No Reference Quality Assessment for Multiply-Distorted Images Based on an Improved Bag-of-Words Model

Yanan Lu, Fengying Xie, Tongliang Liu, Zhiguo Jiang, and Dacheng Tao, *Fellow, IEEE*

***Abstract*—Multiple distortion assessment is a big challenge in image quality assessment (IQA). In this letter, a no reference IQA model for multiply-distorted images is proposed. The features, which are sensitive to each distortion type even in the presence of other distortions, are first selected from three kinds of NSS features. An improved Bag-of-Words (BoW) model is then applied to encode the selected features. Lastly, a simple yet effective linear combination is used to map the image features to the quality score. The combination weights are obtained through lasso regression. A series of experiments show that the feature selection strategy and the improved BoW model are effective in improving the accuracy of quality prediction for multiple distortion IQA. Compared with other algorithms, the proposed method delivers the best result for multiple distortion IQA.**

***Index Terms*—Feature encoding, feature selection, image quality assessment, multiple distortions, no reference.**

I. Introduction

No REFERENCE image quality assessment (IQA) is of crucial importance in many image processing and analysis applications for its ability to predict image quality without access to a reference image. Over the past decade, numerous no reference IQA algorithms for general purpose or specific distortion types have been proposed, among which the general purpose no reference IQA, which predict the visual quality of images without access to reference images or prior knowledge of the distortion types, are the most challenging. Researchers have developed many state-of-the-art algorithms for general purpose no reference IQA. Being required to reflect the visual quality, feature is one of the major issues in developing a robust IQA algorithm. Moorthy *et al.* [1] and Saad *et al.* [2] extracted features in image transformation domains. Mittal *et al.* [3] used features in the spatial domain. He *et al.* [4] introduced sparse representation to IQA, and Kang *et al.* [5] used convolutional neural networks to learn features directly from the raw image pixels instead of using the handcrafted features to assess image quality. These methods can predict image quality in high correlation with subjective scores.

Nevertheless, an important problem still exists in IQA research. The distorted images in the majority of publicly available IQA datasets such as LIVE II [6], IVC [7], CSIQ [8], MICT [9], and TID [10] suffer from a single distortion type. That is to say, most of the existing IQA algorithms are tested on images with a single type of distortion. However, the images available to consumers usually reach them after going through several stages including acquisition, compression, transmission and reception. In this pipeline, they may suffer multiple distortions. The Laboratory for Image & Video Engineering (LIVE) recently built a multiply-distorted image quality dataset LIVE MD [11], for IQA research. This new multiple dataset consists of two parts: blur followed by JPEG compression and blur followed by noise. Experiments in [11] indicate that even some state-of-the-art IQA algorithms show poorer performance on this new dataset compared to their performance on single distortion datasets. In [12], Chandler pointed out that multiply-distorted images are a big challenge for IQA because an IQA algorithm must not only consider the joint effects of these distortions on the image, but must also consider the effects of these distortions on each other.

Compared with IQA images of single distortion type, IQA for multiply-distorted images has thus far received less attention. For example, in [13], Li *et al.* proposed a method named SHANIA using statistical features in the Shearlet domain. In [14], Qian *et al.* proposed to use the multi-scale representation of structure for IQA. In [15], Gao *et al.* proposed a model of learning to rank for IQA. In [16] and [17], Gu *et al.* respectively proposed a five-step and a six-step blind metric (FISBLIM and SISBLIM). The criteria show that the performance of IQA algorithms for multiple distortions ([14]–[17]) is much poorer than that of IQA algorithms for single distortion types ([1]–[5]). A sizable gap exists for multiple distortion IQA research, as it is necessary to learn how to assess the quality of multiply-distorted images precisely and efficiently.

Feature description for image quality assessment is important. In IQA methods, all the assessment metrics are designed according to features that are related to image quality. The more sensitive the feature is to changes in image quality, the more effective the assessment metric will be. In this letter, a no reference IQA method for multiply-distorted images is proposed based on an improved *Bag-of-Words (BoW)* model using *Selected Features*. The algorithm is called BoWSF for short. In the next section, the proposed method is described in

Manuscript received March 06, 2015; revised May 17, 2015; accepted May 20, 2015. Date of publication May 22, 2015; date of current version June 03, 2015. This work was supported by the National Natural Science Foundation of China under Grants 61471016 and 61371134, and by the Australian Research Council Projects DP-140102164, FT-130101457, and LP-140100569. The associate editor coordinating the review of thi smanuscript and apprroving it for publication was Prof. Tolga Tasdizen.

