No-Reference Image Sharpness Assessment in Autoregressive Parameter Space

Ke Gu, *Student Member, IEEE*, Guangtao Zhai, *Member, IEEE*, Weisi Lin, *Senior Member, IEEE*, Xiaokang Yang, *Senior Member, IEEE*, and Wenjun Zhang, *Fellow, IEEE*

Abstract—In this paper, we propose a new no-reference (NR)/ blind sharpness metric in the autoregressive (AR) parameter space. Our model is established via the analysis of AR model parameters, first calculating the energy- and contrast-differences in the locally estimated AR coefficients in a pointwise way, and then quantifying the image sharpness with percentile pooling to predict the overall score. In addition to the luminance domain, we further consider the inevitable effect of color information on visual perception to sharpness and thereby extend the above model to the widely used YIQ color space. Validation of our technique is conducted on the subsets with blurring artifacts from four large-scale image databases (LIVE, TID2008, CSIQ, and TID2013). Experimental results confirm the superiority and efficiency of our method over existing NR algorithms, the state-of-the-art blind sharpness/blurriness estimators, and classical full-reference quality evaluators. Furthermore, the proposed metric can be also extended to stereoscopic images based on binocular rivalry, and attains remarkably high performance on LIVE3D-I and LIVE3D-II databases.

Index Terms—Image sharpness/blurriness, image quality assessment (IQA), no-reference (NR)/blind, autoregressive (AR) parameters, YIQ color space, stereoscopic image, binocular rivalry.

I. INTRODUCTION

NOWADAYS, the expectation of human consumers toward enjoyment of high-quality images is constantly rising. Owing to the limitations of bandwidth and storage media, images however very possibly suffer some typical types of distortions, e.g. white noise and Gaussian blur, before finally reaching to human consumers. Classical full-reference (FR) image quality assessment (IQA), supposing that the original and distorted images are both entirely known, can assess those degradation levels [1]–[6]. But the pristine image is not available in most cases, and thus blind/no-reference (NR)

IQA metrics without access to original references are highly desirable. For noise estimation, these years have witnessed the emergence of quite a few blind algorithms [7], [8]. Though a large set of sharpness/blurriness measures have been developed, their performance indices are far less than the ideal results. Furthermore, this type of approaches are of many valuable applications in image processing, such as automatic contrast enhancement [9], [10], super-resolution [11] and denoising [12]. Therefore, in this work we devote to inducing a high-accuracy blind image sharpness metric.

Early attempts of sharpness/blurriness estimations mainly concentrated on image edges. In [13], a perceptual model was developed based on a pair of edge detectors for vertical and horizontal directions. In [14], Wu *et al.* proposed a blind blur evaluator by computing the point spread function (PSF) from the line spread function (LSF) that is extracted from edges in a blurred image. In [15], the authors computed the edge width in 8×8 blocks before a measure of just-noticeable blur (JNB) factor. Inspired by the successfulness of JNB, the cumulative probability of detecting blur (CPDB) algorithm [16] predicts the image sharpness by calculating the probability of blurriness at each edge.

Over the last few years, there have also existed some blind techniques with some level of success in assessing perceptual sharpness. In [17], the authors combined spatial and transform-based features to induce a hybrid approach, dubbed as spectral and spatial sharpness (S_3). Specifically, the slope of the local magnitude spectrum and total variation is first used to create a sharpness map, and then the scalar index of S_3 is computed as the average of the 1% highest values in that sharpness map. Thereafter, a transform-inspired fast image sharpness (FISH) model [18] was explored with the evaluation of log-energies in high-frequency DWT subbands followed by a weighted mean of the log-energies.

Very recently, Feichtenhofer *et al.* developed a perceptual sharpness index (PSI) [19] by analyzing the edge slopes before integrating an acutance measure to model the influence of local contrast information on the perception to image sharpness. In [20], Wang and Simoncelli analyzed the local phase coherence (LPC) and pointed out that the phases of complex wavelet coefficients constitute a highly predictable pattern in the scale space in the vicinity of sharp image features, and furthermore, the LPC structure was found to be disrupted by image blur. With this concern, Hassen *et al.* designed the valid LPC-based sharpness index (LPC-SI) [21].

Besides, several NR IQA metrics were proved effectively in assessing image blur. The authors in [22] made use of the

recent free energy based brain theory [23] to simulate the internal generative mechanism of the brain, and introduced the NR free energy based quality metric (NFEQM). Distortion Identification-based Image Verity and INtegrity Evaluation (DIIVINE) [24], BLind Image Integrity Notator using DCT Statistics (BLIINDS-II) [25] and Blind/Referenceless Image Spatial Quality Evaluator (BRISQUE) [26] came from the natural scene statistics (NSS) model [27], working with the feature extraction and the training of a regression module via the support vector machine (SVM) [28]. Along this research line, we lately designed the NFSDM [29] and NFERM [30] by systematically integrating two effective reduced-reference (RR)[1] quality metrics in [22] and [31] to eliminate the demand of references.

Differing from previous methods, in this paper we come up with a new blind sharpness measure based on the analysis of autoregressive (AR) model parameters, dubbed as AR-based Image Sharpness Metric (ARISM). Our technique is inspired by the free energy principle and the NFEQM model, built upon the underlying hypothesis that image blurring increases the resemblance of locally estimated AR parameters. Particularly, the proposed ARISM works to separately measure the energy- and contrast-difference of AR model coefficients at each pixel, and then compute the image sharpness with percentile pooling to deduce the overall quality score.

Currently, since three-dimensional (3D) imaging technology works actively from entertainment (e.g. videos and games) to specialized domains (e.g. education and medicine), a growing number of image processing operations have been specifically explored for stereoscopic images, and thereby the necessity of stereoscopic IQA methods shows strongly evident, especially under the NR condition. There have been many related studies extending 2D IQA models to 3D images. In [32], the fusion of 2D quality scores of the left- and right-eye images is used to infer the stereoscopic image quality. In [33], the degradation of edges in the depth map is used as the 3D image quality. In [34]–[36], the authors fused the quality measure of the disparity map with those of left- and right-views to infer the visual quality of stereoscopic images.

Following this research line, we further endeavor to modify the proposed ARISM for the sharpness assessment of stereoscopic images, based on existing studies on binocular rivalry [37]–[39], where it was found that for simple ideal stimuli, a rising contrast advances the predominance of one view against the other. We reasonably suppose that the contrast increases with the difference of AR parameters. Thus, a 3D sharpness measure can be established using the weighted sum of energy- and contrast-differences to weight the ARISM model.

The remainder of this paper proceeds as follows: Section II first reviews our previous related work. Section III introduces the motivation of our approach and describes its framework in detail. A comparison of ARISM with state-of-the-art metrics using blur data sets obtained from four monoscopic

image databases (LIVE [40], TID2008 [41], CSIQ [42], and TID2013 [43]) is given in Section IV. In Section V, the proposed model is extended to a stereoscopic sharpness measure and is verified on LIVE3D-I [44] and LIVE3D-II [45] databases. We finally conclude this paper in Section VI.

II. RELATED WORK

In a recent work [22], the simple yet valid NFEQM method was proposed based on the concept of the free energy theory, which was lately revealed in [23] and it succeeds in explaining and unifying several existing brain theories in biological and physical sciences about human action, perception and learning. The fundamental assumption of the free energy principle is that the cognitive process is controlled by an internal generative model in the brain, similar to the Bayesian brain hypothesis [46]. Depending on this model, the brain is able to use a constructive way to actively infer predictions of the meaningful information from input visual signals and reduce the residual uncertainty.

The aforesaid constructive manner can be approximated by a probabilistic model, which can be separated into a likelihood term and a prior term. For a given scene, the human visual system can deduce its posterior possibilities by inverting the likelihood term. It is natural that there always exists a gap between the real external scene and the brain's prediction, for the reason that the internal generative model cannot be universal everywhere. We believe that this gap between the external input signal and its generative-model-explainable part is highly connected to the quality of visual sensations, and is applicable to the measurement of image sharpness.

Specifically, we postulate that the internal generative model g is parametric for visual sensation, and the perceived scenes can be explained by adjusting the parameter vector $\boldsymbol{\phi}$. Given a visual signal $\mathbf{s}$, its "surprise" (measured by entropy) can be obtained by integrating the joint distribution $p(\mathbf{s}, \boldsymbol{\phi}|g)$ over the space of model parameters $\boldsymbol{\phi}$

$$-\log p(\mathbf{s}|g) = -\log \int p(\mathbf{s}, \boldsymbol{\phi}|g)d\boldsymbol{\phi}. \qquad (1)$$

We bring an auxiliary term $q(\boldsymbol{\phi}|\mathbf{s})$ into both the denominator and numerator in Eq. (1) and derive:

$$-\log p(\mathbf{s}|g) = -\log \int q(\boldsymbol{\phi}|\mathbf{s})\frac{p(\mathbf{s}, \boldsymbol{\phi}|g)}{q(\boldsymbol{\phi}|\mathbf{s})}d\boldsymbol{\phi}. \qquad (2)$$

Here $q(\boldsymbol{\phi}|\mathbf{s})$ is an auxiliary posterior distribution of the model parameters given the input image signal $\mathbf{s}$. It can be thought of as an approximate posterior to the true posterior of the model parameters $p(\boldsymbol{\phi}|\mathbf{s}, g)$ given by the brain. When perceiving the image signal $\mathbf{s}$ or when adjusting the parameters $\boldsymbol{\phi}$ in $q(\boldsymbol{\phi}|\mathbf{s})$ to search for the optimal explanation of $\mathbf{s}$, the brain will minimize the discrepancy between the approximate posterior $q(\boldsymbol{\phi}|\mathbf{s})$ and the true posterior $p(\boldsymbol{\phi}|\mathbf{s}, g)$.

Next, the dependence on the model g will be dropped for simplicity. Using the Jensen's inequality, we can easily get the following relationship from Eq. (2):

$$-\log p(\mathbf{s}) \leq -\int q(\boldsymbol{\phi}|\mathbf{s})\log \frac{p(\mathbf{s}, \boldsymbol{\phi})}{q(\boldsymbol{\phi}|\mathbf{s})}d\boldsymbol{\phi}. \qquad (3)$$

---

[1]RR IQA works under the situation that the partial original image or some extracted features are available as auxiliary information for quality evaluation.
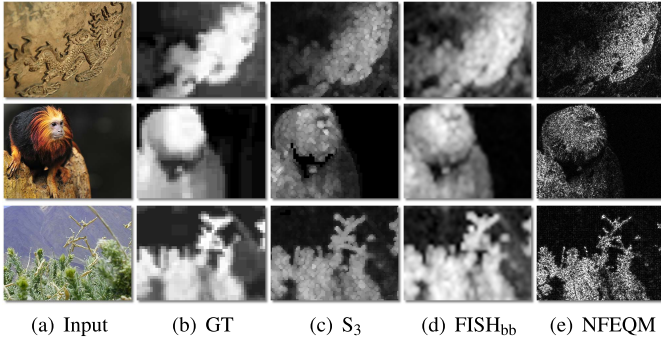
(a) Input	(b) GT	(c) S_3	(d) FISH$_{bb}$	(e) NFEQM

Fig. 1. Comparison of local sharpness maps of ground truth, S_3, FISH$_{bb}$ and NFEQM using representative images in [17].

