SECOND ORDER NATURAL SCENE STATISTICS MODEL OF BLIND IMAGE QUALITY ASSESSMENT

Zeina Sinno, Constantine Caramanis, and Alan Bovik, Fellow, IEEE

Department of Electrical and Computer Engineering
The University of Texas at Austin Austin, Texas, 78712

ABSTRACT

The univariate statistics of bandpass-filtered images provide powerful features that drive many successful image quality assessment (IQA) algorithms. Bivariate Natural Scene Statistics (NSS), which model the joint statistics of multiple bandpass image samples also provide potentially powerful features to assess the perceptual quality of images, by capturing both image and distortion correlations. Here, we make the first attempt to use bivariate NSS features to build a model of no-reference image quality prediction. We show that our bivariate model outperforms existing state of the art image quality predictors.

***Index Terms*—** Second Order Natural Scene Statistics, Blind Image Quality Assessment, Multimedia

1. INTRODUCTION

Digital images have witnessed tremendous growth as a medium for representation and communication. Since human observers are the ultimate receivers of the visual information in images, subjective experiments using human observers remains the most reliable way to assess the quality of an image. Given that 1.3 trillion images were captured in 2017 [1], relying on human observers to assess picture quality is unrealistic. Building models that predict the quality of images in accordance with human observers is a more feasible solution to this problem. The study of blind (no-reference) IQA models involves building learned predictors that deploy low-level image descriptors as inputs. Many models has been developed that extract distortion specific features [2, 3], and [4], while others train learning machines on NSS features computed from distorted images. Examples of this approach include [5], and [6]. Other notable models, such as Ye *et al.* [7] learn visual code words predictive of image quality, and the completely blind model [8], which measures a distance between distorted and pristine Natural Scene Statistics (NSS) features, without requiring any training on either distorted images or on human opinion scores. Saha *et al.* [9] used visibility measured over multiple scales to predict picture quality. In this work, we combine univariate and bivariate NSS features to build a no-reference IQA model that strongly competes with existing models.

We begin in section 2 by presenting the bivariate feature model, section 3 presents the new predictor and in section 4 we evaluate its performance against other models.

2. NORMALIZED BANDPASS IMAGE CORRELATION MODEL

Here we summarize the normalized bandpass image correlation model, on which we define a set of bivariate IQA features.

First deploy a bank of steerable filters [10] to decompose a luminance image. Steerable filters are often used to model bandpass simple cells in primary visual cortex [11]. A steerable filter at a given frequency tuning orientation θ_1 is defined by:

$$F_{\theta_1}(\boldsymbol{x}) = cos(\theta_1)F_x(\boldsymbol{x}) + sin(\theta_1)F_y(\boldsymbol{x}), \quad (1)$$

where $\boldsymbol{x} = (x, y)$, and F_x and F_y are the gradient components of a two-dimensional unit-energy bivariate isotropic gaussian function having a scale parameter σ:

$$G(\boldsymbol{x}) = \frac{1}{2\pi\sigma^2}e^{\frac{-(x^2+y^2)}{2\sigma^2}}, \quad (2)$$

Each analyzed image is passed through a bank of steerable filters of scales $\sigma \in [2, 3, \ ..., 15]$ and over 15 frequency tuning orientations $\theta_1 \in [0, \pi/15, 2\pi/15, ..., \pi]$, yielding 210 bandpass responses. We exclude $\sigma = 1$ since steerable filters become less well defined at that scale.

