

No-Reference Quality Assessment for Image Sharpness and Noise

Lijuan Tang*, Xionguo Min[†], Vinit Jakheti[‡], Ke Gu[‡], Xinfeng Zhang[‡], and Shuai Yang[§]

*School of Information and Electronic Engineering, Jiangsu Vocational College of Business, China

[†]Institute of Image Communication and Network Engineering, Shanghai Jiao Tong University, China

[‡]School of Computer Science and Engineering, Nanyang Technological University, Singapore

[§]Institute of Computer Science and Technology, Peking University, China

Abstract—To blindly evaluate the visual quality of image is of great importance in many image processing and computer vision applications. In this paper, we develop a novel training-free no-reference (NR) quality metric (QM) based on a unified brain theory, namely, free energy principle. The free energy principle tells that there always exists a difference between an input true visual signal and its processed one by human brain. The difference encompasses the “surprising” information between the real and processed signals. This difference has been found to be highly related to visual quality and attention. More specifically, given a distorted image signal, we first compute the aforesaid difference to approximate its visual quality and saliency via a semi-parametric method that is constructed by combining bilateral filter and auto-regression model. Afterwards, the computed visual saliency and a new natural scene statistic (NSS) model are used for modification to infer the final visual quality score. Extensive experiments are conducted on popular natural scene image databases and a recently released screen content image database for performance comparison. Results have proved the effectiveness of the proposed blind quality measure compared with classical and state-of-the-art full- and no-reference QMs.

I. INTRODUCTION

With the widespread utility of digital images in computer vision and multimedia signal processing communities, visual quality has becoming increasingly important. However, images are subject to a wide variety of distortions and artifacts during compression, enhancement [1], [2], tone mapping [3], etc, and thus it is desirable to find a reliable technique for controlling and improving the image quality [4]. Image quality assessment (IQA), due to its consistency with the human visual system (HVS), can be properly used to solve this problem.

From the viewpoint of target, IQA models are basically separately into two categories, i.e. subjective assessment and objective assessment. The former one is decisive since human viewers are the ultimate users, but it is usually laborious, time-consuming, expensive. So it is not suitable for real-time applications. Hence, objective IQA has aroused more and more attention from researchers towards predicting subjective ratings with explicitly designed mathematical models. According to the accessibility of the reference image, objective IQA algorithms can be further classified into full-reference (FR), reduced-reference (RR), and no-reference (NR) methods. FR

IQA models compare a distorted image with its associated reference one [5], [6], [7], [8]. This kind of algorithms work under the condition that reference image can be completely obtained. Instead of using the whole reference image, limited effective features are used to quantify the distortion and evaluate the quality. This type of methods is called RR IQA [9], [10].

In most cases, the reference image (complete or partial) may not be available, so FR and RR IQA metrics are limited by the dependence of the reference image in practical applications. With this view, blind/NR IQA metrics without the help of the original references are highly expected. During the past decade, a large number of blind quality measures have been proposed. One type of NR models operates under the assumption that the degradation type is known beforehand. Typical distortion-specific quality measures are devoted to blockiness [11], [12], sharpness/blurriness [13], [14], [15], [16], [17], [18], noise [19], contrast adjustment [20], etc. The other type of general-purpose blind IQA metrics concentrate on evaluating various types of distortions simultaneously.

With the development of Internet, a huge amount of digital photographs are shared and transmitted everyday. Faithful IQA metrics play an important role in processing of digital photographs. It can be assumed that the most commonly occurring distortions on digital photographs are due to blur and presence of noise. Image noise may appear in the digital photographs due to several reasons. Such as, one reason would be a high International Organization for Standardization (ISO) setting on your camera. Graininess will also be more apparent when camera captures more light to illuminate the scene. Another condition is caused by thermal noise. Image can suffer from the image sensor noise. Long exposures increase the risk of being contaminated by noise too, since the sensor is kept open to gather more image data along with electrical noise. The causes of image blurring are multifold, such as photographing with long exposures, camera out-of-focus, target motion, image compression and so on.