Fig. 2. Comparison of NFEQM and ARISM frameworks.

The right hand side of Eq. (3) is the upper bound by a term called "free energy", which is defined as

$$f(\boldsymbol{\phi}) = -\int q(\boldsymbol{\phi}|\mathbf{s}) \log \frac{p(\mathbf{s}, \boldsymbol{\phi})}{q(\boldsymbol{\phi}|\mathbf{s})} d\boldsymbol{\phi}. \qquad (4)$$

The free energy measures the discrepancy between the input visual signal and its best explanation given by the internal generative model, and thus it can be considered as a natural proxy for psychovisual quality of images. This motivates the use of free energy for the design of NFEQM in the image sharpness/blurriness measure:

$$\text{NFEQM}(\mathbf{s}) = f(\hat{\boldsymbol{\phi}}) \quad \text{with} \quad \hat{\boldsymbol{\phi}} = \arg\min_{\boldsymbol{\phi}} f(\boldsymbol{\phi}|g, \mathbf{s}). \qquad (5)$$

The linear autoregressive (AR) model is used for approximating g, because this model is easy to construct and has a good ability to characterize a wide range of natural scenes by varying its parameters [47]–[49]. For an input visual signal $\mathbf{s}$, we define the AR model as

$$s_n = \mathcal{V}^t(s_n)\,\boldsymbol{v} + \epsilon_n \qquad (6)$$

where s_n is a pixel in question. $\mathcal{V}^t(s_n)$ is a vector of t nearest neighbors of s_n. $\boldsymbol{v} = (v_1, v_2, ..., v_t)^T$ is a vector of AR model coefficients. The superscript "T" means transpose. ϵ_n is the error term. To determine $\boldsymbol{v}$, the linear system can be written in matrix form as

$$\hat{\boldsymbol{v}} = \arg\min_{\boldsymbol{v}} \|\mathbf{s} - \mathbf{V}\boldsymbol{v}\|_2 \qquad (7)$$

where $\mathbf{s} = (s_1, s_2, ..., s_t)^T$ and $\mathbf{V}(i, :) = \mathcal{V}^t(s_i)$. This linear system was solved with the least square method, leading to $\hat{\boldsymbol{v}} = (\mathbf{V}^T\mathbf{V})^{-1}\mathbf{V}^T\mathbf{V}$. Next, we estimated $\hat{\mathbf{s}}$ to be

$$\hat{s}_n = \mathcal{V}^t(s_n)\,\hat{\boldsymbol{v}}. \qquad (8)$$

Referring to the analysis in [22], the process of free-energy minimization is closely related to predictive coding, and it can be finally approximated as the entropy of the prediction residuals between $\mathbf{s}$ and $\hat{\mathbf{s}}$ for a given AR model of fixed orders. Thus, the free energy of the input image signal is quantified by

$$\text{NFEQM}(\mathbf{s}) = -\sum_i p_i(\mathbf{s}_\triangle) \log p_i(\mathbf{s}_\triangle) \qquad (9)$$

where $\mathbf{s}_\triangle$ is the prediction error between the input visual signal and its predicted version. $p_i(\mathbf{s}_\triangle)$ is the probability density of grayscale i in $\mathbf{s}_\triangle$.
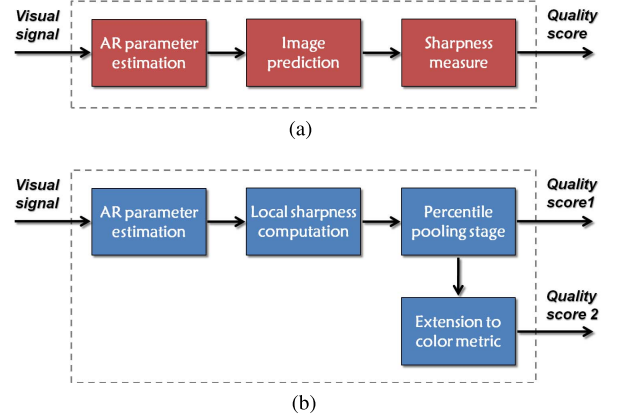
III. IMAGE SHARPNESS MEASURE

A. Motivation

The successfulness of NFEQM implies the effectiveness of AR model in measuring image sharpness. Fig. 1 exhibits the maps from ground truth and three sharpness metrics on three images "dragon", "monkey", and "peak" [17]. As compared to S_3 and FISH$_{bb}$, NFEQM shows fairly good estimation toward the ground truth maps. We can summary the whole process of NFEQM to be a three-step model: AR parameter estimation, image prediction by free energy, and sharpness measure in entropy, as presented in Fig. 2 (a).

However, it can be found that the core of free energy is that the parameters $\boldsymbol{\phi}$ in $q(\boldsymbol{\phi}|\mathbf{s})$ is adjusted to search for the optimal explanation of the visual signal $\mathbf{s}$, thus to minimize the discrepancy of the approximate posterior $q(\boldsymbol{\phi}|\mathbf{s})$ and the true posterior $p(\boldsymbol{\phi}|\mathbf{s}, g)$. So it is reasonable that the distribution of the parameters $\boldsymbol{\phi}$ is more closely related to the working of the brain's perception to image sharpness. Here the distribution of $q(\boldsymbol{\phi}|\mathbf{s})$ is represented by that of the estimated AR parameters, which exhibits a center-peaked appearance. In order to illustrate this, an image and its auxiliary posterior distribution of the model parameters $q(\boldsymbol{\phi}|\mathbf{s})$ computed using the first-order AR model are shown in Fig. 3.

Accordingly, we consider the use of AR model parameters, which were shown to invariant to object transformations (e.g. translation, rotation and scaling) and widely applied in the literature [50], [51], and thus concentrate on the analysis and adoption of AR coefficients in the proposed ARISM method. This is a distinguished difference between our technique and the previous NFEQM metric, which improves the performance in the sharpness measure to a sizable margin. We display this primary framework in Fig. 2 (b), which is composed of AR parameters estimation, local sharpness computation, percentile pooling stage and extension to color metric.

From another point of view, after the estimation of AR coefficients, the aforementioned two models utilize different dimensionality reduction strategies. NFEQM exploits pixels in the input and predicted images for blurriness measure, and thereby works in the spatial domain. In comparison, ARISM estimates the sharpness in the parameter space by analyzing the difference of locally estimated AR parameters
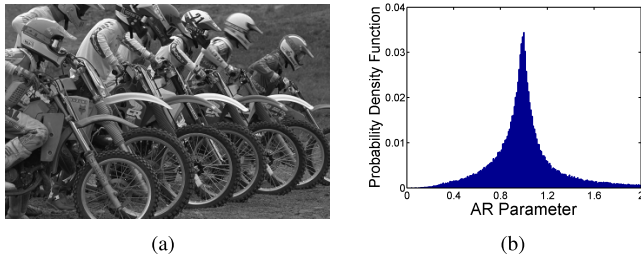
Fig. 3. Illustration of the posterior distribution of the model parameters $q(\phi|\mathbf{s})$ by: (a) a natural image; (b) the associated distribution of $q(\phi|\mathbf{s})$ computed using the first-order AR model.

in a point-wise way, which will be explicitly explained later.

Another distinguishment between our model and existing related methods (including NFEQM) is that ARISM considers the inevitable influence of color information on the sharpness assessment. Most approaches operate on the gray luminance image that is converted from the input color image signal $\mathbf{s}$ by the "rgb2gray" transform matrix:

$$\mathbf{s}_{gray} = [\mathbf{r} \ \mathbf{g} \ \mathbf{b}] \, [c_r \ c_g \ c_b]^T \qquad (10)$$

where $\mathbf{r}$, $\mathbf{g}$ and $\mathbf{b}$ indicate the vectors in $\mathbf{s}$. c_r, c_g and c_b are fixed as 0.299, 0.587 and 0.114. Only using gray information is not reasonable, because some edges might be removed by this transformation, which may result in the disappearance of sharpness in the color image after the above transformation. Thus our technique exploits the simple and widely used YIQ color space [52] for boosting the performance.

B. AR Parameters Estimation

As mentioned in the introduction, higher resemblance of AR parameters corresponding to one particular pixel indicates poorer sharpness of that location. The first step is to estimate the AR model coefficients for each pixel. Instead of using the AR parameters estimation in NFEQM, we employ another easier way to address this problem, which has been efficiently and effectively used for dimensionality reduction [53]. In our ARISM, an 8-th order AR model is trained for each image pixel and its 8-connected neighborhood to derive the optimal AR parameters.

C. Local Sharpness Computation

It is easy to imagine that eight AR model parameters of a pixel will be very close to each other when this pixel is in a comparatively smooth region, and on the other hand, these parameters tend to be obviously distinct when the current pixel belongs to a sharp zone. We pick two classical measures for this. The first one is defined as the difference between the maximum and minimum values of those AR parameters at the location of (i, j) in the input image S:[2]

$$E_{i,j} = |W_{\max} - W_{\min}|^n \qquad (11)$$

[2]For convenience, we use the image matrix S to represent the image signal $\mathbf{s}$ in the following pages. Similarly, the images or maps will be written in the form of matrixes.

where $W_{\max}$ and $W_{\min}$ are computed from the AR parameters as follows:

$$W_{\max} = \max_{(s,t)\in\Omega_{i,j}} (W_{s,t})$$
$$W_{\min} = \min_{(s,t)\in\Omega_{i,j}} (W_{s,t}).$$

where the location pair (s, t) satisfies

$$\Omega_{i,j} = \{(s, t) \mid s \in [i-1, i+1], \ t \in [j-1, j+1], (s, t) \neq (i, j)\}.$$

The max and min operators are independently used to mark the maximum and minimum values from the locally estimated parameters at each pixel location. The exponent n is used to adjust the significance of the difference $E_{i,j}$. In this stage, we select $n = 2$ to measure the energy difference (i.e. the mean-squared error) across the parameters.

