No-Reference Quality Assessment of Contrast-Distorted Images Based on Natural Scene Statistics

Yuming Fang, Kede Ma, Zhou Wang, *Fellow, IEEE*, Weisi Lin, *Senior Member, IEEE*, Zhijun Fang, *Senior Member, IEEE*, and Guangtao Zhai

***Abstract*—Contrast distortion is often a determining factor in human perception of image quality, but little investigation has been dedicated to quality assessment of contrast-distorted images without assuming the availability of a perfect-quality reference image. In this letter, we propose a simple but effective method for no-reference quality assessment of contrast distorted images based on the principle of natural scene statistics (NSS). A large scale image database is employed to build NSS models based on moment and entropy features. The quality of a contrast-distorted image is then evaluated based on its unnaturalness characterized by the degree of deviation from the NSS models. Support vector regression (SVR) is employed to predict human mean opinion score (MOS) from multiple NSS features as the input. Experiments based on three publicly available databases demonstrate the promising performance of the proposed method.**

***Index Terms*—Contrast distortion, image quality assessment, natural scene statistics, no-reference image quality assessment, support vector regression.**

I. Introduction

IMAGE QUALITY ASSESSMENT (IQA) is a fundamental and challenging problem in the field of image processing. Since the Human Visual System (HVS) is the ultimate receiver and interpreter of the visual content, subjective assessment represents the most reliable quality evaluation method. However, subjective evaluation is time-consuming and expensive, making it difficult to be adopted in practical applications, especially when real-time computation is desired. To overcome these limitations, many objective IQA metrics have been proposed during the past decades [1] to predict human perception of image quality. These objective IQA measures are advantageous in terms of repeatability and scalability compared with subjective evaluations [1], [2].

Existing IQA metrics can be classified into three categories based on the availability of the original (reference) image, which is considered to be distorted-free or perfect-quality. Full-reference (FR) metrics require complete access to the reference image [3], [4]; reduced-reference (RR) metrics require only partial information of the reference image [5], [6]; and no-reference (NR) metrics do not require any information of the reference image [7], [8]. In many practical applications, information regarding the reference is not available, and thus NR metrics are highly desirable.

Currently, most existing NR-IQA metrics focus on the quality evaluation of compression, noise or blurring distortions of images [1]. Little investigation is dedicated to the area of NR-IQA for contrast distortion, which is often a determining factor in human perception of image quality. It is worth noting that contrast distortion is often introduced during the image acquisition process. The limitation of the acquisition device or poor lighting condition would cause the loss of contrast and visible details in images [9]. In these scenarios, the acquired image is the original source and there is no perfect-quality reference. Another highly relevant application is image contrast enhancement operations which attempt to improve the visibility of structural details [9]. Nevertheless, how to objectively measure the contrast distortion of images remains an unresolved problem, without which, image contrast enhancement remains ad-hoc, without a clear goal for optimization. In the literature, there are NR metrics designed for general-purpose usages. For example, the authors of [25] and [21] built statistical models of mean subtracted contrast normalized (MSCN) coefficients and spatial relationship between neighboring pixels. The resulting quality models work quite well for noise, blur, and compression distortions, but unfortunately, their performance degrades significantly on contrast distorted images, as will be shown later in this letter. There are several recent studies focusing on contrast distorted images [16], [23], [24], among which the reduced-reference quality metric for contrast-distorted images (RIQMC)

Manuscript received July 09, 2014; revised November 14, 2014; accepted November 17, 2014. Date of publication November 20, 2014; date of current version November 25, 2014. This work was supported in part by the Funds from DoE of Jiangxi Province under Grants GJJ14347 and GJJ14318, the NSF of Jiangxi Province under Grants 20142BAB217011 and 20141BDH80003, and by the Funds from Jiangxi University of Finance and Economics. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Alexandre X. Falcao.

Y. Fang and Z. Fang are with School of Information Technology, Jiangxi University of Finance and Economics, Nanchang, Jiangxi, China (e-mail: fa0001ng@e.ntu.edu.sg; zjfang@gmail.com).

K. Ma and Z. Wang are with Department of Electrical and Computer Engineering, University of Waterloo, Waterloo, ON, Canada (e-mail: k29ma@uwaterloo.ca; Z.Wang@ece.uwaterloo.ca).

W. Lin is with School of Computer Engineering, Nanyang Technological University, Singapore (e-mail: wslin@ntu.edu.sg).

G. Zhai is with Institute of Image Communication and Information Processing, Shanghai Jiao Tong University, Shanghai, China (e-mail: zhaiguangtao@sjtu.edu.cn).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/LSP.2014.2372333

proposed in [16] achieved impressive performance, but unlike our method to be discussed later, this method relies on partial access to the reference image and is not based on natural image statistical models.