Many training-based quality metrics can evaluate image blurriness and noise, but their regression modules are based on machine learning tools which are prone to the problem of overfitting. As for training-free metrics devoted to image

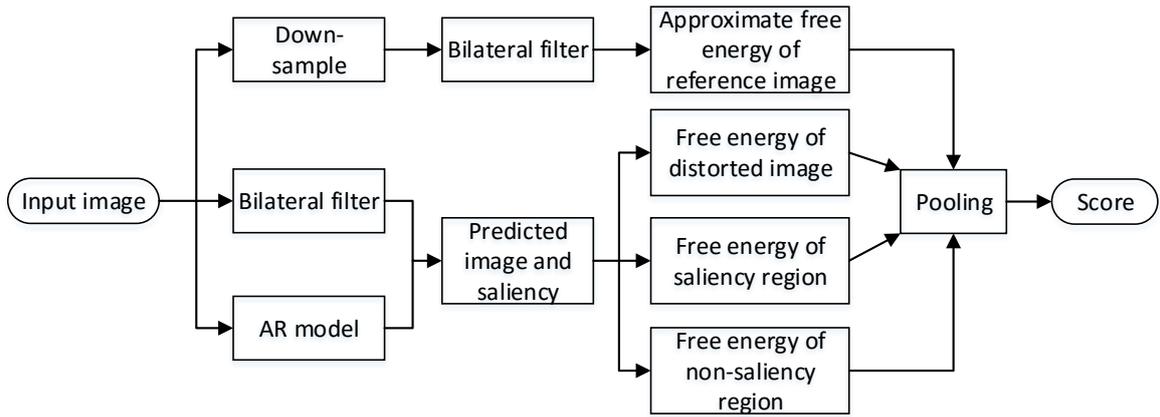


Fig. 1. A framework of the proposed BQISN measure.

blurriness and noise, they are unable to simultaneously handle both these two types. Evaluating the quality of both blurry and noisy images is therefore a meaningful and challenging research topic in this area. To resolve this issue, we propose a new effective blind quality index for image sharpness and noise (BQISN). Our model is based on free energy based brain theory, which indicates that there constantly exists a difference between an input image and its processed one by human brain. This difference contains the “surprising” information between the input and processed signals, and thus it is reasonable to be highly connected to visual quality and attention. In particular, given a distorted image signal, we first compute the above-mentioned difference to estimate its visual quality and saliency via a semi-parametric method that is constructed by combining bilateral filter (BF) and auto-regression (AR) model. Thereafter, the computed visual saliency and a new natural scene statistic (NSS) model are used for modification to yield the final visual quality score.

Compared with the previous works, there are two main contributions made in this paper: 1) to our knowledge, this work is the first one blindly evaluating the quality of blurry and noisy images simultaneously; 2) in contrast to popular FR IQA methods, distortion-specific blind techniques and state-of-the-art NR IQA metrics, the substantial high performance of the proposed method is evaluated on three popular natural scene image databases and even the recently released screen content image database.

The rest of this paper is arranged as follows. Section II first describes the proposed BQISN in detail. In Section III, comparative studies of our BQISN with classical and recently proposed competitors are conducted on four popular image databases [21], [22], [23], [7], confirming the effectiveness of our IQA method. The concluding remarks are given in section IV.

II. THE PROPOSED BLIND QUALITY MEASURE

High quality images are increasingly required by Human viewers at present time. Currently, there are many training-based quality measures which can evaluate image blurriness

and noise, but their regression modules are based on machine learning tools which are prone to the problem of overfitting. As for training-free metrics devoted to image blurriness and noise, they are unable to simultaneously handle both two types.

As a result, we devote to the training-free quality measure BQISN which can evaluate image noise and blurriness simultaneously via a three-step framework illustrated in Fig. 1. The first step is to approximate the reference image’s free energy entropy via the distorted image. The second step utilizes free energy principle to estimate the distorted image’s saliency and free energy entropy. At last, free energy entropy of entire or part of the distorted image and the approximate reference image’s free energy are fused to derive the overall quality score. Details are discussed below.

A. Approximate free energy of the reference image

First, we compute the approximate free energy entropy F_a of the reference image. We adopt the bilateral filtering [26] to preserve the edge features in this paper. It will be introduced in the following part, we also use AR model to simulate the internal generative model in our proposed BQISN. Although the AR model is effective and simple, it is sometimes unreliable when encountered image edges due to the violation of “geometric duality” principle. Hence we should find strategies to complement AR model for preserving edge information. Currently, though there are many edge-preserving filters [36], most of them are iterative and operates in spatial domain. So it may cause issues of instability and fails at edges. The bilateral filtering considers geometric closeness and photometric similarity together. It is simple, non-iterative and effective for preserving image edge information. The bilateral filtering model is defined as

$$\mathbf{y}(x) = \mathbf{k}^{-1}(x) \int \int \mathbf{f}(\delta) \mathbf{g}(\delta, x) \mathbf{p}(\mathbf{f}(\delta), \mathbf{f}(x)) d\delta \quad (1)$$