Inspired by the definition of the famous Michelson contrast [54], we define a second contrast-based measure at the location of $S_{i,j}$:

$$C_{i,j} = \frac{(W_{\max} - W_{\min})^2}{W_{\max}^2 + W_{\min}^2}. \qquad (12)$$

It has been found in [18] that the block-based pooling is an effective way for sharpness evaluation. We further modify E and C into a block-based version:

$$E_{u,v}^{bb} = \frac{1}{M} \sqrt{\sum_{(i,j)\in\Phi_{u,v}} E_{i,j}} \qquad (13)$$

$$C_{u,v}^{bb} = \frac{1}{M} \sqrt{\sum_{(i,j)\in\Phi_{u,v}} C_{i,j}} \qquad (14)$$

where M is the length of the selected square patches. Each of chosen patches $\Phi_{u,v}$ is designated as

$$\Phi_{u,v} = \{(i, j) \mid i \in [(u-1)M+1, uM], \\ j \in [(v-1)M+1, vM]\}$$

where $1 \le u \le \lfloor H/M \rfloor$, $1 \le v \le \lfloor W/M \rfloor$, and W and H are the width and height of the image S, respectively.

It is worthy to stress that using the max and min operators before computing the energy- and contrast-differences is a simple tactic to reduce the dimensionality of AR parameters. Other complicated strategies, such as variance and entropy, are likely to be more effective.

D. Percentile Pooling Stage

In the final, a percentile pooling is taken to calculate the sharpness score. Percentile pooling methods have succeeded in improving the performance accuracy, such as [3] and [17]. As a result, we average the largest $Q_k\%$ values in the k ($k \in \{E, C, E^{bb}, C^{bb}\}$) map to compute the sharpness score ρ_k. We then derive the overall quality index with a linear weighted pooling of those four scores:

$$\rho = \sum_{k\in\Psi} \Theta_k \cdot \rho_k \qquad (15)$$

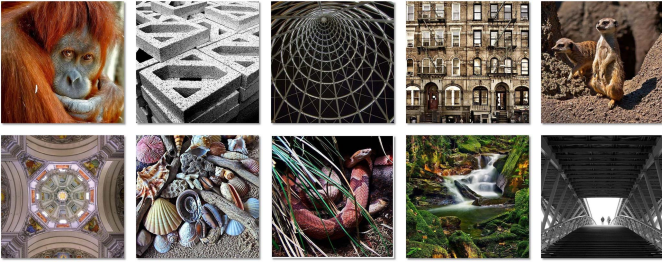where $\Psi = \{E, C, E^{bb}, C^{bb}\}$. Θ_k are positive constants used to adjust the relative importance of each component.

Fig. 4. The selected ten high-quality images from the CUHKPQ image database [55].

Fig. 5. The scatter plot of VIF versus ARISM on the 150 blurred images. The red curve is fitted with the logistic function of Eq. (18).

E. Determination of Parameters

To determine those parameters applied in ARISM, we first selected ten high-quality images, which are of a broad range of scenes (e.g. animals and architectures) from the CUHKPQ database [55] as shown in Fig. 4, and then created 150 blurred images using Gaussian kernels with standard deviation σ_G (from 1 to 5) with Matlab $fspecial$ and $imfilter$ commands. Each of R, G and B image planes was blurred with the same kernel. The CUHKPQ database was chosen for validating the generality and database-independency of our technique, since existing IQA databases [40]–[43] will be used for performance test and comparison in later experiments.

Next, we utilized the visual information fidelity (VIF) [2], which is quantified to be the ratio of the mutual information between the original and distorted images to the information content of the original one itself, owing to its superior performance in the image sharpness measure, to assess the aforesaid 150 images, and then use those objective quality scores to optimize the parameters adopted in ARISM. Spearman rank-order correlation coefficient (SRCC), one of the most popular performance metrics and has been used to find the suitable parameters in quite a few IQA approaches such as [9] and [10], is employed for optimization in this implementation.[3] As given in Fig. 5, we can see from the scatter plot of VIF versus our ARISM model that the sample points are quite clustered to the red fitted curve, with the SRCC value higher than 0.97 (1 is the best).

F. Extension to Color Metric

We further take chrominance information into consideration, as used in the literature [3], [4]. Before the calculation of AR parameters, the simple and widely used YIQ color space [52] is used to transfer an input RGB color image:

$$\begin{bmatrix} Y \\ I \\ Q \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ 0.596 & -0.274 & -0.322 \\ 0.211 & -0.523 & 0.312 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (16)$$

where Y conveys the luminance information, and I and Q contain the chrominance information. We thereby propose the

---

[3]Our ARISM model only applies E, C, and C^{bb} maps (i.e. $\Theta_{E^{bb}} = 0$) since the use of E^{bb} map cannot introduce the performance improvement. The Matlab code of the proposed sharpness metric will be available online at http://sites.google.com/site/guke198701/home.
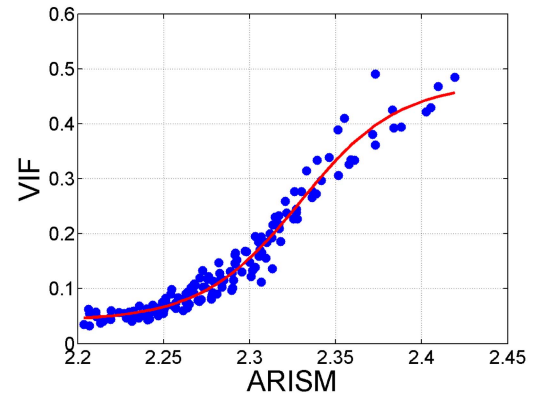
$\mathrm{ARISM_c}$ by extending ARISM to the YIQ space:

$$\rho_c = \sum_{l \in \{Y, I, Q\}} \Delta_l \cdot \rho_l \quad (17)$$

where Δ_l are fixed positive numbers for altering the relative importance of each component, which are optimized with the same method in Section III-E.

IV. EXPERIMENTAL RESULT

In this section we first provide an example of the application of our algorithm using an original natural image "monument" in Fig. 6. We first chose different Gaussian kernels $G(x, y, \sigma)$ with eleven standard deviations σ from 0.5 to 1.5 with an interval of 0.1. Then, eleven blurred images were generated by convolving the original version with each of the selected Gaussian kernels above. Based on the proposed ARISM, we evaluated the sharpness of these eleven blurred images and obtained their quality scores. The sample points of eleven standard deviations versus their corresponding ARISM scores are shown to be very convergent to the red fitted curve in the rightmost scatter plot.

We then calculate and compare the performance of our ARISM model with a large set of relevant methods on blur data sets. First, we used blur image subsets from four large-size LIVE, TID2008, CSIQ and TID2013 databases as testing beds. The most popular LIVE database [40] was developed at the University of Texas at Austin, including 779 lossy images created from 29 pristine ones by corrupting them with five types of distortions. We adopted 145 blurred images and their realigned DMOS (the differential version of MOS) values because realigned DMOSs are more reasonable than original ones [56]. The TID2008 database [41] was provided with a joint international effort between Finland, Italy and Ukraine, which consists of 1,700 images. These images were produced by corrupting 25 original versions with 17 distortion types at 4 different levels. A total of 100 blurred images were applied here. The CSIQ database [42] was released at Oklahoma State University, where 866 images were derived from 30 original counterparts. Six distortion types were considered in CSIQ: white noise, JPEG, JP2K, pink noise, blur, and global contrast decrements. We picked 150 blurred images from this database
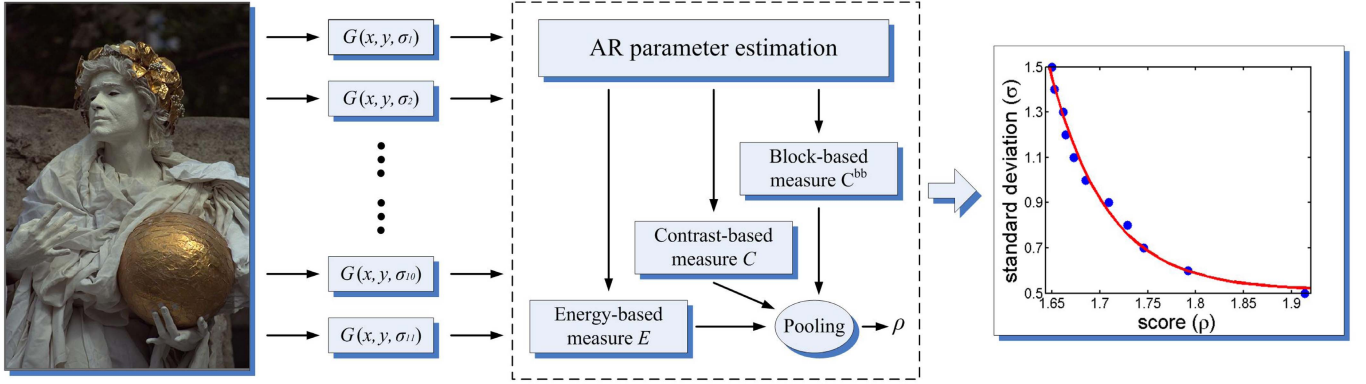
Fig. 6. A simple example of our ARISM model for an original image "monument". We first chose different Gaussian kernels $G(x, y, \sigma)$ with eleven standard deviations σ from 0.5 to 1.5 with an interval of 0.1. By convolving the original image with each of the selected Gaussian kernels above, eleven blurred images were generated. We then estimated the sharpness of these eleven blurred images with our sharpness measure, so as to acquire eleven quality scores. Finally, the rightmost scatter plot shows the well correlation of the eleven standard deviations versus their corresponding ARISM scores.

for testing. The TID2013 [43] contains totally 3,000 images, created by corrupting 25 original ones with 24 categories of distortions at 5 distinct levels. A number of 125 blurred images were used in this study.

Second, we choose fifteen classical FR IQA and state-of-the-art NR/blind algorithms for comparison. They are: 1) Three FR IQA models, peak signal-to-noise ratio (PSNR) that computes the signal energy preservation, structural similarity index (SSIM) that compares luminance, contrast and structural similarities [1], and VIF [2]; 2) Six NR IQA models, NFEQM [22], DIIVINE [24], BLIINDS-II [25], BRISQUE [26], NFSDM [29] and NFERM [30]; 3) Six blind sharpness/blurriness estimators, JNB [15], CPBD [16], S_3 [17], FISH [18], FISH$_{bb}$ [18], and LPC-SI [21]. Notice that the second type of general-purpose NR IQA models are trained on the LIVE database via the SVM, not only for the sharpness assessment.