Next, we apply divisive normalization on all of the steerable filter responses to model the nonlinear adaptive gain control of V1 neuronal responses in the visual cortex [12]. The divisive normalization model is defined as:

$$u_j(\boldsymbol{x}) = \frac{w_j(\boldsymbol{x})}{\sqrt{s + \sum_{\boldsymbol{y}} g(j(\boldsymbol{y})w_j(\boldsymbol{y}))^2}}, \quad (3)$$

where w_j are the steerable filter responses from filter indexed j, u are the coefficients obtained after divisive normalization, and $s = 10^{-4}$ is a stabilizing saturation constant. The weighted sum in the denominator is computed over a spatial neighborhood of pixels from the same sub-band, where

$g(x_i, y_i)$ is a circularly symmetric Gaussian function having unit volume. To match the increase in scale applied at the steerable filtering step (translated by increasing σ), the variance of $g(x_i, y_i)$ is also increased linearly as a function of σ.

Next, define a window at a fixed position within the cropped image and another sliding window of the same dimensions. Denote the distance between the center of the two windows of bandpass, normalized image samples of interest by d, and the angle between them by θ_2. Also define the relative angle $\theta_2 - \theta_1$, where θ_1 is the sub-band tuning orientation relative to the horizontal of the bandpass filter. Then, compute the correlation between the two windows. The two windows are separated by horizontal and vertical separations δ_x and δ_y, which are varied over the integer range 1 to 25, i.e, distances of $\sqrt{\delta_x^2 + \delta_y^2}$ at spatial orientations $\theta_2 = \arctan(\frac{\delta_y}{\delta_x})$ (relative to horizontal). We limited the range $\theta_2 \in [0, \pi[$ since the quantities being measured are symmetrically defined.

The correlation function model expresses a periodic behavior in the relative angle $\theta_2 - \theta_1$, and can be modeled as:

$$\rho(d, \sigma, \theta_2) = A(d, \sigma, \theta_2) cos(2(\theta_2 - \theta_1)) + c(d, \sigma, \theta_2) \quad (4)$$

where $A(d, \sigma, \theta_2)$ is the amplitude, $c(d, \sigma, \theta_2)$ is an offset, d is the spatial separation between the target pixels, σ is the steerable filter spread parameter, and θ_2 is as before.

Next, define the peak correlation function:

$$P(d, \sigma, \theta_2) = max(\rho(d, \sigma, \theta_2)) = A(d, \sigma, \theta_2) + c(d, \sigma, \theta_2). \quad (5)$$

wherein we may rewrite (4) as:

$$\rho(d, \sigma, \theta_2) = A(d, \sigma, \theta_2) cos(2(\theta_2 - \theta_1)) + [P(d, \sigma, \theta_2) - A(d, \sigma, \theta_2)]. \quad (6)$$

Lee, Mumford and Huang [13] systematically observed that the sample covariances of bandpass image pixels follow an approximate reciprocal power law, of the form $\frac{1}{|d|^b}$. Here, we model the peak correlation function as having a more stable form $\frac{1}{|d|^\beta + 1}$. Given a fixed spatial orientation θ_2 and a scale σ, define

$$\hat{P}(d, \sigma, \theta_2) = \frac{1}{(\frac{d}{\alpha_0(\theta_2) * \sigma})^{\beta_0} + 1} \quad (7)$$

where $\{\alpha_0, \beta_0\}$ are parameters that control the shape and falloff of the peak correlation function, which depend on the spatial orientation θ_2.

We model A as the difference of two functions of the form (7):

$$\hat{A}(d, \sigma, \theta_2) = \frac{1}{(\frac{d}{\alpha_1(\theta_2) * \sigma})^{\beta_1(\theta_2)} + 1} - \frac{1}{(\frac{d}{\alpha_2(\theta_2) * \sigma})^{\beta_2(\theta_2)} + 1} \quad (8)$$

Table 1. Model's behavior against several distortions.

	Blur	JPEG	JPEG 2000	Fast Fading	White Noise
ρ	Monotonic ↗	Monotonic ↗	Non-monotonic ↗	Non-monotonic ↗	Slight ↗, then monotonic ↘
P	Monotonic ↗	Monotonic ↗	Non-monotonic ↗	Non-monotonic ↗	Slight ↗, then monotonic ↘
A	Monotonic ↗	Monotonic ↗	Non-monotonic ↗	Non-monotonic ↗	No change

where $\{\alpha_1, \beta_1, \alpha_2, \beta_2\}$ are parameters that control the shape of A and are functions of θ_2.