where

$$\mathbf{k}(x) = \int \int \mathbf{g}(\delta, x) \mathbf{p}(\mathbf{f}(\delta), \mathbf{f}(x)) d\delta \quad (2)$$

where $\mathbf{y}(x)$ and $\mathbf{f}(x)$ denotes a pixel of the output and input image, $\mathbf{g}(\delta, x)$ evaluates the geometric closeness between the

pixel at the center x and a nearby point δ , $p(f(\delta), f(x))$ evaluates the photometric similarity between the center x and a nearby point δ . Thus both domain and range filtering are considered by the bilateral filtering. We define the bilateral filtering as

$$\mathbf{y}_n = \Delta^k(y_n)\boldsymbol{\beta} + \varepsilon_n \quad (3)$$

where y_n is a pixel of the target image, $\Delta^k(y_n)$ defines k member neighborhood vector of y_n , $\boldsymbol{\beta} = (\beta_1, \beta_2, \dots, \beta_k)^T$ is a vector of bilateral filtering coefficients, and ε_n is the error term between the predictions and the input visual signals. Then free energy between the input signal \mathbf{y}_n and the filtered one $\Delta^k(y_n)\boldsymbol{\beta}$ can be calculated. Please refer to [9] for more details about free energy entropy computation.

Since the reference image is not available, we propose to approximate the reference image's free energy using the distorted image. We down-sample the distorted image to one eighth of its original resolution. The down-sampled image is then taken as the target image \mathbf{y}_n and the free energy is computed as F'_a . We believe that a high degree of down-sampling can significantly weaken the influence of various kinds of distortions, and the left information is mainly related to the image content rather than distortions. So we can use the free energy entropy of the distorted image to approximate the free energy entropy of corresponding down-sampled reference image. Then a new natural scene statistic (NSS) model is used to infer free energy entropy of reference image (F_a) from the down-sampled one (F'_a). We conduct natural scene statistic analysis to explore the relationship between F_a and F'_a . Fig. 2 illustrates the scatter plot of all images from the Berkeley database [35]. Note that there is an approximate power relationship between F and F' . Thus the following power function model is used to predict F_a from F'_a :

$$f(F') = a \cdot (F')^b + c \quad (4)$$

where F' denotes free energy of the down-sampled image, and $f(F')$ indicates the estimated free energy of original image; a, b, c are model parameters, which are determined by curve fitting using all images from Berkeley database [35]. We fit the parameters on this database since it has no overlap with existing IQA databases which will be used as test-beds in later experiments. Then free energy entropy of the reference image F_a is approximated as: $F_a = f(F'_a)$.

B. Estimating free energy and saliency of the distorted image

Second, we analyze the distorted image based on human visual system and visual saliency. According to the recent free energy principle which works on the assumption that human cognitive process is decided by an internal generative model [24], the brain can predict the input scene in a constructive way. For example, the human visual system via the internal generative model makes an effort to reduce the uncertain information of an input visual signal. Therefore, difference always exists between the brain's prediction and the input scene. We believe that this difference is highly related to visual saliency. And we utilize the linear autoregressive(AR) model

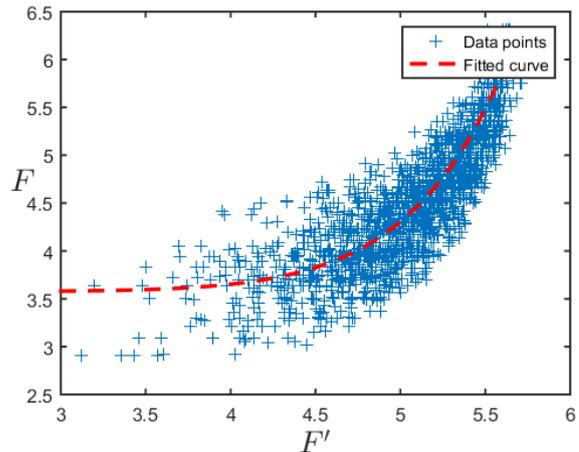


Fig. 2. Scatter plot of free energy entropy of the original image F versus that of corresponding down-sampled image F' . Note that there is an approximate power relationship between F and F' .

to simulate the generative model to predict an input distorted image for AR model is simple and effective to describe a variety of natural scenes [9], [25], [29]. The AR model can be defined as

$$\mathbf{y}_n = \Delta^k(y_n)\boldsymbol{\alpha} + \xi_n \quad (5)$$

where y_n is a pixel of the distorted image, $\Delta^k(y_n)$ is a row-vector which includes k nearest neighbors of y_n , $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \dots, \alpha_k)^T$ is a AR parameter vector, and ξ_n is the error term between the predictions and the input visual signals.