Third, we refer to the suggestion given by the video quality experts group (VQEG) [57], and adopt a nonlinear mapping of the prediction results x to the subjective scores using the four-parameter logistic function:

$$f(x) = \frac{\xi_1 - \xi_2}{1 + \exp(-\frac{x - \xi_3}{\xi_4})} + \xi_2 \qquad (18)$$

where x and $f(x)$ stand for the input score and the mapped score. The free parameters ξ_j $(j = 1, 2, 3, 4)$ are determined during the curve fitting process. Next, four commonly used measures are employed to quantify the performance of those above metrics: 1) SRCC, which computes the monotonicity by ignoring the relative distance between the data:

$$\text{SRCC} = 1 - \frac{6 \sum_{i=1}^{F} d_i^2}{F(F^2 - 1)} \qquad (19)$$

where d_i is the difference between the i-th image's ranks in subjective and objective evaluations, and F represents the number of images in the testing database; 2) Kendall's rank-order correlation coefficient (KRCC), another monotonic-

ity metric used to measure the association between the inputs:

$$\text{KRCC} = \frac{F_c - F_d}{\frac{1}{2} F(F - 1)} \qquad (20)$$

where F_c and F_d separately indicate the numbers of concordant and discordant pairs in the testing data set; 3) Pearson linear correlation coefficient (PLCC), meaning the prediction accuracy:

$$\text{PLCC} = \frac{\sum_i \tilde{f}_i \cdot \tilde{o}_i}{\sqrt{\sum_i \tilde{f}_i^2 \cdot \sum_i \tilde{o}_i^2}} \qquad (21)$$

where $\tilde{o}_i = o_i - \bar{o}$ with o_i and $\bar{o}$ being the subjective scores of the i-th image and the mean of all o_i, and $\tilde{f}_i = f_i - \bar{f}$ with f_i and $\bar{f}$ being the converted objective scores after the nonlinear regression and the mean of all f_i; 4) root-mean-squared error (RMSE), quantifying the difference between f_i and o_i:

$$\text{RMSE} = \sqrt{\frac{1}{F} \sum (f_i - o_i)^2}. \qquad (22)$$

A good measure is expected to attain high values in SRCC, KRCC and PLCC, as well as low values in RMSE. In entire experiments, we merely include blurred images (i.e. original images are excluded).

Table I tabulates the performance measures on those four databases. For each evaluation criterion, we emphasize the top two performed NR/blind metrics with boldface. To provide a straightforward and overall comparison, Table I also computes the average SRCC, KRCC, PLCC and RMSE[4] results for each objective measure over all four databases. Two averages are used: 1) the direct average; 2) the database size-weighted average that computes the mean values based on the size of each data set (145 for LIVE, 100 for TID2008, 150 for CSIQ, and 125 for TID2013). The results of DIIVINE, BLIINDS-II, BRISQUE and NFSDM are not included for the LIVE

---

[4]RMSE is a measure highly related to the range of subjective ratings. Those four databases have different ranges, so the comparison on average should be conducted using all four databases and we do not include the RMSE values of the four training-based NR IQA metrics.

TABLE I

PERFORMANCE EVALUATIONS ON FOUR DATABASES AND TWO AVERAGES. WE BOLD THE TOP TWO PERFORMED NR/BLIND METRICS

LIVE Blur Database (145 Images) [40]				
Metrics	SRCC	KRCC	PLCC	RMSE
PSNR	0.7823	0.5847	0.7835	11.478
SSIM [1]	0.8944	0.7136	0.8743	8.9643
VIF [2]	0.9728	0.8594	0.9743	4.1572
NFEQM [22]	0.8845	0.7052	0.8921	8.3448
DIIVINE [24]	training images			
BLIINDS-II [25]	training images			
BRISQUE [26]	training images			
NFSDM [29]	training images			
NFERM [30]	training images			
JNB [15]	0.7871	0.6069	0.8160	10.677
CPBD [16]	0.9186	0.7634	0.8953	8.2263
S_3 [17]	0.9403	0.7893	0.9503	5.7494
FISH [18]	0.8808	0.7034	0.9043	7.8844
FISH$_{bb}$ [18]	0.9381	0.7898	0.9440	6.0954
LPC-SI [21]	0.9389	0.7785	0.9181	7.3217
ARISM (Pro.)	**0.9511**	**0.8042**	**0.9560**	**5.4176**
ARISM$_c$ (Pro.)	**0.9561**	**0.8161**	**0.9590**	**5.2359**

TID2008 Blur Database (100 Images) [41]				
Metrics	SRCC	KRCC	PLCC	RMSE
PSNR	0.8697	0.7332	0.8729	0.5726
SSIM [1]	0.9386	0.7862	0.9338	0.4200
VIF [2]	0.9540	0.8186	0.9402	0.3998
NFEQM [22]	0.7124	0.4716	0.7179	0.8169
DIIVINE [24]	0.7821	0.6389	0.7375	0.7925
BLIINDS-II [25]	0.8205	0.6245	0.8260	0.6614
BRISQUE [26]	0.7990	0.6233	0.8043	0.6973
NFSDM [29]	0.5534	0.3904	0.5808	0.9552
NFERM [30]	0.8075	0.6002	0.8050	0.6962
JNB [15]	0.6667	0.4951	0.6931	0.8459
CPBD [16]	0.8414	0.6301	0.8237	0.6654
S_3 [17]	0.8328	0.6184	0.8482	0.6216
FISH [18]	0.7828	0.5343	0.8079	0.6915
FISH$_{bb}$ [18]	0.8378	0.6221	0.8519	0.6145
LPC-SI [21]	**0.8561**	**0.6548**	**0.8574**	**0.6040**
ARISM (Pro.)	0.8505	0.6362	0.8428	0.6317
ARISM$_c$ (Pro.)	**0.8681**	**0.6750**	**0.8544**	**0.6098**

CSIQ Blur Database (150 Images) [42]				
Metrics	SRCC	KRCC	PLCC	RMSE
PSNR	0.9291	0.7543	0.9252	0.1087
SSIM [1]	0.9245	0.7665	0.9005	0.1246
VIF [2]	0.9745	0.8661	0.9737	0.0653
NFEQM [22]	0.8939	0.7250	0.9158	0.1151
DIIVINE [24]	0.8716	0.6886	0.8979	0.1262
BLIINDS-II [25]	0.8766	0.6788	0.8930	0.1290
BRISQUE [26]	0.9025	0.7350	0.9274	0.1072
NFSDM [29]	0.7398	0.5453	0.7832	0.1782
NFERM [30]	0.8964	0.7284	0.9218	0.1111
JNB [15]	0.7624	0.5976	0.8061	0.1696
CPBD [16]	0.8853	0.7090	0.8822	0.1349
S_3 [17]	0.8681	0.6868	0.8833	0.1343
FISH [18]	0.8941	0.7323	0.9232	0.1102
FISH$_{bb}$ [18]	0.9177	0.7606	0.9434	0.0950
LPC-SI [21]	0.9071	0.7205	0.9159	0.1151
ARISM (Pro.)	**0.9255**	**0.7581**	**0.9456**	**0.0933**
ARISM$_c$ (Pro.)	**0.9314**	**0.7695**	**0.9481**	**0.0911**

TID2013 Blur Database (125 Images) [43]				
Metrics	SRCC	KRCC	PLCC	RMSE
PSNR	0.9149	0.7884	0.9137	0.5071
SSIM [1]	0.9629	0.8385	0.9577	0.3592
VIF [2]	0.9650	0.8370	0.9530	0.3782
NFEQM [22]	0.7771	0.5347	0.8005	0.7479
DIIVINE [24]	0.8344	0.6180	0.8472	0.6629
BLIINDS-II [25]	0.8367	0.6519	0.8490	0.6593
BRISQUE [26]	0.8143	0.6359	0.8248	0.7057
NFSDM [29]	0.6155	0.4440	0.6538	0.9442
NFERM [30]	0.8498	0.6555	0.8493	0.6588
JNB [15]	0.6902	0.5137	0.7114	0.8770
CPBD [16]	0.8515	0.6462	0.8553	0.6466
S_3 [17]	0.8046	0.5871	0.8432	0.6708
FISH [18]	0.8024	0.5672	0.8327	0.6910
FISH$_{bb}$ [18]	0.8584	0.6462	0.8756	0.6027
LPC-SI [21]	0.8888	0.6839	0.8917	0.5647
ARISM (Pro.)	**0.8980**	**0.7149**	**0.8953**	**0.5560**
ARISM$_c$ (Pro.)	**0.9015**	**0.7174**	**0.8979**	**0.5493**

Direct Average				
Metrics	SRCC	KRCC	PLCC	RMSE
PSNR	0.8740	0.7152	0.8738	3.1667
SSIM [1]	0.9301	0.7762	0.9166	2.4670
VIF [2]	0.9666	0.8452	0.9603	1.2501
NFEQM [22]	0.8170	0.6091	0.8316	2.5062
DIIVINE [24]	0.8294	0.6485	0.8275	-
BLIINDS-II [25]	0.8446	0.6517	0.8560	-
BRISQUE [26]	0.8386	0.6647	0.8521	-
NFSDM [29]	0.6362	0.4599	0.6726	-
NFERM [30]	0.8512	0.6614	0.8587	-
JNB [15]	0.7266	0.5533	0.7567	3.1423
CPBD [16]	0.8742	0.6872	0.8641	2.4183
S_3 [17]	0.8614	0.6704	0.8813	1.7940
FISH [18]	0.8400	0.6343	0.8670	2.3443
FISH$_{bb}$ [18]	0.8880	0.7047	0.9037	1.8519
LPC-SI [21]	0.8977	0.7048	0.8958	2.1514
ARISM (Pro.)	**0.9063**	**0.7330**	**0.9099**	**1.6746**
ARISM$_c$ (Pro.)	**0.9143**	**0.7445**	**0.9148**	**1.6215**

Database Size-Weighted Average				
Metrics	SRCC	KRCC	PLCC	RMSE
PSNR	0.8733	0.7112	0.8729	3.4641
SSIM [1]	0.9281	0.7729	0.9133	2.7027
VIF [2]	0.9678	0.8481	0.9624	1.3459
NFEQM [22]	0.8283	0.6250	0.8434	2.6970
DIIVINE [24]	0.8354	0.6518	0.8382	-
BLIINDS-II [25]	0.8483	0.6554	0.8605	-
BRISQUE [26]	0.8455	0.6722	0.8603	-
NFSDM [29]	0.6487	0.4703	0.6861	-
NFERM [30]	0.8572	0.6699	0.8665	-
JNB [15]	0.7336	0.5603	0.7644	3.3996
CPBD [16]	0.8780	0.6939	0.8681	2.6162
S_3 [17]	0.8662	0.6783	0.8856	1.9227
FISH [18]	0.8469	0.6465	0.8740	2.5294
FISH$_{bb}$ [18]	0.8938	0.7146	0.9097	1.9901
LPC-SI [21]	0.9018	0.7117	0.8994	2.3267
ARISM (Pro.)	**0.9116**	**0.7407**	**0.9166**	**1.7927**
ARISM$_c$ (Pro.)	**0.9189**	**0.7518**	**0.9210**	**1.7356**

database because all of them use that database for training. As a consequence, their average results are calculated over the other three databases only.