Y. Lu, F. Xie, and Z. Jiang are with the Image Processing Center, School of Astronautics, Beihang University, Beijing, China (e-mail: yn_lu@buaa.edu.cn; xfy_73@buaa.edu.cn; jiangzg@buaa.edu.cn).

T. Liu and D. Tao are with the Center for Quantum Computation & Intelligent System and the Faculty of Engineering and Information Technology, University of Technology Sydney, Sydney, Australia (e-mail: tliang.liu@gmail.com; dacheng.tao@uts.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/LSP.2015.2436908

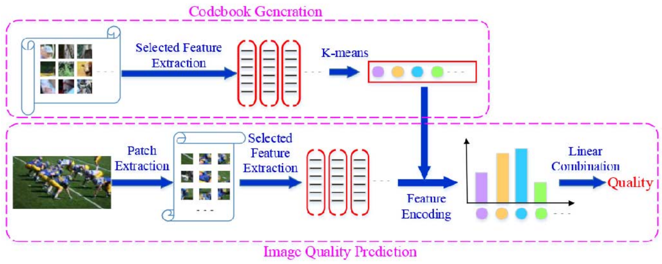

Fig. 1. Flowchart of the proposed algorithm BoWSF.

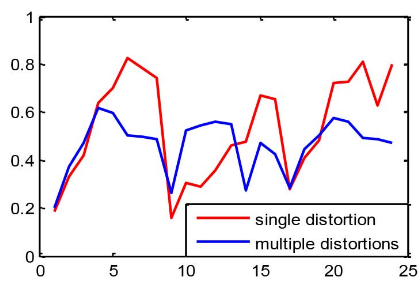

Fig. 2. Correlation coefficient between features and distortion levels.

detail. Experiments and analysis are presented in Section III, and Section IV concludes the letter.

II. Method

Finding IQA models which can handle multiple distortions has proven difficult [12]. In this letter, a quality assessment framework, BoWSF, is proposed for those images impaired by multiple distortions. In the BoWSF model, the correlation between the extracted features and each level of the multiple distortions is analyzed, and the corresponding optimal features are selected based on the correlation analysis. The selected features are then encoded through an improved BoW model. Lastly, the joint effects of multiple distortions are modeled using a simple but effective pooling strategy: linear combination. Fig. 1 shows the flowchart of the proposed BoWSF. In the following subsections, the proposed method is introduced in detail.

A. Feature Correlation Analysis

The particular structures contained in natural scenes can be described through Natural Scene Statistics (NSS) features, which are sensitive to changes in the degree of distortion and are widely used in IQA algorithms [4], [18], [19]. For single distortion assessment, all of the extracted features are usually in high correlation with the distortion levels. However, when two or more distortions appear in the same image, the mutual effects of the multiple distortions lead to complex feature changes, and this high correlation may no longer exist. Taking the widely-known NSS features, which are described in SRNSS [4], as examples: a feature vector of 24 dimensions was formed by extracting the mean, variance, and entropy of wavelet coefficients in each sub-band over 4 scales. The correlation coefficient between the distortion level and the jth feature can be calculated as:

$$r_j = \frac{\sum_i (l_i - \bar{l})(f_{ij} - \bar{f}_j)}{\sqrt{\sum_i (l_i - \bar{l})^2 \sum_i (f_{ij} - \bar{f}_j)^2}} \quad j = 1, 2, \cdots 24 \qquad (1)$$

where, l_i is the distortion level of the ith image, and $\bar{l}$ is the average level for all the images. f_{ij} represents the jth feature extracted from the ith image, and $\bar{f}_j$ is the mean of the jth feature for all the images. Fig. 2 shows the correlation coefficients between features and the distortion levels of images, where the red line represents the 982 single distorted images in the LIVE II dataset [6] and the blue line represents the 480 multiply-distorted images in the LIVE MD dataset [11]. It can be seen that most of the features have a lower correlation coefficient with the distortion levels of multiply-distorted images than those with a single distortion. It will therefore be of great value to select those features that have high correlation with the distortion levels to develop a robust IQA model for multiply-distorted images.