Because of their respective merits, we compute the free energy entropy F_b between the input scene and the predicted one by using the semi-parametric model which combines both the non-parametric model bilateral filtering and the parametric model AR. In detail, we use the bilateral filter described by Eq. (3) and the AR model described in Eq. (5) to predict the distorted image:

$$\mathbf{z}_n = w_1 \Delta^k(y_n)\boldsymbol{\beta} + w_2 \Delta^k(y_n)\boldsymbol{\alpha} \quad (6)$$

where \mathbf{z}_n is the predicted image; y_n is a pixel of the distorted image; $\Delta^k(y_n)\boldsymbol{\beta}$ and $\Delta^k(y_n)\boldsymbol{\alpha}$ are estimated by Eq. (3) and Eq. (5), respectively; w_1, w_2 control the relative importance of two components. Free energy entropy of the distorted image F_b is then calculated between \mathbf{z}_n and \mathbf{y}_n .

In fact, the internal generative model can be approximated by a probabilistic model, which includes a prior term and a likelihood term. It is clear that there always exists a gap between the input signal and the internal generative model's prediction. And we believe the gap between the input signal and the corresponding prediction is highly related to the human visual perception, and thereby can be utilized for saliency detection. Visual saliency, which are generally "pop-outs" of images, thus can be predicted by the local entropy of the gap. For more details about calculating saliency using free energy principle, please refer to [25]. Visual saliency helps us to discriminate between textured and smooth regions of

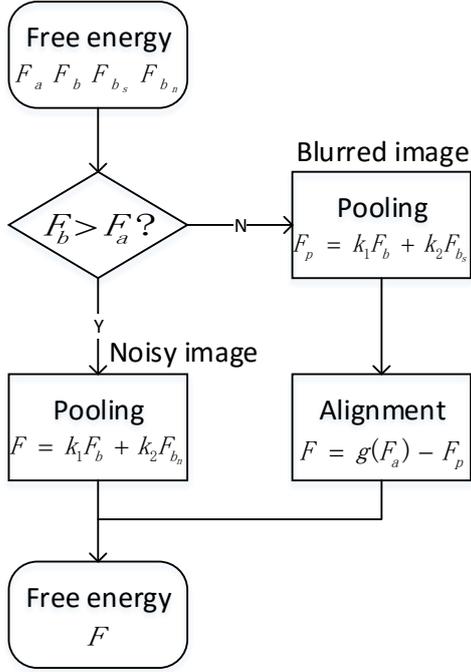


Fig. 3. Flowchart of the pooling process. We adopt different pooling strategies for different distortion types.

the image according to the distortion type, which makes our model more effective. After saliency detection, we compute the free energy entropy of the distorted image's saliency and non-saliency areas (denoted as F_{b_s} and F_{b_n} , respectively), which will be fused with F_a and F_b in the following stage.

C. Free energy pooling

At last, we fuse the afore-computed free energy entropy according to the distortion types. Fig 3 illustrates a flowchart of the pooling process. We first compare the two free energy entropy F_b and F_a . According to the compare results, we adopt different strategies. If F_b is smaller than F_a , it means that the target image is a blurry one, then we fuse the free energy entropy of saliency region (F_{b_s}) and entire image (F_b). Otherwise, the target image is corrupted by noise, and we integrate the free energy entropy of non-saliency region (F_{b_n}) and entire image (F_b). We adopt such pooling strategies based on the phenomenon that noise in the smooth region and blur in the textured region do much more harm to the perceptual quality. So we focus more on saliency region when image is corrupted by blur, whereas pay more attention to non-saliency region when the degradation is noise. For both conditions, linear fusion is adopted because of its simplicity.

The last problem is the alignment of noisy and blurry image's free energy entropy. Image with more noise generally has a larger free energy entropy, whereas it will be smaller for image with a stronger blurriness. So we need to align them in a common scale to assess the quality of noisy and blurry images simultaneously. We adopt the following mapping function to

blurry images:

$$F = g(F_a) - F_p \quad (7)$$

where F, F_p are free energy entropy of the blurry image after and before alignment; $g(*)$ is a simple linear function.

Finally, quality of both blurry and noisy image can be described by the fused free energy entropy F :

$$F = \begin{cases} k_1 F_b + k_2 F_{b_n} & \text{if } F_b > F_a \\ g(F_a) - (k_1 F_b + k_2 F_{b_s}) & \text{otherwise} \end{cases} \quad (8)$$

where k_1, k_2 are weighting parameters; $g(*)$ is a linear function; $F_a, F_b, F_{b_s}, F_{b_n}$ have been described above.