We can observe that the proposed ARISM model correlates highly with human visual perception to image sharpness, and it is remarkably superior to those testing NR/blind techniques on average. In general, FR IQA metrics are considered hardly matchable with NR/blind approaches owing to the existence of original references. Although this comparison is unfair to ARISM, our metric is still better than the FR PSNR, while is a little inferior to the FR SSIM on average.

Moreover, we should mention that the average performance improvement (using SRCC) of the proposed ARISM is larger than 1.9% relative to the second best LPC-SI algorithm. As compared to the previous NFEQM method, our technique has achieved noticeable performance gain, about 8.1% on
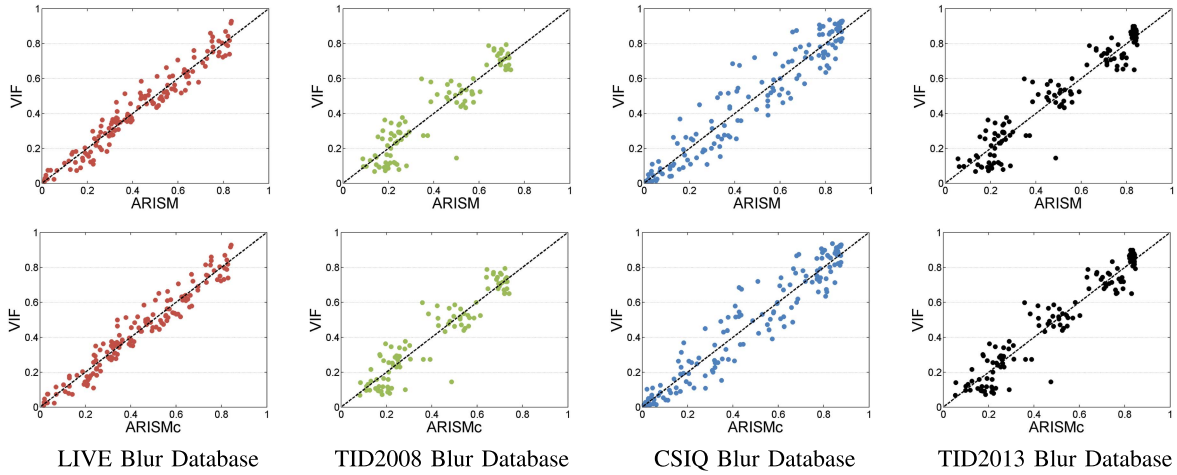
Fig. 7. Scatter plots of FR VIF versus our ARISM/ARISM$_c$ on LIVE, TID2008, CSIQ and TID2013 blur subsets.

TABLE II
SRCC AND PLCC COMPARISON BETWEEN FR VIF AND THE PROPOSED ARISM/ARISM$_c$ ON FOUR BLUR SUBSETS

ARISM / VIF	LIVE	TID2008	CSIQ	TID2013
SRCC	0.9753	0.8911	0.9512	0.9318
PLCC	0.9724	0.9267	0.9573	0.9529

ARISM$_c$ / VIF	LIVE	TID2008	CSIQ	TID2013
SRCC	0.9743	0.9005	0.9557	0.9364
PLCC	0.9714	0.9338	0.9606	0.9571

TABLE III
PERFORMANCE MEASURES OF OUR ARISM, ARISM-S, ARISM$_c$, AND ARISM$_c$-S MODELS ON FOUR BLUR SUBSETS

Models	LIVE		TID2008		CSIQ		TID2013		Direct Average		Weighted Average	
	SRCC	PLCC	SRCC	PLCC	SRCC	PLCC	SRCC	PLCC	SRCC	PLCC	SRCC	PLCC
ARISM	0.9511	0.9560	0.8505	0.8428	0.9255	0.9456	0.8980	0.8953	0.9063	0.9099	0.9116	0.9166
ARISM-S	0.9522	0.9570	0.8519	0.8419	0.9257	0.9460	0.8995	0.8946	0.9073	0.9099	0.9126	0.9167
ARISM$_c$	0.9561	0.9590	0.8681	0.8544	0.9314	0.9481	0.9015	0.8979	0.9143	0.9148	0.9189	0.9210
ARISM$_c$-S	0.9576	0.9606	0.8722	0.8561	0.9306	0.9477	0.9038	0.8989	0.9161	0.9158	0.9205	0.9219

LIVE, 21.9% on TID2008, 4.2% on CSIQ, 16.0% on TID2013, 11.9% on the direct average, and 10.9% on the database size-weighted average. This also demonstrates the superiority of the proposed scheme used in ARISM over that used in NFEQM for the sharpness measure.

To confirm the proposed ARISM/ARISM$_c$, we also test how well it predicts FR VIF, which is of substantially high accuracy in assessing blurred images. We in Fig. 7 illustrate the scatter plots acquired using all four data sets, where each sample point indicates one test image and the vertical and horizontal axes correspond to FR VIF and ARISM/ARISM$_c$, respectively. The points lie on the black diagonal dash line for the perfect prediction. To provide a quantitative comparison, Table II lists SRCC and PLCC values between VIF and our metric on each data set. It can be viewed that the points are scattered fairly close to the black diagonal lines in Fig. 7 and correlation performance results are almost above 0.9, meaning the good prediction performance of our technique. It should be noted that high correlation between the proposed model and FR VIF or subjective opinion scores strongly suggests the effectiveness of our hypothesis that image blurring increases the resemblance of locally estimated AR parameters.

We further notice that using max and min operators and two classical energy and contrast measures, or in other words a simple dimensionality reduction method, is just an easy and empirical tactic to analyze the AR parameters. It is obvious that other measures (e.g. variance and entropy) or machine learning-based technologies (e.g. principal component analysis and popular deep learning network) might be of more superior performance in the image sharpness estimation.

In addition, we also employ the DMOSs of blurry images in the LIVE database to optimize the parameters used in the proposed method, dubbed as ARISM-S and ARISM$_c$-S, since the objective quality metric leans to the mean human scores. The performance results of the proposed ARISM, ARISM-S, ARISM$_c$, and ARISM$_c$-S models are compared using LIVE, TID2008, CSIQ, TID2013 databases and the two means, as reported in Table III. In contrast to ARISM and ARISM$_c$ that are optimized using the high-accuracy VIF metric, ARISM-S and ARISM$_c$-S built upon subjective scores perform better on most of testing databases and two averages.

Statistical significance analysis based on the variance-based hypothesis testing shows additional information regarding the relative performance of different quality algorithms [56]. The hypothesis behind such analysis is that the residual difference between the subjective score and its objective prediction is Gaussian distributed. In reality, this assumption is not always met perfectly, whereas is somewhat reasonable because the Central Limit Theorem comes into play and the distribution of the residual difference approximates the Gaussian distri-

TABLE IV

STATISTICAL SIGNIFICANCE COMPARISON OF ARISM/ARISM$_c$/ARISM-3 AND TESTING MODELS WITH F-TEST

Algorithms	ARISM				ARISM$_c$				ARISM-3			
	LIVE	TID2008	CSIQ	TID2013	LIVE	TID2008	CSIQ	TID2013	LIVE	TID2008	CSIQ	TID2013
NFEQM [22]	+1	+1	+1	+1	+1	+1	+1	+1	+1	+1	+1	+1
DIIVINE [24]	-	+1	+1	+1	-	+1	+1	+1	-	+1	+1	+1
BLIINDS-II [25]	-	+1	+1	+1	-	+1	+1	+1	-	+1	+1	+1
BRISQUE [26]	-	+1	+1	+1	-	+1	+1	+1	-	+1	+1	+1
NFSDM [29]	-	+1	+1	+1	-	+1	+1	+1	-	+1	+1	+1
NFERM [30]	-	+1	+1	+1	-	+1	+1	+1	-	+1	+1	+1
JNB [15]	+1	+1	+1	+1	+1	+1	+1	+1	+1	+1	+1	+1
CPBD [16]	+1	+1	+1	+1	+1	+1	+1	+1	+1	+1	+1	+1
S$_3$ [17]	0	0	+1	+1	0	0	+1	+1	0	0	+1	+1
FISH [18]	+1	+1	+1	+1	+1	+1	+1	+1	+1	+1	+1	+1
FISH$_{bb}$ [18]	+1	0	+1	+1	+1	0	+1	+1	0	0	+1	+1
LPC-SI [21]	+1	0	+1	0	+1	0	+1	0	+1	0	+1	0

TABLE V

PERFORMANCE AND TIME COMPARISON OF ARISM-si AS WELL AS LPC-SI AND S$_3$ APPROACHES

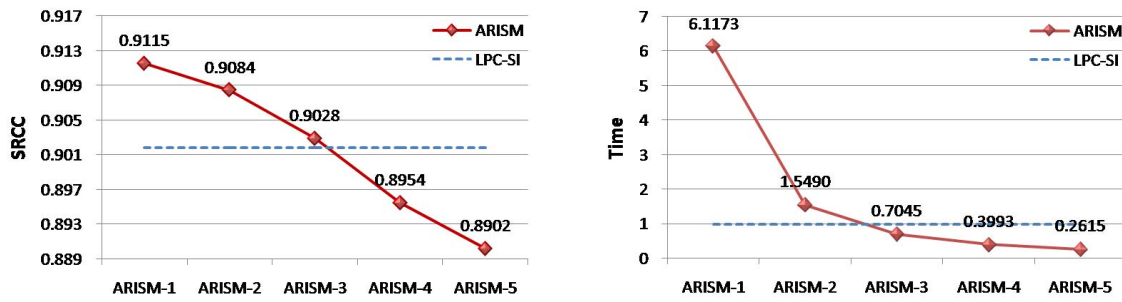| Algorithm	ARISM-1	ARISM-2	ARISM-3	ARISM-4	ARISM-5	S$_3$	LPC-SI
SRCC	0.9115	0.9084	**0.9028**	0.8954	0.8902	0.8662	0.9018
PLCC	0.9158	0.9138	**0.9045**	0.8957	0.8922	0.8856	0.8994
Time (second / image)	6.1173	1.5490	**0.7045**	0.3993	0.2615	9.1518	0.9780

Fig. 8. Plots of performance and computational time (second/image) of our ARISM-si and LPC-SI approaches.

bution with the large number of sample points. For a given image database, the F-test is applied to compare the variances of two sets of prediction residuals by two objective methods, in order to determine whether the two sample sets are of the same distribution. As such, we can make a statistically sound judgment regarding superiority or inferiority of one objective method against another. Results of statistical significance are listed in Table IV. A symbol "0" denotes that the two objective methods are statistically indistinguishable, "+1" denotes our method is statistically better than that of the column, and "−1" denotes that our method is statistically worse than that of the column. A symbol "-" denotes the unfeasible analysis since learning-based DIIVINE, BLIINDS-II, BRISQUE and NFSDM are trained on LIVE. It is found that our model is statistically indistinguishable from S$_3$ for LIVE and TID2008, from FISH$_{bb}$ for TID2008, from LPC-SI for TID2008 and TID2013, and better than all other blind algorithms.