B. Feature Extraction and Selection

The NSS features used in SRNSS [4], BRISQUE-L [18] and BIQI [19] are representative and fast for calculation. In this letter, we extract these three features as original features for the following feature selection. The SRNSS and BIQI features are both extracted in the wavelet domain. As mentioned above, SRNSS features include the mean, variance, and entropy of wavelet coefficients in 4 sub-bands. In BIQI, the sub-band coefficients are parameterized using a generalized Gaussian distribution (GGD), and the parameters are then used as features. The BRISQUE-L feature is extracted in the spatial domain over 2 scales. Following image preprocessing, a GGD is utilized in BRISQUE-L to estimate the distribution, and the estimated distribution parameters are taken to be features. The dimensions of SRNSS, BRISQUE-L and BIQI features are 24, 36, and 18, respectively. We put the three kinds of features together to form a 78-D original feature f_{all}.

$$f_{all} = [f_{SRNSS}, f_{BRISQUE-L}, f_{BIQI}] \qquad (2)$$

The feature selection is applied to f_{all} to locate the features which are sensitive to one distortion type even in the presence of another distortion. Suppose there are x images suffering from multiple distortions (type A and type B), and each of the distortions has y levels. When the distortion level of type B is fixed, the good features for A should be in high correlation with the changes in the distortion level of type A. Similarly, for a fixed distortion level of type A, the good features for B should be sensitive to the distortion level of type B. We separate the x images into y groups, and the images in each group have the same distortion level as type B. For each group, the features for distortion type A are selected as: 1) calculate the Pearson linear correlation coefficient (LCC) and the Spearman rank-order correlation coefficient (SROCC) between the level of distortion A and every feature dimension. LCC and SROCC are both between 0 and 1–the closer the value of LCC and SROCC to 1, the higher the correlation between the two variables; 2) average the value of LCC and SROCC, and take the top 3 features with the highest 3 coefficients as the selected features. All the features chosen for each group are put together to form the final selected features f_A for distortion A. To avoid one feature appearing more than once in the selected features f_A, the duplicate features should be removed. f_A have high correlation with the distortion level of type A because of the selection strategy described above. In the same way, a group of selected features f_B for distortion B can also be obtained.

C. Feature Representation Based on BoW

As an image representation method, the BoW model has been widely used in image classification. In this model, an image is seen as 'a bag' of certain 'words'. The histogram formed by the statistical numbers of all words is a representation of this image. It is true that certain regions of an image can hide distortions better than other regions, which is the result of visual masking

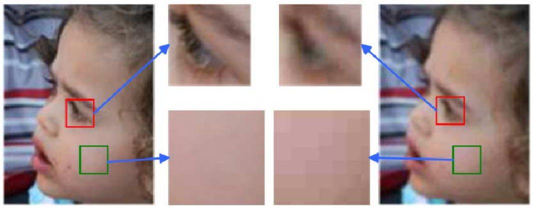

Fig. 3. Multiply-distorted image.

[20]. For example, Fig. 3 shows the original reference image and the distorted vision corrupted by blur and JPEG compression, and it can be seen that the blocking artifacts are usually seen in smooth regions such as the face and the forehead, while most of the blur appears in edge regions like the eye and hair. The visual perceptual quality of the various patches in a multiply-distorted image is different. If the image patches of different distortion types or levels are seen as 'words', then a multiply-distorted image can also be seen as 'a bag' of 'words'.

In this letter, codebooks for distortion types A and B, represented by $codebook_A$ and $codebook_B$ respectively, are generated by applying a k-means algorithm to a large number of image patches using their corresponding selected features f_A and f_B. The patches are from the supplementary multiply-distorted images described in the experiment section, and in this letter, the number of patches is 46800. Since the selected features are in high correlation with the distortion level, it is easy to cluster the image blocks of the same distortion level in one class by k-means. That is to say, the words in one codebook indicate the corresponding distortion level to some extent.

