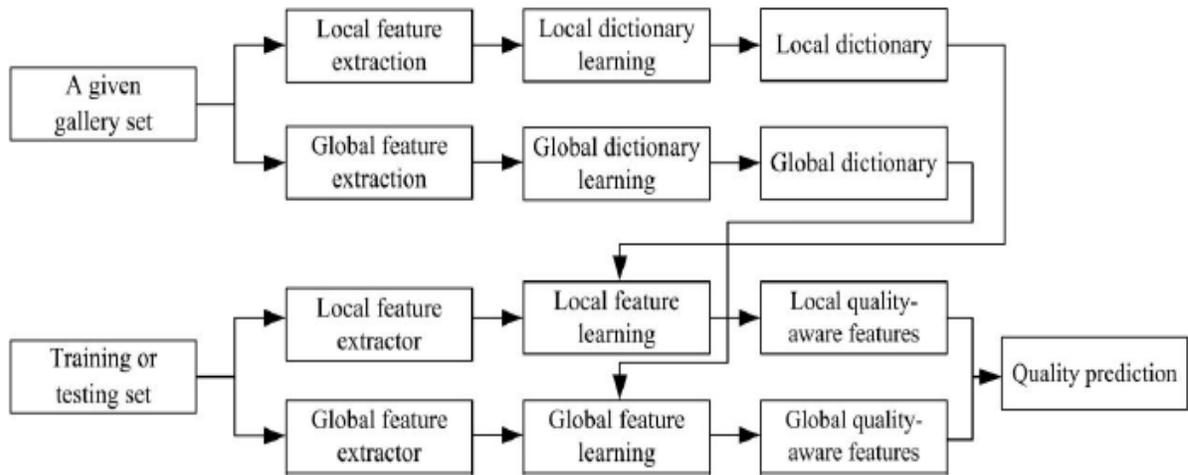consider both the local and global statistical properties   of the images when designing the proposed metric.

Fig. 1 A flow chart of the proposed blind metric.

Semantic structural alteration can generally reflect degra- dations in visual quality. Recent work indicated that structure descriptors (e.g., local binary patterns (LBPs) [44]) can effectively and efficiently represent the semantic structural information of visual signals and can be considered as the binary approximation of the semantic structural information primitives in the primary visual cortex [24]. The global quality feature learning method implemented in this work incorporates global dictionary learning and global feature representation. The HVS is highly sensitive to the edge profile representation that is often encountered in images [17], [18]. Because cortical and retinal neurons in the primary stage of vision respond to stimulus frequency and orientation, multiscale orientation filter responses are similar to the orientation and scale sensitivities of the visual receptive fields of HVS [45]. In this work, perceptually similar log-Gabor filters (kernels) are utilized for global feature extraction [46]. A phase change generally induces significant visual distortion. In this study, we first readjust the phase range to [0, 360). Motivated by the local binary pattern (LBP) strategy [44], we consider two cases when comparing feature types: different feature types or identical/similar feature types. To improve the stability of the binary quantification, the normalized phase range is divided into K intervals (here, we set $K= 4$). The local feature types are considered similar when they belong to the same interval; otherwise, they are considered different.

IV. RESULTS AND DISCUSSION

We conducted several experiments using three IQA datasets: the newly released SIQAD dataset [11], categorical subjec- tive image quality (CSIQ) dataset [51], and cross-content- type (CCT) dataset [31] to validate the effectiveness of the proposed blind metric. The SIQAD dataset contains 20 reference SCIs altered by seven types of distortion, each with seven levels of degradation. Thus, this dataset consists of 140 distorted SCIs, comprising 20 reference SCIs distorted by Gaussian noise (GN), motion blur (MB), Gaussian blur (GB), contrast change (CC), JPEG2000 compression (JP2K), JPEG compres- sion (JPEG), and layer-segmentation-based coding (LSC). For further details on this dataset, see [10]. The CSIQ dataset consists of 866 NISs and six types of distortion: white noise (WN), JPEG, JP2K, additive Gaussian pinknoise (PN), GB, and global contrast decrements (GCD). The difference mean opinion score (DMOS) of each NSI is included.