Furthermore, we compare the effectiveness and efficiency with the top two blind sharpness metrics (LPC-SI and S$_3$). Clearly, it requires much computational cost in the estimation of AR parameters, and thus the proposed ARISM model needs a great amount of time[5] that is the average value using

100 blurred images of the same size 512×384 in the TID2008 database with a computer of Intel Core i7 CPU at 3.40 GHz, as provided in Table V. But we notice that the neighboring AR coefficients are highly similar, and hence we choose the sampling method for computational time reduction. That is to say, the AR model parameters are evaluated once every a few pixels in both horizontal and vertical directions. In this way, the computational quantity can be largely reduced to about $\frac{1}{si^2}$, where si means the value of the sampling interval. Here we calculate the database size-weighted average performance on all four databases (using SRCC, PLCC) and the computational cost of five ARISM-si ($si = 1, 2, 3, 4, 5$), and report those results in Table V. A better metric is expected to take less time. It can be readily viewed that the high efficiency is attained via a bit of loss in the prediction accuracy. Therefore we can flexibly select a proper ARISM-si model for the effectiveness- or efficiency-dominant environment.

Apart from the self-comparison, we find a good compromise (ARISM-3) between effectiveness and efficiency, in contrast to the top two performed LPC-SI and S$_3$ metrics, whose results are reported in Table V. Because the ARISM-3 is an effective and efficient metric, we compute the F-test of ARISM-3 with other NR/blind algorithms, as tabulated in Table IV. Results tell that ARISM-3 is statistically indistinguishable from S$_3$ and FISH$_{bb}$ for LIVE and TID2008, from LPC-SI for TID2008 and TID2013, and superior to

---

[5]Due to the limited performance gain yet much extra computational load of the color information, we hereinafter do not consider the use of color space.
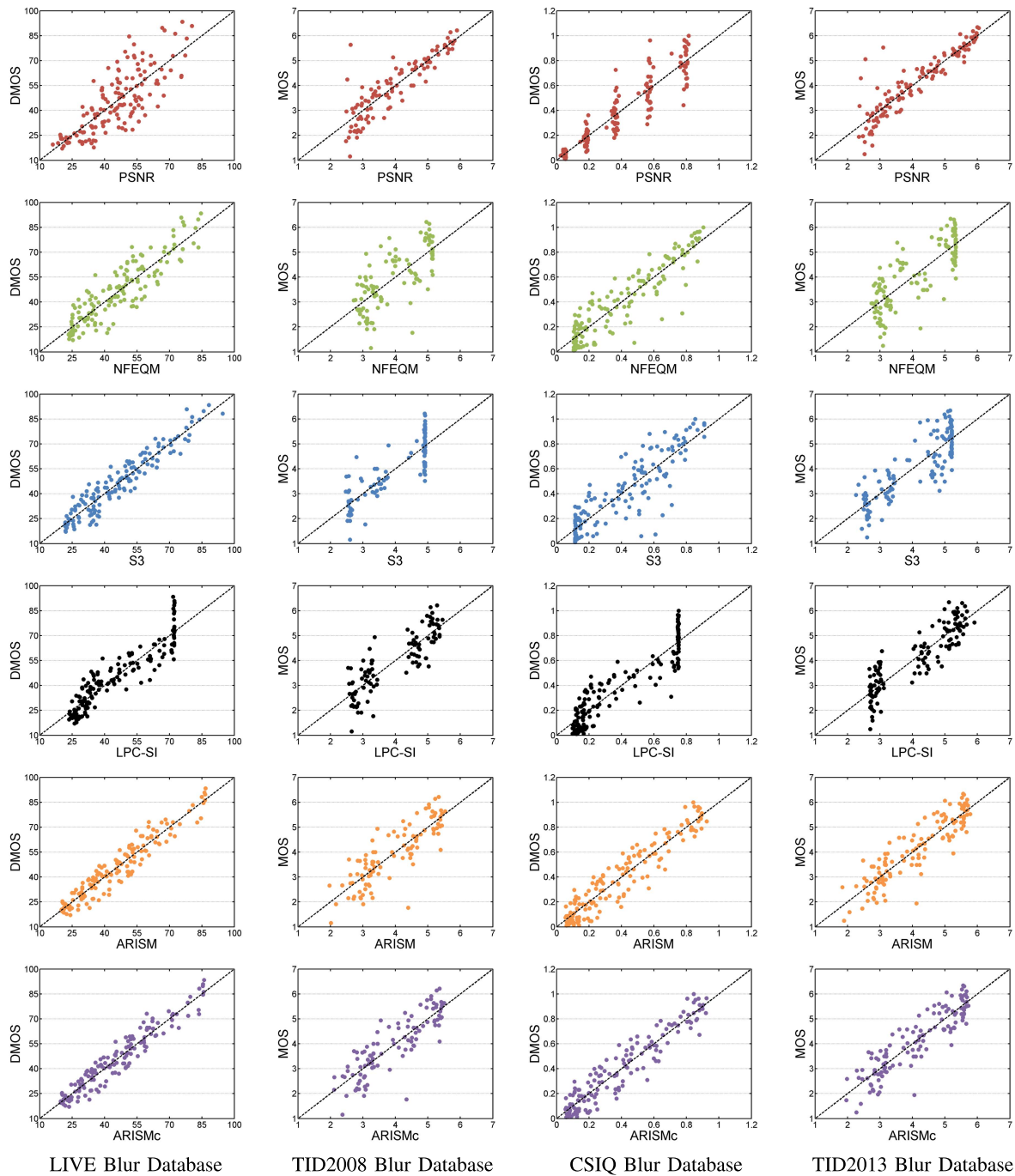
Fig. 9. Scatter plots of MOS/DMOS versus FR PSNR, NR/blind NFEQM and S₃, as well as our ARISM and ARISM$_c$ models (after the nonlinear regression) on four blur image's subsets from LIVE, TID2008, CSIQ and TID2013 databases.

all other models for all databases. For a clear show, Fig. 8 exhibits two plots of performance and computational time of those five ARISM-*si* models and the LPC-SI metric.

Finally, we show the scatter plots of MOS/DMOS versus objective quality predictions of representative FR PSNR, NR/blind NFEQM, S₃, and the proposed ARISM, ARISM$_c$ metrics (after the nonlinear mapping) on all four databases in Fig. 9. Our technique generally provides reasonable quality measures, where the sample points tend to be clustered closer to the black diagonal lines (meaning perfect prediction) than other testing methods under comparison.

V. EXTENSION TO SHARPNESS ASSESSMENT OF STEREOSCOPIC IMAGES

The 3D imaging technology is nowadays greatly important, because the number of digital 3D pictures and movies for human consumption has dramatically increased over the recent years. Therefore, how to validly monitor, control and improve the visual quality of stereoscopic images become an urgent problem and thus accurate stereoscopic IQA methods are highly desirable. One type of classical schemes is to integrate the 2D IQA measures of the left- and right-views with/without the quality of the disparity map to yield the final

TABLE VI

COMPARISON ON LIVE3D-I AND LIVE3D-II. THE TOP TWO ALGORITHMS ARE HIGHLIGHTED BY BOLD FONT

Metrics	Type	LIVE3D-I Blur Subset [44]			LIVE3D-II Blur Subset [45]			Average		
		SRCC	PLCC	RMSE	SRCC	PLCC	RMSE	SRCC	PLCC	RMSE
PSNR	FR	**0.902**	0.916	5.820	0.852	0.928	5.194	0.877	0.922	5.507
SSIM [1]	FR	0.879	0.918	5.752	0.836	0.842	7.518	0.858	0.880	6.635
You [34]	FR	0.882	0.919	5.679	0.813	0.784	8.649	0.848	0.852	7.164
Hewage [33]	RR	0.690	0.798	8.748	0.028	0.450	12.44	0.359	0.624	10.59
Akhter [32]	NR	0.555	0.617	11.39	0.682	0.795	8.450	0.619	0.706	9.920
BRISQUE [26]	NR	0.860	**0.926**	**5.473**	0.862	0.862	**4.323**	0.861	0.894	**4.898**
Chen [46]	NR	0.878	0.917	5.898	**0.900**	**0.941**	4.725	**0.889**	**0.929**	5.312
SARISM (Pro.)	NR	**0.903**	**0.932**	**5.230**	0.877	**0.960**	**3.917**	**0.890**	**0.946**	**4.573**

TABLE VII

STATISTICAL SIGNIFICANCE COMPARISON BETWEEN OUR SARISM AND TESTING STEREOSCOPIC IQA METRICS

SARISM	PSNR	SSIM [1]	You [34]	Hewage [33]	Akhter [32]	BRISQUE [26]	Chen [46]
LIVE3D-I Database	+1	+1	+1	+1	+1	0	+1
LIVE3D-II Database	+1	+1	+1	+1	+1	+1	0

quality prediction of the 3D image [32]–[36]. Using a similar yet more reasonable and effective strategy, in this paper we further extend the proposed model to the sharpness assessment of stereoscopic images with a few small modifications.

Early researches presented somewhat conflict observations and opinions concerning the effect of asymmetric distortions. In [58], evidence shows that the quality of asymmetric blurred images is heavily dominated by the higher-quality view. So the key point is how to well fuse the quality scores of the left- and right-eye images. Several existing studies on binocular rivalry [37]–[39] tell that for simple ideal stimuli, a growing contrast increases the predominance of one view against the other. In general, the contrast of a visual stimulus having complicated scenes increases with the difference of AR parameters, which motivates a sound hypothesis that the level of view dominance in binocular rivalry of 3D images is rising with the difference of AR model coefficients in two views.

To specify, given an input blurry image pair S_L and S_R, the energy- and contrast-based maps E and C as well as the associated block-based versions E^{bb} and C^{bb} are computed by using Eqs. (11)-(14). We can obtain their quality scores Y_L and Y_R in the luminance component with the proposed percentile pooling stage in Eq. (15). Notice that the higher values of E and C, the larger difference of AR parameters. Based on the assumption that the view dominance of stereoscopic images improves with the difference of AR coefficients and thus with the energy- and contrast-differences, a straightforward method is to integrate E and C to compute the weights of the view dominance as follows:

$$V_{Y_L} = M_{Y_L,E}^{\alpha_Y} + M_{Y_L,C}^{\beta_Y} \tag{23}$$

$$V_{Y_R} = M_{Y_R,E}^{\alpha_Y} + M_{Y_R,C}^{\beta_Y} \tag{24}$$

where $M_{Y_L,E}$ and $M_{Y_L,C}$ are the mean of E and C of the left-eye image, while $M_{Y_R,E}$ and $M_{Y_R,C}$ are the mean of E and C of the right-eye one; α_Y and β_Y are positive weights for adjusting the relative importance of energy- and contrast-measures. Then the 3D image quality score in the luminance

component can be expressed by

$$Q_Y = V_{Y_L} \cdot Y_L + V_{Y_R} \cdot Y_R. \tag{25}$$

Using Eqs. (16)-(17), we similarly compute the scores in two chrominance components, and finally infer the overall quality of the stereoscopic image to be

$$Q_S = \sum_{l \in \{Y,I,Q\}} \Delta_l \cdot Q_l \tag{26}$$

where Δ_l are fixed positive weights. It needs to emphasize that, except the newly introduced parameters α_l and β_l, other parameters are the same with those used in our proposed 2D image sharpness metric.