With the selected features and codebooks, an image can be represented with a bag of words. In a traditional BoW model [21], an image patch is seen as its nearest words, which is effective for solving the classification problem. However, in IQA, the degree of distortion can be continuous. Note that the distance between an image patch and the words can indicate the possibility that the patch suffers from the same level of distortion. The closer the distance, the bigger the possibility, thus we propose to use the reciprocal of distance as the weight when representing images with BoW model. Suppose an image is divided into N patches $x = \{x_1, x_2 \cdots, x_N\}$, the BoW feature of this image f_{bow} is defined as:

$$f_{bow} = [f_1, f_2 \cdots, f_M] \tag{3}$$

$$f_j = \sum_{i=1}^{N} \frac{1}{1 + d_{ij}}, \quad j = 1, 2 \cdots, M \tag{4}$$

where d_{ij} represents the distance between the ith image patch and the jth word in the codebook. M is the number of words in one codebook. The words in a codebook indicate the distortion type and distortion level, and the value of each dimension of BoW feature f_{bow} indicates how many patches in the image suffer from this distortion.

The BoW feature representation procedure is applied separately for distortion types A and B. The group of BoW features f_{bow_A} for distortion A can be obtained by using the corresponding features f_A and codebook $codebook_A$. Similarly, the BoW features f_{bow_B} for distortion B can be obtained by using f_B and $codebook_B$. The two groups of BoW features are put together to form the final representation for multiply-distorted images.

$$f_{bow\_multi} = [f_{bow\_A}, f_{bow\_B}] \tag{5}$$

D. Image Quality Prediction

Taking BoW features as the representation of the image, the dimensions of which can be explained as the number of different distortion levels, the image quality can be predicted by simply using the linear combination of each dimension of the BoW features.

$$p = \sum w f_{bow\_multi}^T \tag{6}$$

where, $f_{bow\_multi}^T$ is the transposition of BOW features f_{bow_multi}, w is the weights for linear combination. In the training stage, the features are normalized and then the weights are obtained through lasso regression [22]. As is known, the lasso is a shrinkage and selection method for linear combination, which minimizes the usual sum of squared errors with a bound on the sum of the absolute values of the coefficients. p is the human subjective score in the training stage and the objective prediction in the test stage.

III. Experiments and Analysis

To quantitatively validate the performance of the proposed BoWSF, experiments were conducted in regards to three aspects using the LIVE MD dataset: 1) the effectiveness of the feature selection strategy; 2) the effectiveness of the feature encoding; 3) the comparison with state-of-the-art algorithms. The performances of the competing IQA algorithms were compared using two evaluation criteria: LCC and SROCC.

A. Experiments Data

The LIVE MD dataset contains 480 images in total, which is generated from 15 reference images through adding two kinds of multiple distortion: blur followed by JPEG and blur followed by noise. This is insufficient for the learning-based BoWSF model, therefore we supplement the image samples from the LIVE II dataset in this letter.

LIVE II is a single distortion dataset, including blur, JPEG, noise, etc. There are 29 reference images in LIVE II, which are all different with the 15 reference images in LIVE MD. Therefore we choose the 29 reference images of LIVE II as source images. Then the same multiple distortions as LIVE MD are added to the source images according to the instructions in [11]. As with LIVE MD, the supplementary dataset has two parts: part1 blur followed by JPEG compression and part2 blur followed by noise. Four levels of blur, JPEG compression and noise were considered. The distorted parameters for levels 1, 2 and 3 can be found in [11], and level 0 indicates no distortion. Finally, the supplementary dataset contains 928 multiply-distorted images in total (464 distorted images for each part).

The distortion levels of the images in the supplementary dataset are known but their quality ground truth is unknown. In this letter, the supplementary dataset is used for feature selection and codebook generation. The images in the LIVE MD dataset are used for training the combination weight described in Section II-D and to test the effectiveness of the algorithm. Experiments are conducted separately on the two kinds of distortion combination (part1 blur with JPEG and part2 blur with noise).