III. EXPERIMENTAL RESULTS AND ANALYSIS

A. Experiment settings

1) *Testing databases:* In our experiments, three popular natural scene image databases and a recently released screen content image database are used to evaluate the performance of the proposed method. These four databases are LIVE [21], Tampere Image Database 2013 (TID2013) [23], Categorical Subjective Image Quality database (CSIQ) [22], and Screen Image Quality Assessment Database (SIQAD) [7]. The numbers of noisy images in LIVE, CSIQ, TID2013 and SIQAD are 145, 150, 125, and 140, respectively. While, numbers of blurry images in LIVE, CSIQ, TID2013 and SIQAD are 145, 150, 125, and 140, respectively. Thus 560 noisy images, 560 blurry images, a total of 1120 images are used in the experiments.

2) *Compared algorithms:* We select several classical FR IQA and recent general-purpose and distortion-specific N-R/blind IQA methods for comparison. They are: 1) Two FR IQA metrics, structural similarity index (SSIM) that compares contrast, structural similarities and luminance [5] and local-tuned-global (LTG) [6] which measures global quality degradation, salient local distortions in luminance and chrominance information; 2) Five general-purpose NR IQA metrics, BRISQUE [28], NFERM [29], NIQE [30], IL-NIQE [31] and SISBLIM [32]; 3) Five distortion-specific NR IQA models, S3 [16], FISH [17], FISHbb [18], ARISM_c [?] and SINE [19].

3) *Evaluation criteria:* We adopt the video quality experts group (VQEG)'s suggestion [27] and use four criteria to compare the performance of the quality metrics. They are Spearman rank order correlation coefficient (SRCC), Pearson linear correlation coefficient (PLCC), Root mean square error (RMSE), and Kendall's rank correlation coefficient (KRCC). PLCC and RMSE are used to evaluate the prediction accuracy. SRCC and KRCC are used to evaluate the prediction monotonicity. A good IQA method will produce a higher values of SRCC, KRCC, PLCC, and a lower values of RMSE. In our experiment, we adopt the four-parameter nonlinear logistic function to map the prediction results x to the subjective scores:

$$f(x) = \frac{\alpha_1 - \alpha_2}{1 + e^{-\frac{(x - \alpha_3)}{\alpha_4}}} + \alpha_2 \quad (9)$$

where $\alpha_1, \alpha_2, \alpha_3, \alpha_4$ are the parameters to be fitted, $f(x)$ and x indicate the mapped score and the input score.

4) *Testing process*: We use both noise and blur subsets as testing beds. Quality measures are required to evaluate both distortions simultaneously. Quality scores of the distorted images are computed in each database respectively. Then the predicted scores are compared with the subjective scores to judge the performance of the quality measure.

B. Experimental results

In Table I, the performances on the whole noise and blur image subsets of four databases are presented. To facilitate overall comparison, Table I also provides the direct average and the database size-weighted average (compute the mean values according to the size of each image subset, i.e., 290 for LIVE, 300 for CSIQ, 250 for TID2013 and 280 for SIQAD) performances for all compared measures. Since NFERM and BRISQUE use LIVE database for training, we do not list their performances on LIVE database and the average performances for these two measures are computed over CSIQ, TID2013 and SIQAD. We also do not compare the average RMSE values because RMSE is highly related to the range of subjective ratings.

From Table I, we can see that the proposed method is superior to the compared general-purpose and specific-distortion NR/blind IQA methods. Generally, NR/blind IQA measures are not compared with FR IQA metrics. We can still observe that the proposed BQISN model is even better than the FR SSIM, while it is a little inferior to the FR LTG on average. Furthermore, from an average perspective, the proposed BQISN performs substantially better than the second best NFERM measure according to all criteria except for PLCC.

IV. CONCLUSION

In this paper we put forward a novel training-free no-reference quality metric (QM) for image sharpness and noise. It is based on a unified brain theory - the free energy principle, which tells that there always exists a difference between an input true visual signal and its processed one by human brain. The proposed method first computes the visual quality and saliency between the real and processed signals via a semi-parametric method that is constructed by combining bilateral filter and auto-regression model. Afterwards, the computed visual saliency and a new natural scene statistic (NSS) model are used for modification to infer the final visual quality score. Extensive experiments have been done to evaluate the performance of the proposed BQISN across three popular natural scene image databases and a recently released screen content image database. The experiment results have proved the effectiveness of the proposed blind quality measure compared with classical and state-of-the-art full- and no-reference image quality metrics.