Our goal next is then to find, for a fixed spatial orientation θ_2, the values of the parameters $\{\alpha_0, \beta_0\}$ that produce the best fit to (7) and the parameters $\{\alpha_1, \beta_1, \alpha_2, \beta_2\}$, yielding the best fit to (8), resulting in the least mean squared error. We form two optimization systems for P and A that account for scale to find those optimal values, that minimize the error. Denote by D the set of distances for a given spatial orientation θ_2. For the case $\theta_2 = 0$ or $\pi/2$, $D = \{0, 1, 2, 3, ..., 24, 25\}$. Our optimization systems are then expressed as:

$$\begin{aligned} \min_{\alpha_0, b_0} \quad & \sum_{d \in D} \sum_{\sigma=2}^{15} (P(d, \sigma, \theta_2) - \hat{P}(d, \sigma, \theta_2))^2 \\ \min_{\alpha_1, b_1, \alpha_2, \beta_2, b_2} \quad & \sum_{d \in D} \sum_{\sigma=2}^{15} (A(d, \sigma, \theta_2) - \hat{A}(d, \sigma, \theta_2))^2 \end{aligned} \quad (9)$$

We derived and validated the bandpass correlation model in [14, 15, 16] and we were able to verify that we can reconstruct the correlation with very low mean squared error. As a further step, we studied the behavior of the model in the presence of distortions and found that the model parameters vary very systematically in the presence of distortions. We summarize in Table 1 the behavior of ρ, P and A as the level of distortions increases. In [14], we studied the behavior of the model against different parameters and found that the changes are reflected thru the behavior of $\{\alpha_0, \beta_0, \alpha_1, \beta_1, \alpha_2, \beta_2\}$ also consistently vary with distortions.

3. APPLICATION TO BLIND IMAGE QUALITY PREDICTION

3.1. Bivariate Features

Motivated by the observation that distortions lead to systematic and predictable perturbations of our correlation models' features, it is natural to consider whether they can be used to predict the perceptual quality of images. We studied the quality-predictive efficacies of the parameters $\{\alpha_0, \beta_0, \alpha_1, \beta_1, \alpha_2, \beta_2\}$ over multiple spatial angles θ_2. We found α_0, β_0 α_1, and α_2 to be the most responsive to distortion, and hence the most useful for quality prediction. We also found that only using parameters computed along the horizontal, vertical, and diagonal was sufficient; adding more angles did not further boost performance.

We noticed that applying non-linear operations to these raw features boosts performance. Denote by $\{F_1, F_2, ..., F_8\}$ the quality predictive features derived from our correlation model. Noting that $\alpha_i(\theta_2)$, $\beta_i(\theta_2)$; $i = 1, 2, 3$ are all functions of θ_2, then define $F_1 = {\beta_0}^{-1}(0)$; $F_2 = {\beta_0}^{-1}(\frac{\pi}{4})$; $F_3 = {\beta_0}^{-1}(\frac{\pi}{2})$; $F_4 = {\beta_0}^{-1}(\frac{3\pi}{4})$ where $\beta^{-1} = \frac{1}{\beta}$; and also $F_5 = {\alpha_0}^{-1}(0){\alpha_0}^{-1}(\frac{\pi}{2})$; $F_6 = {\beta_0}^{-1}(0){\beta_0}^{-1}(\frac{\pi}{2})$; $F_7 = {\alpha_1}^{-1}(0){\alpha_1}^{-1}(\frac{\pi}{2})$; and $F_8 = {\alpha_2}^{-1}(0){\alpha_2}^{-1}(\frac{\pi}{2})$.