Two popular stereoscopic image databases (LIVE3D-I [44] and LIVE3D-II [45]) are adopted in this work for the performance measure. LIVE3D-I includes 20 reference stereoscopic images and the associated 365 distorted stereoscopic pairs. Five types of distortions, including JPEG, JPEG2000, blur, noise and fastfading, are symmetrically exerted on the original left- and right-views at different levels. LIVE3D-II consists of 120 symmetrically distorted stereoscopic pairs and 240 asymmetrically pairs generated from 8 source pairs. The same five distortion types are symmetrically and asymmetrically applied to the reference left- and right-eye images at various degradation levels. In this work we consider the blur image sets in the aforesaid two databases. Three performance indices (SRCC, PLCC and RMSE) are used to quantify the correlation performance of the proposed SARISM model. As listed in Table VI, we can observe from the results that our approach has attained fairly high performance accuracy.

A comparison of our model with seven competitive quality metrics, including FR PSNR, SSIM [1], You [34], RR Hewage [33], and NR/blind Akhter [32], BRISQUE [26], Chen [45] is given in Table VI. In the LIVE3D-I database which is composed of symmetrically stereoscopic image pairs, our SARISM has obtained the top performance, and this is also due to the superiority of the proposed 2D image sharpness metric. In the LIVE3D-II database consisting of asymmetrically stereoscopic image pairs, there exist substantial differences across various methods in the correlation
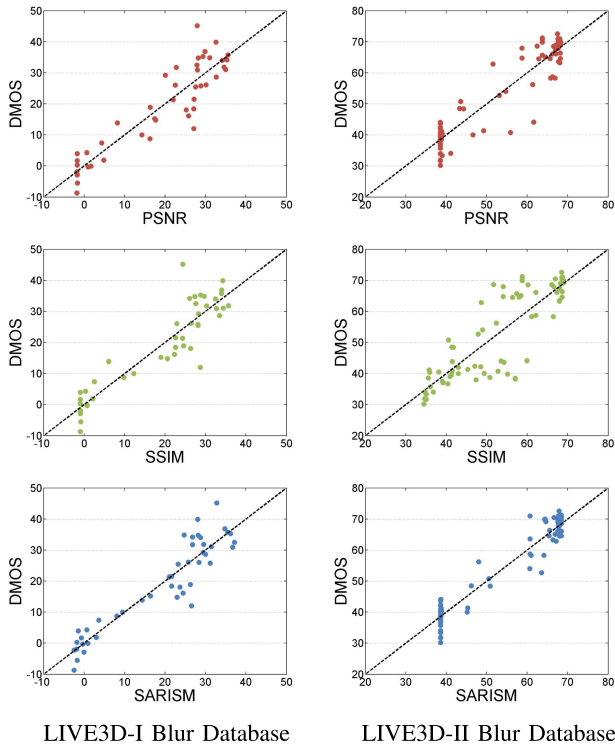
Fig. 10. Scatter plots of DMOS versus FR PSNR, SSIM and the proposed SARISM metrics (after the nonlinear regression) on blur subsets from LIVE3D-I and LIVE3D-II databases.

performance with human opinions. First, without access to the original 3D image, our blind sharpness measure is superior to the four testing FR and RR-IQA metrics that need the help of reference information for predicting visual quality. Second, between the two training-free NR algorithms, the proposed SARISM is remarkably better than the Akhter model. Third, in comparison to the two training-based BRISQUE and Chen methods, our metric outperforms BRISQUE while is a little inferior to Chen in the measure of monotonicity but superior to it in the measure of prediction accuracy. This phenomenon is possibly because the Chen method has a complicated binocular rivalry model that encompasses a SSIM-based stereo algorithm for estimating disparity map and a set of multi-scale Gabor filters. In contrast, our SARISM only uses some intermediate results as weights to combine the sharpness measures of the left- and right-views with basic matrix operations. The average performance indices are also shown in Table VI, which confirms the effectiveness of the proposed metric over all the tested algorithms.

The F-test is further applied to the statistical significance comparison of the proposed SARISM and testing metrics, as listed in Table VII. Although it is statistically equivalent to BRISQUE on LIVE3D-I and to Chen on LIVE3D-II, overall, the proposed method is statistically better than all approaches considered. Fig. 10 further illustrates a visualized comparison of the scatter plots between DMOS versus PSNR, SSIM and our SARISM model on LIVE3D-I and LIVE3D-II databases. The proposed technique generally presents reasonable quality predictions, where the sample points tend to be clustered closer to the black diagonal lines (meaning perfect prediction) as compared to other metrics under comparison.

VI. CONCLUSION

In this paper, we have proposed a new simple yet effective blind sharpness measure via parameter analysis of classical autoregressive (AR) image model. Our method is established upon the assumption that higher resemblance of the locally estimated AR model coefficients means lower sharpness. We further extend ARISM to the simple and widely used YIQ color space, and thus introduce ARISM$_c$ taking into account the color effect on the image sharpness assessment. Results of experiments conducted on blur data sets from four large-size monoscopic image databases have demonstrated that the proposed ARISM and ARISM$_c$ enjoy superior performance relative to mainstream NR IQA metrics, state-of-the-art blind sharpness/blurriness evaluators, and FR quality evaluations. We also extend the proposed model to the sharpness assessment of stereoscope images with a few small modifications. In contrast to related popular quality methods, our developed stereoscopic sharpness measure performs effectively on two recently released 3D image databases.

Furthermore, we want to highlight two points: 1) this paper explores a new framework based on the free energy principle and the AR model, introducing remarkable performance gain with respective to the previous NFEQM metric; 2) we only design a simple and empirical scheme via the analysis of AR model parameters while other advanced technologies based on machine learning will be researched in the future.

REFERENCES

[1] 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.

[2] H. R. Sheikh and A. C. Bovik, "Image information and visual quality," *IEEE Trans. Image Process.*, vol. 15, no. 2, pp. 430–444, Feb. 2006.

[3] 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.*, Oct. 2014, pp. 506–510.

[4] L. Zhang, L. Zhang, X. Mou, and D. Zhang, "FSIM: A feature similarity index for image quality assessment," *IEEE Trans. Image Process.*, vol. 20, no. 8, pp. 2378–2386, Aug. 2011.

[5] A. Liu, W. Lin, and M. Narwaria, "Image quality assessment based on gradient similarity," *IEEE Trans. Image Process.*, vol. 21, no. 4, pp. 1500–1512, Apr. 2012.

[6] J. Wu, W. Lin, G. Shi, and A. Liu, "Perceptual quality metric with internal generative mechanism," *IEEE Trans. Image Process.*, vol. 22, no. 1, pp. 43–54, Jan. 2013.

[7] D. Zoran and Y. Weiss, "Scale invariance and noise in natural images," in *Proc. IEEE 12th Int. Conf. Comput. Vis.*, Sep./Oct. 2009, pp. 2209–2216.

[8] W. Liu and W. Lin, "Additive white Gaussian noise level estimation in SVD domain for images," *IEEE Trans. Image Process.*, vol. 22, no. 3, pp. 872–883, Mar. 2013.

[9] 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.*, to be published.

[10] K. Gu, G. Zhai, W. Lin, and M. Liu, "The analysis of image contrast: From quality assessment to automatic enhancement," *IEEE Trans. Cybern.*, to be published.

[11] J. Sun, Z. Xu, and H.-Y. Shum, "Gradient profile prior and its applications in image super-resolution and enhancement," *IEEE Trans. Image Process.*, vol. 20, no. 6, pp. 1529–1542, Jun. 2011.

[12] X. Zhu and P. Milanfar, "Automatic parameter selection for denoising algorithms using a no-reference measure of image content," *IEEE Trans. Image Process.*, vol. 19, no. 12, pp. 3116–3132, Dec. 2010.

[13] P. Marziliano, F. Dufaux, S. Winkler, and T. Ebrahimi, "A no-reference perceptual blur metric," in *Proc. IEEE Int. Conf. Image Process.*, vol. 3. Sep. 2002, pp. III-57–III-60.

[14] S. Wu, W. Lin, S. Xie, Z. Lu, E. P. 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.

[15] R. Ferzli and L. J. 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.

[16] 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, Sep. 2011.

[17] C. T. Vu, T. D. Phan, and D. M. Chandler, "S_3: 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.

[18] P. V. Vu and D. M. Chandler, "A fast wavelet-based algorithm for global and local image sharpness estimation," *IEEE Signal Process. Lett.*, vol. 19, no. 7, pp. 423–426, Jul. 2012.

[19] C. Feichtenhofer, H. Fassold, and P. Schallauer, "A perceptual image sharpness metric based on local edge gradient analysis," *IEEE Signal Process. Lett.*, vol. 20, no. 4, pp. 379–382, Apr. 2013.

[20] Z. Wang and E. P. Simoncelli, "Local phase coherence and the perception of blur," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 16. May 2004, pp. 1–8.

[21] R. Hassen, Z. Wang, and M. M. A. Salama, "Image sharpness assessment based on local phase coherence," *IEEE Trans. Image Process.*, vol. 22, no. 7, pp. 2798–2810, Jul. 2013.

[22] G. Zhai, X. Wu, X. Yang, W. Lin, and W. Zhang, "A psychovisual quality metric in free-energy principle," *IEEE Trans. Image Process.*, vol. 21, no. 1, pp. 41–52, Jan. 2012.

[23] K. Friston, "The free-energy principle: A unified brain theory?" *Nature Rev. Neurosci.*, vol. 11, pp. 127–138, Feb. 2010.

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

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

[26] 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.

[27] E. P. Simoncelli and B. A. Olshausen, "Natural image statistics and neural representation," *Annu. Rev. Neurosci.*, vol. 24, no. 1, pp. 1193–1216, Mar. 2001.

[28] C.-C. Chang and C.-J. Lin, "LIBSVM: A library for support vector machines," *ACM Trans. Intell. Syst. Technol.*, vol. 2, no. 3, 2011, Art. ID 27. [Online]. Available: http://www.csie.ntu.edu.tw/~cjlin/libsvm

[29] K. Gu, G. Zhai, X. Yang, W. Zhang, and L. Liang, "No-reference image quality assessment metric by combining free energy theory and structural degradation model," in *Proc. IEEE Int. Conf. Multimedia Expo*, Jul. 2013, pp. 1–6.