The CCT dataset, contains 1,320 distorted images (e.g., SCIs, NSIs, and others) and associated DMOS values generated from 72 pristine images distorted with two distortion types, at various levels of distortion.
Two typical performance criteria were employed to evaluate the IQA metrics tested in these experiments: i) Pearson's linear correlation coefficient (PLCC), which evaluates prediction accuracy, and ii) Spearman's rank order correlation coeffi- cient (SROCC), which reflects the prediction monotonic of IQA metrics. Higher PLCC and SROCC values values indicate superior correlation performance. Thus, a high-performance objective model has the PLCC and SROCC values are close to 1. As the proposed blind metric requires training and testing, a cross-validation test was implemented by randomly splitting each dataset into two non-overlapping subsets: training

(80%) and test (20%). To offset the performance bias as much as possible, this 80:20 split of the data was iterated 1,000 times and the median PLCC and SROCC values of the 1,000 trials are reported.

TABLE I PARAMETERS SELECTION EXPERIMENT RESULTS

Parameters		Criteria	
a	b	PLCC	SROCC
4	-7	0.816	0.776
4	-6	0.828	0.787
4	-5	0.828	0.788
4	-4	0.824	0.784
3	-4	0.822	0.783
2	-4	0.820	0.777

TABLE II PERFORMANCE OF THE PROPOSED BLIND METRIC AND THE OTHER TWELVE METRICS USING THE SIQAD DATASET

Metrics	Criteria	
	PLCC	SROCC
PSNR	0.587	0.560
SSIM	0.591	0.582
SPQA	0.858	0.842
SFUW	0.891	0.880
ESIM	0.879	0.863
QAC	0.375	0.301
NIQE	0.342	0.370
IL_NIQE	0.388	0.322
BQMS	0.755	0.722
GM_LOG	0.720	0.663
DIIVINE	0.691	0.659
LGP	0.784	0.754
Proposed	**0.828**	**0.788**

TABLE III MEAN AND STANDARD DEVIATION OF THE SROCC VALUES ACROSS 1,000 TRIALS

Criteria	GM_LOG	LGP	Proposed
Mean	0.663	0.753	0.787
Std	0.0715	0.0666	0.0359

To comprehensively evaluate the improvement in prediction of the proposed blind metric, the performance in terms of the SROCC and PLCC was compared with that of the following previous objective metrics: five FR IQA metrics (peak signal to noise ratio (PSNR), structural similarity index (SSIM) [11], SPQA [10], SFUW [12], and ESIM [17]) and six blind IQA metrics (QAC [20], NIQE [21], IL_NIQE [22], GM_LOG [23], local gradient patterns (LGP) [24], DIIVINE [52], and BQMS [29]). The PLCC and SROCC values for the given SIQAD are summarized in Table II. The best results in each case are highlighted in bold font. Hence, these experimental results indicate that the proposed blind metric can achieved comparative and reasonably encouraging quality predictions for distorted SCIs compared to other blind metrics. Specifically, it is well known that GM_LOG and LGP perform well for natural images. However, these metrics do not perform well when applied

to SCIs. Moreover, compared with FR metrics, the proposed blind metric achieves competitive performance with the top-performing FR metric because it considers the local and global statistical properties of the SCIs. Furthermore, Table III shows the mean and standard deviation of the SROCC values across the 1,000 trials; higher means with lower standard deviations indicate outstanding prediction performance. In summary, the proposed blind metric quantifies and predicts perceptual distortions in SCIs stably.