As a first test, we trained a Support Vector Regression (SVR) model using a radial basis function and 80-20 split on the LIVE Challenge images using 8 features as input, obtaining Pearsons linear correlation coefficient (PLCC) and Spearmans rank ordered correlation coefficient (SROCC) both in the range of 0.3. Clearly, taken alone the bivariate features are insufficient predictors of picture quality. However, we find that they are usefully complementary to existing univariate features for the IQA task.

3.2. Mean Subtracted Contrast Normalized Features

Rather than using our model features in isolation, we combine them with univariate NSS that have been used successfully for blind picture quality prediction. The motivation behind this reasoning is that the bivariate correlation model is not standalone, rather it extends existing univariate NSS models and completes a bivariate density model. We computed the mean subtracted, contrast normalized (MSCN) coefficients as used in the BRISQUE model [5]. Given an image I, process luminances via local mean subtraction and divisive normalization, similarly to (3). Mittal *et al.* [5] showed that those MSCN coefficients are disturbed by the presence of distortion. We extracted the arithmetic mean $\dot{\mu}$, sample kurtosis κ, and skewness γ, from the luminance image to obtain ($\dot{\mu}_L$, κ_L, γ_L), and form the chrominance component a* from the CIELAB space ($\dot{\mu}_a$, κ_a, γ_a), and used them as additional features in our predictor.

Furthermore, [5] showed that the histogram of the MSCN coefficients of both pristine and distorted images are modeled as fitting a zero-mean generalized gaussian density (GGD):

$$f(x; \phi; \gamma^2) = \frac{\phi}{2\eta\Gamma(1/\phi)} \exp[-(\frac{|x|}{\eta})^{\phi}] \tag{10}$$

where

$$\eta = \gamma\sqrt{\frac{\Gamma(\frac{1}{\phi})}{\Gamma(\frac{3}{\phi})}} \tag{11}$$

and $\Gamma(\cdot)$ is the gamma function:

$$\Gamma(a) = \int_0^{\infty} t^{a-1}e^{-t}dt \quad a > 0. \tag{12}$$

The shape parameter ϕ controls the shape of the distribution, while η controls its variance. We used the moment matching approach to estimate these two parameters from the histograms of each considered image's MSCN coefficients [17].

Denote by ϕ_L and ι_L the shape and the variance features at scale 1 of the luminance component, and by ϕ_Y and ι_Y the scale and shape parameters of scale 1 from the yellow color channel component. The yellow color channel component is computed on an RGB image I as:

$$Y = \frac{R+G}{2} - \frac{|R-G|}{2} - B \tag{13}$$

Sinno [18] *et al.* observed that the height of the peak at zero of the histogram of the MSCN coefficients is highly correlated with how well exposed the image is. A small peak indicates that the image is well-exposed, whereas a high peak means that the image is poorly-exposed (under exposed or over exposed). Furthermore, they used this information to correct for underexposed and overexposed regions in an image using Laplacian pyramid fusion of multiple shots of the same scene, but varying in exposure. We used the peak at zero of the histogram of MSCN coefficients as a feature in our model, and denote it by δ.

We also considered the pairwise products of neighboring MSCN coefficients along four orientations (H), vertical (V), main-diagonal ($D1$) and secondary-diagonal ($D2$), similarly to [5]. As shown in [5], the histograms of the pairwise MSCN coefficients are well modeled as asymmetrical generalized gaussian distributed (AGGD):

$$f(x; \nu, \eta_l^2, \eta_r^2) = \begin{cases} \frac{\nu}{(\eta_l + \eta_r)}\exp[-(\frac{-x}{\eta_l})] & \text{x} < 0 \\ \frac{\nu}{(\eta_l + \eta_r)}\exp[-(\frac{-x}{\eta_r})] & \text{x} \geq 0 \end{cases} \tag{14}$$

where

$$\eta_l = \iota_l\sqrt{\frac{\Gamma(\frac{1}{\nu})}{\Gamma(\frac{3}{\nu})}} \tag{15}$$

$$\eta_r = \iota_r\sqrt{\frac{\Gamma(\frac{1}{\nu})}{\Gamma(\frac{3}{\nu})}}. \tag{16}$$