ACKNOWLEDGMENT

This work is kindly supported by National Science Foundation of China (61379143), the Fundamental Research Funds for the Central Universities (2015QNA66) and Science and Technology Planning Project of Nantong (BK2014022).

- [1] K. Gu, G. Zhai, X. Yang, W. Zhang, and C. W. Chen, "Automatic contrast enhancement technology with saliency preservation," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 25, no. 9, pp. 1480-1494, Sept. 2015.
- [2] K. Gu, G. Zhai, W. Lin, and M. Liu, "The analysis of image contrast: From quality assessment to automatic enhancement," *IEEE Trans. Cybernetics*, vol. 46, no. 1, pp. 284-297, Jan. 2016.
- [3] K. Gu, S. Wang, G. Zhai, S. Ma, X. Yang, W. Lin, W. Zhang, and W. Gao, "Blind quality assessment of tone-mapped images via analysis of information, naturalness, and structure," *IEEE Trans. Multimedia*, vol. 18, no. 3, pp. 432-443, Mar. 2016.
- [4] W. Lin and C.-C. Jay Kuo, "Perceptual visual quality metrics: A survey," *J. Vis. Commun. Image Represent.*, vol. 22, no. 4, pp. 297-312, 2011.
- [5] 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.
- [6] K. Gu, G. Zhai, X. Yang, and W. Zhang, "An efficient color image quality metric with local-tuned-global model," in *Proc. IEEE Int. Conf. Image Process.*, pp. 506-510, Oct. 2014.
- [7] 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.
- [8] K. Gu, S. Wang, H. Yang, W. Lin, G. Zhai, X. Yang, and W. Zhang, "Saliency-guided quality assessment of screen content images," *IEEE Trans. Multimedia*, vol. 18, no. 6, pp. 1-13, Jun. 2016.
- [9] S. Wang, K. Gu, X. Zhang, W. Lin, L. Zhang, S. Ma, and W. Gao, "Subjective and objective quality assessment of compressed screen content images," *IEEE J. Emerg. Sel. T. Circuits Syst.*, 2016.
- [10] 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 and Syst.*, pp. 1095-1098, May 2013.
- [11] S. A. Golestaneh and D. M. Chandler, "No-reference quality assessment of JPEG images via a quality relevance map," *IEEE Sig. Process. Lett.*, vol. 21, no. 2, pp. 155-158, Feb. 2014.
- [12] X. Min, G. Zhai, K. Gu, Y. Fang, X. Yang, X. Wu, J. Zhou, X. Liu, "Blind quality assessment of compressed images via pseudo structural similarity," in *Proc. IEEE Int. Conf. Multimedia and Expo*, pp. 1-6, Jul. 2016.
- [13] S. Wu, W. Lin, S. Xie, Z. Lu, E. Ong, and S. Yao, "Blind blur assessment for vision-based applications," *J. Vis. Commun. Image Represent.*, vol. 20, no. 4, pp. 231-241, May 2009.
- [14] R. Ferzli and L. Karam, "A no-reference objective image sharpness metric based on the notion of just noticeable blur (JNB)," *IEEE Trans. Image Process.*, vol. 18, no. 4, pp. 717-728, Apr. 2009.
- [15] N. D. Narvekar and L. J. Karam, "A no-reference image blur metric based on the cumulative probability of blur detection (CPBD)," *IEEE Trans. Image Process.*, vol. 20, no. 9, pp. 2678-2683, Sept. 2011.
- [16] C. Vu, T. Phan, and D. Chandler, "S₃: A spectral and spatial measure of local perceived sharpness in natural images," *IEEE Trans. Image Process.*, vol. 21, no. 3, pp. 934-945, Mar. 2012.
- [17] P. Vu and D. Chandler, "A fast wavelet-based algorithm for global and local image sharpness estimation," *IEEE Sig. Process. Lett.*, vol. 19, no. 7, pp. 423-426, Jul. 2012.
- [18] K. Gu, G. Zhai, W. Lin, X. Yang, and W. Zhang, "No-reference image sharpness assessment in autoregressive parameter space," *IEEE Trans. Image Process.*, vol. 24, no. 10, pp. 3218-3231, Oct. 2015.
- [19] D. Zoran and Y. Weiss, "Scale invariance and noise in natural images," in *Proc. IEEE Int. Conf. Comput. Vis.*, pp. 2209-2216, Sept. 2009.
- [20] K. Gu, W. Lin, G. Zhai, X. Yang, W. Zhang, and C. W. Chen, "No-reference quality metric of contrast-distorted images based on information maximization," *IEEE Trans. Cybernetics*, 2016.
- [21] H. R. Sheikh, Z. Wang, L. Cormack, and A. C. Bovik, "LIVE image quality assessment Database Release 2," [Online]. Available: <http://live.ece.utexas.edu/research/quality>
- [22] E. C. Larson and D. M. Chandler, "Categorical image quality (CSIQ) database," [Online]. Available: <http://vision.okstate.edu/csiq>
- [23] N. Ponomarenko *et al.*, "Image database TID2013: Peculiarities, results and perspectives," *SPIC*, vol. 30, pp. 57-55, Jan. 2015.
- [24] K. Friston, "The free-energy principle: A unified brain theory?" *Nature Reviews Neuroscience*, vol. 11, pp. 127-138, 2010.
- [25] K. Gu, G. Zhai, W. Lin, X. Yang and W. Zhang, "Visual saliency detection with free energy theory," *IEEE Signal Process. Lett.*, vol. 22, no. 10, pp. 1552-1555, Oct. 2015.