B. Parameter Selection

There are two key parameters for the BoW model: the patch size and the number of words in the codebook. In the classification domain, the patch size of the BoW model is small (usually from 5×5 to 15×15) [23]. However, for IQA, the NSS

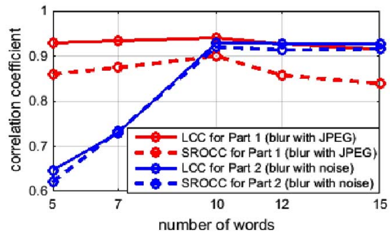

Fig. 4. Selection of number of words.

TABLE I
LCC AND SROCC OF ALL THE FEATURES AND SELECTED FEATURES

	Part 1 (Blur & JPEG)		Part 2 (Blur & Noise)	
	LCC	SROCC	LCC	SROCC
AF + SVR	0.0627	0.1301	0.2216	0.4166
SF + SVR	0.6376	0.6631	0.5263	0.6499
AF + LC	0.6404	0.6772	0.7002	0.7363
SF + LC	**0.7641**	**0.7261**	**0.8700**	**0.8604**

AF: All Features,SF: Selected Features,LC: Linear Combination.

TABLE II
LCC AND SROCC FOR FEATURE ENCODING

	Part 1 (Blur & JPEG)		Part 2 (Blur & Noise)	
	LCC	SROCC	LCC	SROCC
SF + LC	0.7641	0.7261	0.8700	0.8604
SF + Traditional BoW + LC	0.9396	0.8682	0.8701	0.8380
SF + Improved BoW + LC	**0.9508**	**0.8949**	**0.9284**	**0.9210**

SF: Selected Features,LC: Linear Combination.

features are usually extracted from the whole image [4]–[19]. Taking these two aspects into consideration, the patch size is 100×100 in this letter.

The number of words in $codebook_A$ and $codebook_B$ is the same, denote as n_c. We changed n_c from 5 to 15, and for different n_c, the LCC and SROCC are shown in Fig. 4. All the LCCs and SROCCs reported in this letter including Fig. 4, as well as Section III-C and III-D are the median values after applying the training and testing stage 1000 times. It can be seen that the four correlation coefficients reach a maximum when $n_c = 10$. Therefore $n_c = 10$ is the optimal parameter for the proposed algorithm in this letter.

C. Effectiveness of the Feature Selection and Encoding

A feature selection strategy is proposed in this letter for multiple distortion IQA. We choose linear combination and the widely used support vector regressor (SVR) [24] to map the image features to the quality score. The images in the LIVE MD dataset are generated from 15 reference images. In this experiment,10 reference images are selected randomly,and their corresponding distorted images are used for training. The remaining reference images and their corresponding distorted versions are used for testing. Table I gives the LCC and SROCC of all the features and the selected features. It can be seen that the quality prediction performance is effectively improved through feature selection. The results in Table I also indicate that simple linear combination is more effective than SVR for multiple distortions.

An improved BoW model is applied for multiple distortion IQA in the proposed algorithm. The quality prediction without BoW feature encoding and the quality prediction using the traditional BoW model are used as comparisons. The results of LCC and SROCC are shown in Table II. It can be seen that the improved BoW can significantly improve prediction accuracy for both part1 (blur with JPEG) and part2 (blur with noise).

TABLE III
STATISTICS RESULT OF IQA ALGORITHMS

	Part 1 (Blur & JPEG)			Part 2 (Blur & Noise)		
	LCC	SROCC	KRCC	LCC	SROCC	KRCC
PSNR	0.6890	0.6514	0.4618	0.7218	0.6378	0.4689
SSIM	0.7615	0.7443	0.5430	0.7473	0.7022	0.5251
BIQI	0.6398	0.6542	0.4609	0.6634	0.4902	0.3499
QAC	0.3758	0.3959	0.2761	0.4054	0.4707	0.3532
BRISQUE	0.8454	0.7902	0.5953	0.2868	0.2992	0.2107
BRISQUE-2	0.9462	**0.9214**	-	0.9226	0.8934	-
SHANIA	0.7629	0.8014	-	0.7073	0.7528	-
FISBLIM	0.8920	0.8583	-	0.8725	0.8548	-
SISBLIM	0.8651	0.8749	0.6926	0.8795	0.8802	0.6976
GAO	0.9265	0.8939	0.7208	0.9015	0.8983	0.7226
Proposed-1	**0.9508**	**0.8949**	**0.7333**	**0.9284**	**0.9210**	**0.7527**
Proposed-2	**0.9483**	**0.9075**	**0.7365**	**0.9371**	**0.9043**	**0.7418**