TABLE IV OVERALL PERFORMANCE OF ELEVEN METRICS FOR EACH TYPE OF DISTORTION (PLCC)

Distortion	FR					Blind						
	PSNR	SSIM	SPQA	SFUW	ESIM	QAC	DIIVINE	NIQE	IL_NIQE	GM_LOG	LGP	Proposed
GN	0.905	0.866	0.892	0.887	0.899	0.853	0.872	0.805	0.815	0.898	0.892	**0.903**
GB	0.860	0.901	0.906	0.923	0.923	0.559	0.853	0.480	0.540	0.886	0.856	**0.911**
MB	0.704	0.806	0.832	0.878	0.889	0.378	0.803	0.201	0.359	0.678	0.789	**0.837**
CC	0.753	0.756	0.799	0.829	0.764	0.074	0.672	0.118	0.138	0.664	**0.731**	0.660
JPEG	0.770	0.749	0.770	0.757	0.800	0.302	0.631	0.427	0.268	0.689	0.628	**0.762**
JP2K	0.789	0.775	0.825	0.815	0.789	0.189	0.636	0.185	0.418	0.571	0.619	**0.668**
LSC	0.781	0.731	0.796	0.759	0.792	0.337	0.519	0.418	0.196	0.565	**0.748**	0.683

TABLE V PERFORMANCE OF ELEVEN METRICS FOR EACH TYPE OF DISTORTION (SROCC)

Distortion	FR					Blind						
	PSNR	SSIM	SPQA	SFUW	ESIM	QAC	DIIVINE	NIQE	IL_NIQE	GM_LOG	LGP	Proposed
GN	0.879	0.869	0.886	0.869	0.876	0.842	0.862	0.824	0.816	0.878	0.868	**0.879**
GB	0.858	0.893	0.912	0.917	0.924	0.624	0.849	0.567	0.456	0.879	0.841	**0.894**
MB	0.713	0.804	0.844	0.874	0.894	0.338	0.799	0.375	0.446	0.665	0.779	**0.832**
CC	0.683	0.641	0.638	0.722	0.611	0.075	0.498	0.069	0.044	0.551	**0.615**	0.487
JPEG	0.757	0.758	0.771	0.750	0.799	0.145	0.626	0.447	0.287	0.691	0.611	**0.744**
JP2K	0.775	0.760	0.844	0.812	0.783	0.134	0.628	0.247	0.381	0.576	0.602	**0.645**
LSC	0.793	0.737	0.859	0.754	0.796	0.187	0.514	0.344	0.168	0.561	**0.723**	0.666

TABLE VI PERFORMANCE COMPARISON ON THE CSIQ AND CCT DATASETS

Dataset	Criteria	DIIVINE	CORNIA	NIQE	IL_NIQE	LGP	Proposed
CSIQ	PLCC	0.898	0.918	0.716	0.917	0.921	**0.922**
	SROCC	0.876	0.893	0.627	0.889	0.893	**0.906**
CCT	PLCC	0.555	0.581	0.548	0.501	0.659	**0.674**
	SROCC	0.508	0.487	0.234	0.166	0.621	**0.641**

TABLE VII PERFORMANCE OF EACH FEATURE IN THE PROPOSED BLIND METRIC FOR ALL DISTORTIONS ON THE SIQAD DATASET

Criteria	Metric A	Metric B	Proposed
PLCC	0.821	0.809	0.828
SROCC	0.783	0.769	0.788

TABLE VIII PERRFORMANCE OF EACH FEATURE IN THE PROPOSED BLIND METRIC FOR EACH DISTORTION TYPES (PLCC) ON THE SIQAD DATASET

Distortion	Metric A	Metric B	Proposed
GN	0.892	**0.914**	0.903
GB	0.902	0.887	**0.911**
MB	0.826	**0.875**	0.837
CC	**0.671**	0.335	0.660
JPEG	0.740	**0.819**	0.762
JP2K	0.678	0.666	**0.668**
LSC	0.682	**0.784**	0.683

TABLE IX PERFORMANCE OF EACH FEATURE IN THE PROPOSED METRIC FOR EACH DISTORTION TYPES (SROCC) ON THE SIQAD DATASET

Distortion	Metric A	Metric B	Proposed
GN	0.866	**0.900**	0.879
GB	0.885	0.869	**0.894**
MB	0.824	**0.870**	0.832
CC	**0.492**	0.312	0.487
JPEG	0.720	**0.809**	0.744
JP2K	0.661	0.642	**0.645**
LSC	0.660	**0.790**	0.666