TABLE I
PERFORMANCE COMPARISONS OF 13 IQA MEASURES ON FOUR DATABASES. WE HIGHLIGHT THE TOP TWO NR MEASURES.

LIVE Noise and Blur Database (290 Images) [21]				
Metrics	SRCC	KRCC	PLCC	RMSE
SSIM [5]	0.9115	0.7388	0.8865	11.008
LTG [6]	0.9736	0.8579	0.8927	7.2393
BRISQUE [28]	training images			
NFERM [29]	training images			
NIQE [30]	0.9283	0.7655	0.7891	8.6431
IL-NIQE [31]	0.9509	0.8130	0.8699	7.1303
SISBLIM [32]	0.9390	0.7918	0.7989	7.6708
S ₃ [16]	0.2089	0.0996	0.4857	16.521
FISH [17]	0.2332	0.1350	0.3456	16.203
FISH _{bb} [17]	0.1961	0.0868	0.1441	16.816
ARISM _c [18]	0.1193	0.0055	0.2484	24.318
SINE [19]	0.4155	0.3144	0.6773	15.960
BQISN (Pro.)	0.9472	0.7986	0.9199	8.3552

CSIQ Noise And Blur Database (300 Images) [22]				
Metrics	SRCC	KRCC	PLCC	RMSE
SSIM [5]	0.8616	0.6603	0.7568	0.1477
LTG [6]	0.9657	0.8337	0.8888	0.0758
BRISQUE [28]	0.9076	0.7447	0.9302	0.0889
NFERM [29]	0.9024	0.7379	0.9260	0.0923
NIQE [30]	0.8396	0.6481	0.8657	0.1226
IL-NIQE [31]	0.8477	0.6641	0.7971	0.1121
SISBLIM [32]	0.8756	0.6978	0.8399	0.0965
S ₃ [16]	0.3081	0.2069	0.3878	0.1589
FISH [17]	0.3092	0.2095	0.5585	0.1502
FISH _{bb} [17]	0.4022	0.2879	0.6835	0.2455
ARISM _c [18]	0.2891	0.1923	0.2991	0.1455
SINE [19]	0.1301	0.1213	0.1302	0.2386
BQISN (Pro.)	0.8778	0.6935	0.8137	0.1122

TID2013 Noise And Blur Database (250 Images) [22]				
Metrics	SRCC	KRCC	PLCC	RMSE
SSIM [5]	0.8036	0.5995	0.7761	0.65
LTG [6]	0.9326	0.7754	0.86	0.3879
BRISQUE [28]	0.8038	0.6225	0.7793	0.6068
NFERM [29]	0.8410	0.6531	0.8224	0.5736
NIQE [30]	0.6772	0.4673	0.6446	0.7908
IL-NIQE [31]	0.8148	0.6244	0.8391	0.5692
SISBLIM [32]	0.7527	0.5489	0.7657	0.6634
S ₃ [16]	0.3062	0.1759	0.3387	0.7051
FISH [17]	0.3029	0.1606	0.4752	0.7215
FISH _{bb} [17]	0.4111	0.2702	0.6299	0.6952
ARISM _c [18]	0.3405	0.2119	0.2345	0.6852
SINE [19]	0.0014	0.0458	0.0202	1.0401
BQISN (Pro.)	0.8611	0.6748	0.7971	1.0626