The shape parameter ν controls the 'shape' of the distribution while η_l^2 and η_r^2 are scale parameters that control the spread on each side of the mode, respectively. The parameters $(\nu, \eta_l^2, \eta_r^2)$ are also estimated using moment-matching [19]. Next, we also created a reduced resolution version of the luminance image by low pass filtering followed by downsampling by a factor of two, then followed the same procedure as above to obtain $(\nu, \eta_l^2, \eta_r^2)$ at the new scale. In our predictor, we used (ν, η_l^2) as features over the four orientations H, V, $D1$, and $D2$, which we denote by $(\nu_H, {\eta_l^2}_H)$, $(\nu_V, {\eta_l^2}_V)$, $(\nu_{D1}, {\eta_l^2}_{D1})$ and $(\nu_{D2}, {\eta_l^2}_{D2})$. This yields 8 additional features.

Combining the correlation features $F_1 - F_2$ with the MSCN features yields a total of 27 features, as summarized in Table 2.

Table 2. Features used in the bivariate image quality prediction model.

Feature ID	Feature Description
F_1 - F_4	$\beta_0{}^{-1}(0),\ \beta_0{}^{-1}(\frac{\pi}{4}),\ \beta_0{}^{-1}(\frac{\pi}{2}),\ \beta_0{}^{-1}(\frac{3\pi}{4})$
F_5	$\alpha_0{}^{-1}(0)\alpha_0{}^{-1}(\frac{\pi}{2})$
F_6	$\beta_0{}^{-1}(0)\beta_0{}^{-1}(\frac{\pi}{2})$
F_7	$\alpha_1{}^{-1}(0)\alpha_1{}^{-1}(\frac{\pi}{2})$
F_8	$\alpha_2{}^{-1}(0)\alpha_2{}^{-1}(\frac{\pi}{2})$
F_9 - F_{11}	$\dot{\mu}_L,\ \kappa_L,\ \gamma_L$
F_{12} - F_{14}	$\dot{\mu}_a,\ \kappa_a,\ \gamma_a$
F_{15}	δ
F_{16} - F_{17}	$\phi_L,\ \iota_L$
F_{18} - F_{19}	$\phi_Y,\ \iota_Y$
F_{20} - F_{21}	$\nu_H,\ \eta^2_{l\ H}$
F_{22} - F_{23}	$\nu_V,\ \eta^2_{l\ V}$
F_{24} - F_{25}	$\nu_{D1},\ \eta^2_{l\ D1}$
F_{26} - F_{27}	$\nu_{D2},\ \eta^2_{l\ D2}$

4. QUALITY EVALUATION

As a resource to learn a blind IQA model using our model, we used the recent LIVE in the Wild Image Quality Challenge Database ("LIVE Challenge") [20]. This database contains 1162 images captured using mobile devices. This database is a unique and difficult test of blind IQA predictors. Using a regression module, we constructed a mapping from the feature space (Table 2) to human ratings, resulting in a measure of image quality. We used a support vector regressor (SVR) [21] that has been successfully deployed in many prior image quality models [22, 5]. We used the LIBSVM package [23] to implement the SVR with a radial basis function (RBF) kernel and to predict the MOS scores. We split the images randomly and used 80% of it for training and the rest for testing, then we normalized our features, and fed them into the SVR module to predict the MOS score. We repeated the process 50 times. We obtained a median Pearsons linear correlation coefficient (PLCC) of 0.73 and a Spearmans rank ordered correlation coefficient (SROCC) of 0.69 against MOS.

Table 3 compares the performances of various reported algorithms. By using only 27 features, our correlation-enhanced model was able to outperform the other leading models, demonstrating the power of the bivariate features. The performance of our model was only approached by FRIQUEE, which uses a large number of features (more than $20\times$ as many). Notably, the correlation features substantially boosted the performance of simple BRISQUE [5]. A very interesting extension will be to apply the model to temporal pictures, as for example on video frame differences which present highly regular statistical structures [24].