TABLE X RESULTS OF F-TEST COMPARING SROCC VALUES OF VARIOUS METRICS ON THE SIQAD DATASET

	IL_NIQE	BQMS	GM_LOG	LGP	Proposed
IL_NIQE	0	-1	-1	-1	-1
BQMS	1	0	1	-1	-1
GM_LOG	1	-1	0	-1	-1
LGP	1	1	1	0	-1
Proposed	1	1	1	1	0

To more comprehensively evaluate an IQA metric's abil- ity to predict the perceptual quality of SCIs caused by different types of distortions, we examined the prediction performance of our blind metric against the competing met- rics (PSNR, SSIM, SPQA, SFUW, ESIM, QAC, NIQE,

IL_NIQE, GM_LOG, DIIVINE, and LGP) on specific types of distortions. The PLCC and SROCC results are presented (Tables IV and V, respectively). From the tables, it can be see that, compared with existing IQA metrics, the proposed blind metric can better handle GN, GB, MB, JPEG, and JP2k distortions. Possibly because of the quality features, which cannot efficiently and effectively represent the visual quality distortion of contrast or shape changes, the proposed blind metric is not good at handling CC and LSC distortions. However, although some blind metrics are good at handling certain individual distortions (CC and LSC), the proposed blind metric is clearly competitive with the blind metric that obtain very promising performance on CC and LSC distortions. In general, the proposed blind metric performs better than, or has comparable prediction performance to, the classical FR metrics.

In this section, we tested our proposed blind metric on the CSIQ and CCT datasets to demonstrate that the proposed model has a strong ability to handle cross-content-type images. Table VI summarizes the test results on these two datasets, which lead to several useful findings. First, on the CSIQ dataset, which contains NISs only, the results indicate that the learned local and global quality features are efficient for NSIs. Second, on the CCT dataset, which contains cross-content-type images (e.g., NSIs, SCIs, and others), it is interesting to note that the conventional handcrafted features cannot represent cross-content-type images, which contain not only natural scenes but also text, tables, icons, and graphics. However, the learning-based features are able to catch the specific information with various degrees of distortion for text, tables, icons, and graphics. It is efficient for the analysis of NSIs and SCIs

under these circumstance not to include any prior knowledge about the image type. In summary, the results show that the proposed blind metric has a generalization ability that can be applied to cross-content-type images, including SCIs, NSIs, and other types. To better understanding the individual contributions of the learned local and global features in the proposed blind metric, we designed two test metrics, A and B, in which only the learned local or global features, respectively, were used to estimate the visual quality. Tables VII, VIII, and IX show the performance of metrics A and B, along with that of the proposed blind metric on the SIQAD dataset. For all distortions, it can be see that the performance can be promoted by correctly and properly integrating the learned local and global features. For each type of distortion, we were surprised to see that metric B obtained promising performances for the GN, MB JPEG, and LSC distortions, while metric A performed very poorly. Therefore, the learned local and global features are complementary, reflecting different attributes of the HVS for quality prediction.

To evaluate the statistical significance of the proposed blind metric's advantages over other blind IQA metrics, F-tests were conducted at the 5% significance level. The validation results on the SIQAD dataset are summarized in Table X. Symbol "1" denotes that the row metric is significantly better than the column metric, "–1" means that the row metric is significantly worse than the column metric, and "0" indicates that the row metric is significantly indistinguishable with the column metric. It is clear that outstanding performance was achieved by our proposed metric, whose results are "1" for all comparisons. The computational complexity is another important consideration when evaluating the computational performance of our proposed blind metric. The testing environment consisted of a 2.70 GHz Intel Core i5 CPU processor, 8 GB RAM, and the MATLAB R2010 platform. The experimental results on the SIQAD dataset are listed in Table XI, where we list the average computation time per SCI image. As Table XI clearly demonstrates, our proposed blind metric has medium computational complexity.

1) The current computational complexity of the proposed metric is not very low. The most time-consuming part of the metric is the extraction of the log-Gabor base features. We could replace these features with less computationally expensive features (e.g., gradient based features) in future work. Furthermore, because the local and global feature meth- ods can be executed simultaneously, and parallel computing could be used to improve the speed of the proposed metric.