D. Comparison with other IQA Algorithms

In this experiment,the performance of the proposed BoWSF is compared with the performance of other IQA algorithms,including full reference IQA methods PSNR,SSIM [25],state-of-the-art no reference IQA algorithms BIQI [19],QAC [26],BRISQUE [3],BRISQUE-2 [11] and three algorithms developed for multiple distortions SHANIA [13],FISBLIM [16],SISBLIM [17] and GAO [15]. For better comparison,we add another performance criterion,Kendall's Rank Correlation Coefficient (KRCC). As with LCC and SROCC,a larger value of KRCC indicates a better performance. Table III shows the experiment results. To avoid mistakes during realization, the results of SHQNIA and FISBLIM come from the original paper,while the results of the other methods, except for PSNR,are obtained through the demos provided by the authors. Table III is the statistics result. Proposed-1 and Proposed-2 are our proposed models. Proposed-1 used different models with their corresponding codebooks for Part1 and Part2. Proposed-2 is a new model trained with all the codebooks,which can estimate the image quality in LIVE MD without knowing the distortion combination information. It can be seen that our proposed algorithm delivers the best quality prediction accuracy for each part of the LIVE MD dataset except for SROCC of Part1.

IV. CONCLUSION

Multiple distortions have complex influences on image quality and are a big challenge for IQA. A no reference IQA method,BoWSF for multiply-distorted images,is proposed in this letter. The features sensitive to each of the multiple distortions are selected from NSS features. An improved BoW model is then used to encode the selected image features. The encoding process is conducted separately for each type of multiple distortion in the images,and the representation of the multiply-distorted image is then formed. Lastly,a simple linear combination is applied to map the image features to the quality score, and the combination weights are obtained through lasso regression. Not only can the proposed method be used in the two kinds of multiple distortion types in LIVE MD, but it can also be used in other multiple distortion combinations by using the corresponding selected features and codebooks. Comprehensive experiment results show that the proposed feature selection and encoding process are beneficial for improving the accuracy of quality prediction for multiply-distorted images,and at the same time,the quality assessment results are in high agreement with the subjective score.

REFERENCES

[1] A. K. Moorthy and A. C. Bovik, "Blind image quality assessment: From natural scene statistics to perceptual quality[J]," *IEEE Trans. Image Process.*, vol. 20, no. 12, pp. 3350–3364, 2011.

[2] M. A. Saad, A. C. Bovik, and C. Charrier, "Blind image quality assessment: A natural scene statistics approach in the DCT domain[J]," *IEEE Trans. Image Process.*, vol. 21, no. 8, pp. 3339–3352, 2012.

[3] A. Mittal, A. K. Moorthy, and A. C. Bovik, "No-reference image quality assessment in the spatial domain[J]," *IEEE Trans. Image Process.*, vol. 21, no. 12, pp. 4695–4708, 2012.

[4] L. He, D. Tao, and X. Li *et al.*, "Sparse representation for blind image quality assessment[C]," in *IEEE Conf. Computer Vision and Pattern Recognition (CVPR), 2012*, 2012, pp. 1146–1153, IEEE.

[5] L. Kang, P. Ye, and Y. Li *et al.*, "Convolutional neural networks for no-reference image quality assessment[C]," in *IEEE Conf. Computer Vision and Pattern Recognition (CVPR), 2014*, 2014, pp. 1733–1740, IEEE.

[6] H. R. Sheikh, Z. Wang, L. Cormack, and A. C. Bovik, "Live image quality assessment data set release 2," [Online]. Available: http://www.live.ece.utexas.edu/ accessed 28, Mar. 2013

[7] L. C. Patrick, "Florent Autrusseau Subjective quality assessment IRCCyN/IVC database," [Online]. Available: http://www.irccyn.ec-nantes.fr/ivcdb/

[8] 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. 011006–011006-21, 2010.