We also tested the performance of our model on the LIVE IQA database [28]. The results are summarized in Table 4. Our predictor also outperformed on this database too.

Furthermore, we performed the p statistical significance test on the different groups of features used by our predictor and we were able to verify that the features deliver statistically

Table 3. Comparison of Image Quality Models on the LIVE Challenge Database.

	Number of Features	PLCC	SROCC
Bivariate Model	27	**0.73**	0.69
FRIQUEE [25]	584	0.72	**0.72**
BRISQUE [5]	36	0.61	0.60
C-DIVIINE [26]	82	0.66	0.63
DIVIINE [22]	88	0.56	0.51
BLIINDS-II [27]	**24**	0.45	0.48
NIQE [8]	36	0.48	0.42

Table 4. Comparison of Image Quality Models on the LIVE IQA Database.

	Number of Features	PLCC	SROCC
Bivariate Model	27	**0.96**	**0.96**
FRIQUEE [25]	584	0.93	0.95
BRISQUE [5]	36	0.94	0.94
C-DIVIINE [26]	82	0.94	0.95
DIVIINE [22]	88	0.93	0.92
BLIINDS-II [27]	**24**	0.92	0.91
NIQE [8]	36	0.92	0.91

significant superior performance. As an additional test, we removed each group of features in our predictor to understand their individual contributions. Removing the BRISQUE derived luminance based features had the greatest impact on performance, followed by our bivariate NSS features. This is not unexpected, because the univariate NSS model in BRISQUE is complemented by our bivariate NSS correlation model.

5. CONCLUSION

We built a new predictor for the IQA problem by combining quality-predictive features from a new bivariate NSS correlation model with known BRISQUE univariate NSS features. The resulting new IQA model was shown to outperform top performing blind image quality assessment models. As a next step, we plan to use those bivariate NSS features to build predictors to asses the quality of different modalities such as millimeter wave, X-ray and infra red images.

6. REFERENCES

[1] S. Heyman, "Photos, photos everywhere," *New York Times*, Jul 2015. [Online]. Available: https://www.nytimes.com/2015/07/23/arts/international/photos-photos-everywhere.html

[2] J. Chen, Y. Zhang, L. Liang, S. Ma, R. Wang, and W. Gao, "A no-reference blocking artifacts metric using selective gradient and plainness measures," pp. 894–897, 2008.

[3] S. A. Golestaneh and D. M. Chandler, "No-reference quality assessment of *JPEG* images via a quality rele-

vance map," *IEEE Sign. Process. Lett.*, vol. 21, no. 2, pp. 155–158, 2014.

[4] R. Barland and A. Saadane, "Reference free quality metric using a region-based attention model for *JPEG*-2000 compressed images," *Electronic Imaging 2006*, pp. 605 905–605 905, 2006.

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

[6] H. Tang, N. Joshi, and A. Kapoor, "Learning a blind measure of perceptual image quality," *IEEE Conf. on Comput. Vis. Patt. Rec.*, pp. 305–312, 2011.

[7] P. Ye and D. Doermann, "No-reference image quality assessment using visual codebooks," *IEEE Trans. Imag. Proc.*, vol. 21, no. 7, pp. 3129–3138, 2012.

[8] A. Mittal, R. Soundararajan, and A. C. Bovik, "Making a completely blind image quality analyzer," *IEEE Sign. Process, Lett.*, vol. 20, no. 3, pp. 209–212, 2013.

[9] A. Saha and Q. M. J. Wu, "Utilizing image scales towards totally training free blind image quality assessment," *IEEE Trans. on Imag. Proc.*, vol. 24, no. 6, pp. 1879–1892, 2015.