SIQAD Noise And Blur Database (280 Images) [7]				
Metrics	SRCC	KRCC	PLCC	RMSE
SSIM [5]	0.6817	0.4792	0.6329	10.630
LTG [6]	0.8716	0.6759	0.7935	7.1283
BRISQUE [28]	0.6689	0.4757	0.6743	11.086
NFERM [29]	0.7923	0.6015	0.7825	9.2679
NIQE [30]	0.445	0.3042	0.4402	13.531
IL-NIQE [31]	0.1788	0.1358	0.1754	14.571
SISBLIM [32]	0.6755	0.4672	0.5873	10.925
S ₃ [16]	0.1922	0.1283	0.0365	12.886
FISH [17]	0.1337	0.0908	0.2595	12.499
FISH _{bb} [17]	0.2445	0.181	0.3943	15.081
ARISM _c [18]	0.2047	0.1443	0.0976	15.17
SINE [19]	0.3628	0.2602	0.3485	13.958
BQISN (Pro.)	0.7731	0.5571	0.7807	9.1597

Direct Average				
Metrics	SRCC	KRCC	PLCC	RMSE
SSIM [5]	0.8146	0.6195	0.7631	-
LTG [6]	0.9359	0.7857	0.8588	-
BRISQUE [28]	0.7934	0.6143	0.7946	-
NFERM [29]	0.8452	0.6642	0.8436	-
NIQE [30]	0.7225	0.5463	0.6849	-
IL-NIQE [31]	0.6981	0.5593	0.6704	-
SISBLIM [32]	0.8107	0.6264	0.7480	-
S ₃ [16]	0.2539	0.1527	0.3122	-
FISH [17]	0.2448	0.1490	0.4097	-
FISH _{bb} [17]	0.3135	0.2065	0.4630	-
ARISM _c [18]	0.2384	0.1385	0.2199	-
SINE [19]	0.2275	0.1854	0.2941	-
BQISN (Pro.)	0.8648	0.6810	0.8279	-

Database Size-Weighted Average				
Metrics	SRCC	KRCC	PLCC	RMSE
SSIM [5]	0.8166	0.6218	0.7637	-
LTG [6]	0.9368	0.7875	0.8596	-
BRISQUE [28]	0.7958	0.6171	0.7984	-
NFERM [29]	0.8468	0.6663	0.8464	-
NIQE [30]	0.7277	0.5522	0.6901	-
IL-NIQE [31]	0.6999	0.5617	0.6699	-
SISBLIM [32]	0.8146	0.6313	0.7496	-
S ₃ [16]	0.2530	0.1525	0.3144	-
FISH [17]	0.2442	0.1496	0.4100	-
FISH _{bb} [17]	0.3114	0.2052	0.4596	-
ARISM _c [18]	0.2355	0.1363	0.2212	-
SINE [19]	0.2334	0.1892	0.3019	-
BQISN (Pro.)	0.8838	0.7038	0.8379	-

- [26] C. Tomasi and R. Manduchi, "Bilateral filtering for gray and color images," *ICCV*, pp. 836-846, Jan. 1998.
- [27] VQEG, "Final report from the video quality experts group on the validation of objective models of video quality assessment," Mar. 2000, <http://www.vqeg.org/>.
- [28] 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.
- [29] 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.
- [30] A. Mittal, R. Soundararajan and A. C. Bovik, "Making a completely blind image quality analyzer," *IEEE Signal Process. Lett.*, vol. 20, no. 3, pp. 209-212, Mar. 2013.
- [31] L. Zhang, L. Zhang, and A. C. Bovik, "A feature-enriched completely blind image quality evaluator," *IEEE Trans. on Image Process.*, vol. 24, no. 8, pp. 2579-2591, Aug. 2015.
- [32] K. Gu, G. Zhai, X. Yang, and W. Zhang, "Hybrid no-reference quality metric for singly and multiply distorted images," *IEEE Trans. on Broadcasting*, vol. 60, no. 3, pp. 555-567, Aug. 2014.
- [33] 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.
- [34] Methodology for the subjective assessment of the quality of television pictures. International Telecommunication Union Recommendation ITUR BT.500-13, 2012.
- [35] D. Martin, C. Fowlkes, D. Tal and J. Malik, "A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics," *ICCV*, pp. 416-423, 2001.
- [36] V. Jakhethiya, O. C. Au, S. Jaiswal, L. Jia and H. Zhang, "Fast and efficient intra-frame deinterlacing using observation model based bilateral filter," *ICASSP*, pp. 5819-5823, 2014.