2) Blind quality evaluation for NSIs and SCIs is still in a preliminary stage. Hence, we have sufficient space for improvement. In this work, only local and global visual characteristics for NSIs and SCIs were investigated, while the visual characteristics of NSIs and SCIs such as visual saliency and the visual statistical model are different from those of natural images. Further exploration of human visual physiology and psychology may bring more inspiration for establishing a model with more specific characteristics.

3) Table VII shows that the adding global features only improves the performance a little for all distortions; however, Tables VIII and IX show that the main contribution of the performances on GN, MB JPEG, and LSC distortions may come from the global feature. Metric B is marginally inferior to metric A for all distortions, perhaps because the global feature cannot efficiently and effectively represent the visual quality distortion for contrast change. Further, how to construct a more effective global feature for CC distortion should be considered.

V. FUTURE SCOPE AND CONCLUSION

In this paper, we have put forward an effective blind quality predictor for distorted images by incorporating the learning of both local and global quality features. The novelty of our research resides in combining the complementary behaviors of the locality-constrained linear coding (LLC)-based local qual- ity features and collaborative representation (CR)-based global quality features to achieve a fused representation for images.

To extract the learned local features, we use a strategy of LLC with max pooling. Further, the learned global features of the distorted images can be obtained using the CR method. Compared with competing IQA metrics, experimental results show that the proposed blind metric can obtain significantly higher statistical consistency with evaluations from human subjects, confirming that the devised metric is a very robust blind IQA metric for NSIs and SCIs. Various aspects of the present research merit further inves- tigation and will, therefore, be considered in future work. In the feature extraction stage, we will focus on mining the special characteristics of the NSIs and SCIs more deeply. In the perceptual-quality prediction stage, we intend to design blind IQA metrics for NSIs and SCIs based on deep learning approaches.

REFERENCES

[1] S. Wang, L. Ma, Y. Fang, W. Lin, S. Ma, and W. Gao, "Just noticeable difference estimation for screen content images," *IEEE Trans. Image Process.*, vol. 25, no. 8, pp. 3838–3851, Aug. 2016.

[2] 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.

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

[4] Z. Ma, W. Wang, M. Xu, and H. Yu, "Advanced screen content coding using color table and index map," *IEEE Trans. Image Process.*, vol. 23, no. 10, pp. 4399–4412, Oct. 2014.

[5] W. Zhu, W. Ding, J. Xu, Y. Shi, and B. Yin, "Hash-based block matching for screen content coding," *IEEE Trans. Multimedia*, vol. 17, no. 7, pp. 935–944, Jul. 2015.

[6] S. Wang, X. Zhang, X. Liu, J. Zhang, S. Ma, and W. Gao, "Utility- driven adaptive preprocessing for screen content video compression," *IEEE Trans. Multimedia*, vol. 19, no. 3, pp. 660–667, Mar. 2017.

[7] C.-C. Chen and W.-H. Peng, "Intra line copy for HEVC screen content intra-picture prediction," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 27, no. 7, pp. 1568–1579, Jul. 2017.

[8] S. Wang *et al.*, "Subjective and objective quality assessment of com- pressed screen content images," *IEEE J. Emerg. Sel. Topics Circuits Syst.*, vol. 6, no. 4, pp. 532–543, Dec. 2016.

[9] H. Yang, Y. Fang, Y. Yuan, and W. Lin, "Subjective quality evaluation of compressed digital compound images," *J. Vis. Commun. Image Represent.*, vol. 26, pp. 105–114, Jan. 2015.

[10] 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.

[11] 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.

[12] 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.

[13] K. Gu *et al.*, "Saliency-guided quality assessment of screen content images," *IEEE Trans. Multimedia*, vol. 18, no. 6, pp. 1098–1110, Jun. 2016.

[14] 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.

[15] 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.

[16] X. Guo, L. Huang, K. Gu, L. Li, Z. Zhou, and L. Tang, "Naturalization of screen content images for enhanced quality evaluation," *IEICE Trans. Inf. Syst.*, vol. 100, no. 3, pp. 574–577, 2017.

[17] 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.

[18] 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.*, vol. 28, no. 1, pp. 1–14, Jan. 2018.