[10] W. T. Freeman and E. H. Adelson, "The design and use of steerable filters," *IEEE Trans. Pattern Anal. Machine Intell.*, no. 9, pp. 891–906, 1991.

[11] M. Clark and A. C. Bovik, "Experiments in segmenting texton patterns using localized spatial filters," *Patt. Recog.*, vol. 22, no. 6, pp. 707–717, 1989.

[12] D. L. Ruderman and W. Bialek, "Statistics of natural images: Scaling in the woods," *Phys. Rev. Lett.*, vol. 73, no. 6, p. 814, 1994.

[13] A. B. Lee, D. Mumford, and J. Huang, "Occlusion models for natural images: A statistical study of a scale-invariant dead leaves model," *Int. J. Comput. Vision*, vol. 41, no. 1-2, pp. 35–59, 2001.

[14] Z. Sinno, C. Caramanis, and A. C. Bovik, "Towards a closed form second-order natural scene statistics model," *IEEE Trans. Imag. Proc.*, in revision.

[15] Z. Sinno and A. C. Bovik, "Relating spatial and spectral models of oriented bandpass natural images," *Southw. Symp. Image Anal. Interpret.*, Mar 2016.

[16] Z. Sinno and A. C. Bovik, "Generalizing a closed-form correlation model of oriented bandpass natural images," *IEEE Glob. Conf. Sig. Info. Proc.*, Dec 2015.

[17] K. Sharifi and A. Leon-Garcia, "Estimation of shape parameter for generalized gaussian distributions in subband decompositions of video," *IEEE Trans. on Circ. Sys. Vid. Tech.*, vol. 5, no. 1, pp. 52–56, 1995.

[18] Z. Sinno, C. Bampis, and A. C. Bovik, "Detecting, localizing and correcting exposure-staurated regions using a natural image statistics model," *Ann. Meet. of Vis. Sci. Soc.*, 2017.

[19] N. E. Lasmar, Y. Stitou, and Y. Berthoumieu, "Multiscale skewed heavy tailed model for texture analysis," in *Proc. IEEE Int. Conf. Imag. Process.* IEEE, 2009, pp. 2281–2284.

[20] D. Ghadiyaram and A. C. Bovik, "Massive online crowdsourced study of subjective and objective picture quality," *IEEE Trans. Image Process.*, vol. 25, no. 1, pp. 372–387, 2016.

[21] B. Schölkopf, A. J. Smola, R. C. Williamson, and P. L. Bartlett, "New support vector algorithms," *Neur. Comp.*, vol. 12, no. 5, pp. 1207–1245, 2000.

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

[23] C. C. Chang and C. J. Lin, "LIBSVM: A library for support vector machines," *ACM Trans. Intell. Sys. Tech. (TIST)*, vol. 2, no. 3, p. 27, 2011.

[24] H. R. Sheikh and A. C. Bovik, "A visual information fidelity approach to video quality assessment," *First Intern. Worksh. Vid. Proc. Quall. Met. Cons. Electr.*, pp. 23–25, 2005.

[25] D. Ghadiyaram and A. C. Bovik, "Scene statistics of authentically distorted images in perceptually relevant color spaces for blind image quality assessment," *IEEE Int. Conf. Imag. Process.*, pp. 3851–3855, 2015.

[26] Y. Zhang, A. K. Moorthy, D. M. Chandler, and A. C. Bovik, "C-DIIVINE: No-reference image quality assessment based on local magnitude and phase statistics of natural scenes," *Sig. Proc.: Image Commun.*, vol. 29, no. 7, pp. 725–747, 2014.

[27] M. A. Saad, A. C. Bovik, and C. Charrier, "Blind image quality assessment: A natural scene statistics approach in the dct domain," *IEEE Trans. on Imag. Proc.*, vol. 21, no. 8, pp. 3339–3352, 2012.

[28] Z. Wang, A. Bovik, H. Sheikh, and E. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *IEEE Trans. Image Processing*, vol. 13, no. 4, pp. 600–612, April 2004.