[9] Y. Horita, K. Shibata, Y. Kawayoke, and Z. M. P. Sazzad, *MICT image quality evaluation dabase 2000*, Mar. 2013 [Online]. Available: http://www.mict.eng.utoyama.ac.jp/mictdb.html

[10] N. Ponomarenko, V. Lukin, and A. Zelensky *et al.*, "TID2008-A database for evaluation of full-reference visual quality assessment metrics," *Adv. Mod. Radioelectron.*, vol. 10, no. 4, pp. 30–45, 2009.

[11] D. Jayaraman, A. Mittal, and A. K. Moorthy *et al.*, "Objective quality assessment of multiply distorted images[C]," in *Conf. Rec. Forty Sixth Asilomar Conf. Signals, Systems and Computers (ASILOMAR), 2012*, 2012, pp. 1693–1697, IEEE.

[12] D. M. Chandler, "Seven challenges in image quality assessment: Past,present,and future research[J]," *Signal Process.*, 2013.

[13] Y. Li, L. M. Po, and X. Xu *et al.*, "No-reference image quality assessment using statistical characterization in the shearlet domain[J]," *Signal Process.: Image Commun.*, vol. 29, no. 7, pp. 748–759, 2014.

[14] J. Qian, D. Wu, and L. Li *et al.*, "Image quality assessment based on multi-scale representation of structure[J]," *Digital Signal Process.*, vol. 3, no. 3, pp. 125–133, 2014.

[15] F. Gao, D. Tao, and X. Gao *et al.*, "Learning to rank for blind image quality assessment [J]," *arXiv preprint arXiv:1309.0213*, 2013.

[16] K. Gu, G. Zhai, and M. Liu *et al.*, "FISBLIM: A FIve-Step BLInd Metric for quality assessment of multiply distorted images[C]," in *IEEE Workshop on Signal Processing Systems (SiPS), 2013*, vol. 201, no. 3, pp. 241–246, IEEE.

[17] K. Gu, G. Zhai, and X. Yang *et al.*, "Hybrid no-reference quality metric for singly and multiply distorted images [J]," *IEEE Trans. Broadcast.*, vol. 60, no. 3, pp. 555–567, 2014.

[18] A. Mittal, A. K. Moorthy, and A. C. Bovik, "Making image quality assessment robust[C]," in *Conf. Rec. Forty Sixth Asilomar Conf. Signals,Systems and Computers (ASILOMAR), 2012*, 2012, pp. 1718–1722, IEEE.

[19] A. K. Moorthy and A. C. Bovik, "A two-step framework for constructing blind image quality indices[J]," *IEEE Signal Process. Lett.*, vol. 17, no. 5, pp. 513–516, 2010, IEEE.

[20] G. E. Legge and J. M. Foley, "Contrast masking in human vision[J]," *J. OSA*, vol. 70, no. 12, pp. 1458–1471, 1980.

[21] L. Fei-Fei and P. Perona, "A Bayesian hierarchical model for learning natural scene categories[C]," in *IEEE Computer Society Conf. Computer Vision and Pattern Recognition, 2005 CVPR 2005*, 2005, no. 2, pp. 524–531, IEEE.

[22] R. Tibshirani, "Regression shrinkage and selection via the lasso," *J. Roy. Statist. Soc. B*, vol. 58, no. 1, pp. 267–288, 1996.

[23] L. Liu, P. Fieguth, and D. Clausi *et al.*, "Sorted random projections for robust rotation-invariant texture classification[J]," *Patt. Recognit.*, vol. 45, no. 6, pp. 2405–2418, 2012.

[24] A. J. Smola and B. Schölkopf, "A tutorial on support vector regression," *Statist. Comput.*, vol. 14, no. 3, pp. 199–222, 2004.

[25] Z. Wang, A. C. Bovik, and H. R. Sheikh *et al.*, "Image quality assessment: From error visibility to structural similarity[J]," *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600–612, 2004.

[26] W. Xue, L. Zhang, and X. Mou, "Learning without human scores for blind image quality assessment[C]," in *IEEE Conf. Computer Vision and Pattern Recognition (CVPR), 2013*, 2013, pp. 995–1002, IEEE.