In this study, we aim to develop a no-reference quality metric for contrast-distorted images (NR-CDIQA) based on natural scene statistics (NSS) [14], [20]. It is widely accepted that natural images have strong statistical regularities across different visual content and NSS models have been designed to capture those statistical properties [14], [20]. Our method is based on the hypothesis that contrast distortion would temper the statistical regularities of natural images, leading to unnatural appearance that degrades perceived image quality.

The moment features of images have been widely used and proved promising in many studies related to image contrast [10], [11], [12], [13]. It is demonstrated that the first raw moment (mean) of image intensity can be used to represent the overall brightness of images [10]. Recently, the second central moment (variance) of image intensity was used to calculate the expected context-free contrast for optimal contrast-tone mapping [11]. The connections between human perception of image surface and higher central moments (skewness and kurtosis) have also been demonstrated in previous studies [12], [13]. Inspired by the wide use of moment features in image contrast research [10], [11], [12], [13] and the entropy in image processing applications, we extract moment and entropy features from images, and build NSS models upon them using a large-scale image database. The distortion of each contrast-distorted image is measured by its likelihood of naturalness based on the NSS models. Support Vector Regression (SVR) is adopted to map the likelihood feature to perceived quality. Experimental results and comparisons with state-of-the-art FR-IQA and RR-IQA metrics demonstrate the promising performance of the proposed method.

II. Proposed Method

Presumably an image with good quality should look natural, but quantifying the naturalness of an image is not a trivial problem. In the literature, there have been many studies on the statistics of natural images and their relationship with visual perception and image quality [14], [20]. Nevertheless, the current results are still distance away from a comprehensive statistical model of natural images. In this study, since our major focus is on contrast distortion only, we are not attempting to develop a generic NSS model for images. Instead, we focus on those features that are most likely affected by contrast distortion, and build our NSS models based on these features. In particular, we carry out statistics on moment and entropy features of image intensities using the SUN2012 database [15], which includes 16873 images that cover a large variety of image content. We then build NSS models based on these statistics, and subsequently combine these models using a learning based feature fusion method, leading to an NR-IQA algorithm that predicts perceived quality of contrast-distorted images.

A. Feature Extraction and NSS Models

Let $\mathcal{M}$ denotes the sample mean operator. The sample mean $m(\mathbf{I})$, standard deviation $d(\mathbf{I})$, skewness $s(\mathbf{I})$ and kurtosis $k(\mathbf{I})$ of image intensity are computed as:

$$m(\mathbf{I}) = \mathcal{M}(\mathbf{I}), \tag{1}$$

$$d(\mathbf{I}) = \sqrt{\mathcal{M}[(\mathbf{I} - \mathcal{M}(\mathbf{I}))^2]}, \tag{2}$$

$$s(\mathbf{I}) = \frac{\mathcal{M}[(\mathbf{I} - \mathcal{M}(\mathbf{I}))^3]}{d(\mathbf{I})^3}, \tag{3}$$

and

$$k(\mathbf{I}) = \frac{\mathcal{M}[(\mathbf{I} - \mathcal{M}(\mathbf{I}))^4]}{d(\mathbf{I})^4} - 3. \tag{4}$$

The entropy of the image $\mathbf{I}$ can be estimated as:

$$e(\mathbf{I}) = -\sum_j p_j(\mathbf{I}) \log_2 p_j(\mathbf{I}), \tag{5}$$

where $p_j\mathbf{I}$ denotes the frequency of intensity value j occurs in $\mathbf{I}$. All these features are extracted based on image intensity distributions, which do not directly reflect the regularities of structural details in natural images, and thus are not sufficient to establish a generic NSS model. However, since our focus is on contrast distortion, which will certainly affect image intensity distributions, NSS models built upon these features would be useful to capture changes in the naturalness of images undergoing contrast distortions.

We compute the sample mean, standard deviation, skewness, kurtosis and entropy of all images from SUN2012 database [15]. The histograms of these features are shown in Fig. 1. Since the database includes a large number of images of substantially different visual content, the histograms roughly reflect the distribution of these features in natural images.