[30] 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.

[31] K. Gu, G. Zhai, X. Yang, and W. Zhang, "A new reduced-reference image quality assessment using structural degradation model," in *Proc. IEEE Int. Symp. Circuits Syst.*, May 2013, pp. 1095–1098.

[32] R. Akhter, J. Baltes, Z. M. P. Sazzad, and Y. Horita, "No-reference stereoscopic image quality assessment," *Proc. SPIE*, vol. 7524, p. 75240T, Feb. 2010.

[33] C. T. E. R. Hewage and M. G. Martini, "Reduced-reference quality metric for 3D depth map transmission," in *Proc. 3DTV-Conf., True Vis.-Capture, Transmiss., Display 3D Video*, Jun. 2010, pp. 1–4.

[34] J. You, L. Xing, A. Perkis, and X. Wang, "Perceptual quality assessment for stereoscopic images based on 2D image quality metrics and disparity analysis," in *Proc. Int. Workshop Video Process. Quality Metrics Consum. Electron.*, 2010, pp. 1–6.

[35] F. Shao, W. Lin, S. Gu, G. Jiang, and T. Srikanthan, "Perceptual full-reference quality assessment of stereoscopic images by considering binocular visual characteristics," *IEEE Trans. Image Process.*, vol. 22, no. 5, pp. 1940–1953, May 2013.

[36] F. Shao, W. Lin, S. Wang, G. Jiang, and M. Yu, "Blind image quality assessment for stereoscopic images using binocular guided quality lookup and visual codebook," *IEEE Trans. Broadcast.*, to be published.

[37] W. J. M. Levelt, "The alternation process in binocular rivalry," *Brit. J. Psychol.*, vol. 57, nos. 3–4, pp. 225–238, 1966.

[38] R. Blake, "Threshold conditions for binocular rivalry," *J. Experim. Psychol., Human Perception Perform.*, vol. 3, no. 2, pp. 251–257, May 1977.

[39] J. Wang, K. Zeng, and Z. Wang, "Quality prediction of asymmetrically distorted stereoscopic images from single views," in *Proc. IEEE Int. Conf. Multimedia Expo*, Jul. 2014, pp. 1–6.

[40] 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, accessed 2006.

[41] N. Ponomarenko, V. Lukin, A. Zelensky, K. Egiazarian, M. Carli, and F. Battisti, "TID2008—A database for evaluation of full-reference visual quality assessment metrics," *Adv. Modern Radioelectron.*, vol. 10, pp. 30–45, 2009.

[42] 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, p. 011006, Mar. 2010. [Online], Available: http://vision.okstate.edu/csiq

[43] N. Ponomarenko *et al.*, "Image database TID2013: Peculiarities, results and perspectives," *Signal Process., Image Commun.*, vol. 30, pp. 57–77, Jan. 2015.

[44] A. K. Moorthy, C.-C. Su, A. Mittal, and A. C. Bovik, "Subjective evaluation of stereoscopic image quality," *Signal Process., Image Commun.*, vol. 28, no. 8, pp. 870–883, Sep. 2013.

[45] M.-J. Chen, L. K. Cormack, and A. C. Bovik, "No-reference quality assessment of natural stereopairs," *IEEE Trans. Image Process.*, vol. 22, no. 9, pp. 3379–3391, Sep. 2013.

[46] D. C. Knill and A. Pouget, "The Bayesian brain: The role of uncertainty in neural coding and computation," *Trends Neurosci.*, vol. 27, no. 12, pp. 712–719, 2004.

[47] X. Wu, G. Zhai, X. Yang, and W. Zhang, "Adaptive sequential prediction of multidimensional signals with applications to lossless image coding," *IEEE Trans. Image Process.*, vol. 20, no. 1, pp. 36–42, Jan. 2011.

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

[49] 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.

[50] I. Sekita, T. Kurita, and N. Otsu, "Complex autoregressive model for shape recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 14, no. 4, pp. 489–496, Apr. 1992.

[51] Y. Nakatani, D. Sasaki, Y. Iiguni, and H. Maeda, "Online recognition of handwritten Hiragana characters based upon a complex autoregressive model," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 21, no. 1, pp. 73–76, Jan. 1999.

[52] C. C. Yang and S. H. Kwok, "Efficient gamut clipping for color image processing using LHS and YIQ," *Opt. Eng.*, vol. 42, no. 3, pp. 701–711, Mar. 2003.

[53] S. T. Roweis and L. K. Saul, "Nonlinear dimensionality reduction by locally linear embedding," *Science*, vol. 290, no. 5500, pp. 2323–2326, 2000.

[54] E. Peli, "Contrast in complex images," *J. Opt. Soc. Amer. A*, vol. 7, no. 10, pp. 2032–2040, Oct. 1990.

[55] X. Tang, W. Luo, and X. Wang, "Content-based photo quality assessment," *IEEE Trans. Multimedia*, vol. 15, no. 8, pp. 1930–1943, Dec. 2013.

[56] 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.

[57] VQEG. (Mar. 2000). *Final Report From the Video Quality Experts Group on the Validation of Objective Models of Video Quality Assessment*. [Online]. Available: http://www.vqeg.org/

[58] D. V. Meegan, L. B. Stelmach, and W. J. Tam, "Unequal weighting of monocular inputs in binocular combination: Implications for the compression of stereoscopic imagery," *J. Experim. Psychol., Appl.*, vol. 7, no. 2, pp. 143–153, Nov. 2001.

Ke Gu received the B.S. degree in electronic engineering from Shanghai Jiao Tong University, Shanghai, China, in 2009. He is currently pursuing the Ph.D. degree. He is the Reviewer of some IEEE TRANSACTIONS and journals, including the IEEE TRANSACTIONS ON IMAGE PROCESSING, the IEEE TRANSACTIONS ON CYBERNETICS, the IEEE SIGNAL PROCESSING LETTERS, *Neurocomputing*, the *Journal of Visual Communication and Image Representation*, and *Signal, Image and Video Processing*.

He was a Visiting Student with the Department of Electrical and Computer Engineering, University of Waterloo, Canada, in 2014. From 2014 to 2015, he was a Visiting Student with the School of Computer Engineering, Nanyang Technological University, Singapore. His research interests include quality assessment, visual saliency detection, and contrast enhancement.

Guangtao Zhai (M'10) received the B.E. and M.E. degrees from Shandong University, Shandong, China, in 2001 and 2004, respectively, and the Ph.D. degree from Shanghai Jiao Tong University, Shanghai, China, in 2009. He is currently a Research Professor with the Institute of Image Communication and Information Processing, Shanghai Jiao Tong University.

He was a Visiting Student with the Department of Electrical and Computer Engineering, McMaster University, Hamilton, ON, Canada, from 2008 to 2009, where he was a Post-Doctoral Fellow from 2010 to 2012. From 2012 to 2013, he was a Humboldt Research Fellow with the Institute of Multimedia Communication and Signal Processing, Friedrich Alexander University of Erlangen–Nuremberg, Germany. He received the Award of National Excellent Ph.D. Thesis from the Ministry of Education of China in 2012. His research interests include multimedia signal processing and perceptual signal processing.

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

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

Dr. Lin served as an Associate Editor of the IEEE TRANSACTIONS ON IMAGE PROCESSING, the IEEE TRANSACTIONS ON MULTIMEDIA, the IEEE SIGNAL PROCESSING LETTERS, and *Journal of Visual Communication and Image Representation*. He is also on six IEEE Technical Committees and Technical Program Committees of a number of international conferences. He was the lead Guest Editor for a special issue on perceptual signal processing of the IEEE JOURNAL OF SELECTED TOPICS IN SIGNAL PROCESSING in 2012. He is a Chartered Engineer in the U.K., a fellow of the Institution of Engineering Technology, and an Honorary Fellow of the Singapore Institute of Engineering Technologists. He was the Co-Chair of the IEEE Multimedia Communications Technical Committee special interest Group on Quality of Experience. He was an Elected Distinguished Lecturer of APSIPA in 2012/2013.

Xiaokang Yang (SM'04) received the B.S. degree from Xiamen University, Xiamen, China, in 1994, the M.S. degree from the Chinese Academy of Sciences, Shanghai, China, in 1997, and the Ph.D. degree from Shanghai Jiao Tong University, Shanghai, in 2000. He is currently a Full Professor and the Deputy Director of the Institute of Image Communication and Information Processing with the Department of Electronic Engineering, Shanghai Jiao Tong University.

He was a Research Fellow with the Centre for Signal Processing, Nanyang Technological University, Singapore, from 2000 to 2002. From 2002 to 2004, he was a Research Scientist with the Institute for Infocomm Research, Singapore. He has authored over 80 refereed papers, and has filed six patents. His research interests include video processing and communication, media analysis and retrieval, perceptual visual processing, and pattern recognition. He actively participates in the international standards, such as MPEG-4, JVT, and MPEG-21. He is a member of Visual Signal Processing and Communications Technical Committee of the IEEE Circuits and Systems Society. He received the Microsoft Young Professorship Award 2006, the Best Young Investigator Paper Award at the IS&T/SPIE International Conference on Video Communication and Image Processing, and awards from A-STAR and Tan Kah Kee Foundations. He was the Special Session Chair of Perceptual Visual Processing of the IEEE ICME2006. He is the local Co-Chair of ChinaCom2007 and the Technical Program Co-Chair of the IEEE SiPS2007.

Wenjun Zhang (F'11) received the B.S., M.S., and Ph.D. degrees in electronic engineering from Shanghai Jiao Tong University, Shanghai, China, in 1984, 1987, and 1989, respectively.

He worked as a Post-Doctoral Fellow with Philips Kommunikation Industrie AG, Nuremberg, Germany, from 1990 to 1993, where he was actively involved in developing HD-MAC system. He joined as a Faculty Member with Shanghai Jiao Tong University in 1993, and became a Full Professor with the Department of Electronic Engineering, in 1995. As the National HDTV TEEG Project Leader, he successfully developed the first Chinese HDTV prototype system in 1998. He was one of the main contributors to the Chinese Digital Television Terrestrial Broadcasting Standard issued in 2006 and has been leading team in designing the next generation of broadcast television system in China since 2011. He holds more than 40 patents and has authored over 90 papers in international journals and conferences. His main research interests include digital video coding and transmission, multimedia semantic processing, and intelligent video surveillance. He is a Chief Scientist of the Chinese National Engineering Research Centre of Digital Television, an industry/government consortium in DTV technology research and standardization, and the Chair of Future of Broadcast Television Initiative Technical Committee.