For the mean, standard deviation, and skewness features, their histograms, as given in Fig. 1(a)–1(c), respectively, can be well fitted by Gaussian probability density functions given by

$$p_m = \frac{1}{\sqrt{2\pi}\sigma_m} \exp\left[-\frac{(m-\mu_m)^2}{2\sigma_m^2}\right], \tag{6}$$

$$p_d = \frac{1}{\sqrt{2\pi}\sigma_d} \exp\left[-\frac{(d-\mu_d)^2}{2\sigma_d^2}\right], \tag{7}$$

and

$$p_s = \frac{1}{\sqrt{2\pi}\sigma_s} \exp\left[-\frac{(s-\mu_s)^2}{2\sigma_s^2}\right], \tag{8}$$

where p_m, p_d, and p_s stand for the likelihood of an image being natural given its mean m, standard deviation d and skewness s, respectively. The fitting parameters are found to be $\mu_m = 118.559$, $\sigma_m = 26.063$, $\mu_d = 57.274$, $\sigma_d = 12.858$, $\mu_s = 0.180$, and $\sigma_s = 0.632$, respectively. The fitting curves of mean, standard deviation and skewness are also shown in Figs. 1(a), 1(b) and 1(c), respectively.

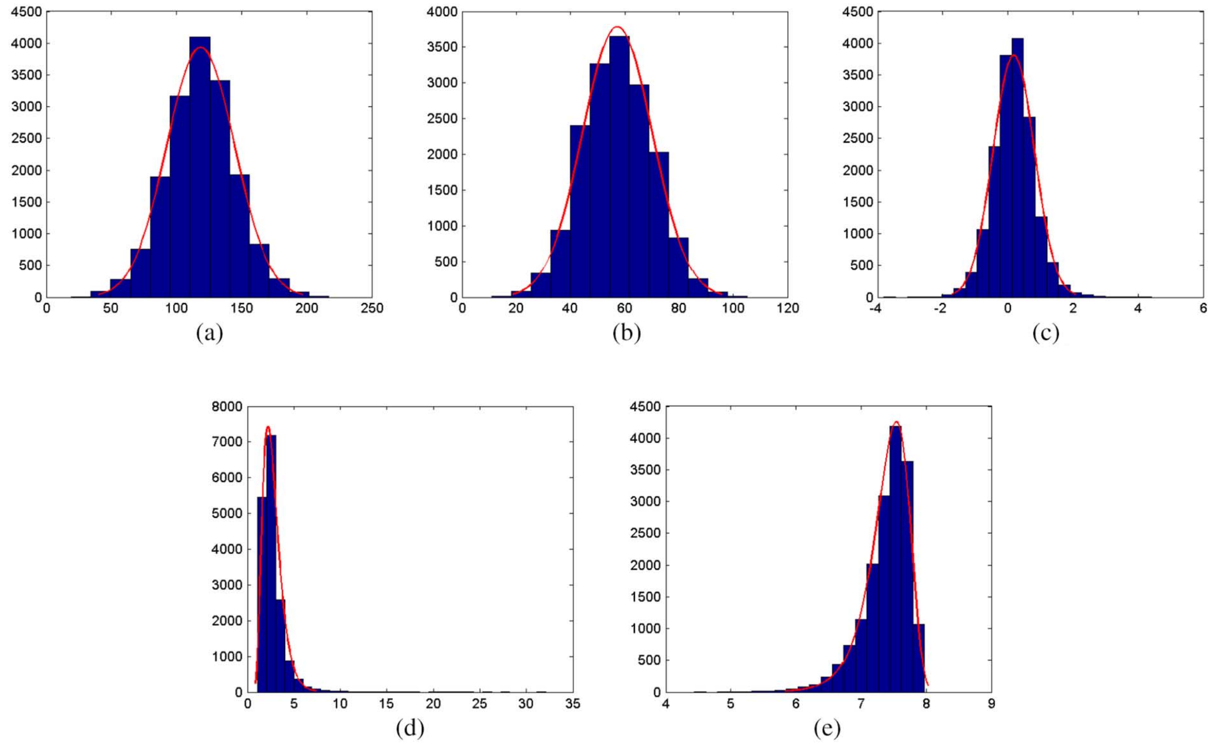

Fig. 1. Histograms and the corresponding fitting curves of different features based on images in SUN2012 database [15] (a) Mean (b) Standard Deviation (c) Skewness (d) Kurtosis (e) Entropy.

The histogram of the kurtosis feature in Fig. 1(d) is found to be well fitted by an inverse Gaussian probability density function given by

$$p_k = \sqrt{\frac{\lambda_k}{2\pi k^3}} \exp\left[\frac{-\lambda_k (k-\mu_k)^2}{2\mu_k^2 k}\right], \tag{9}$$

where p_k denotes the likelihood of an image being natural given its kurtosis value k. λ_k and μ_k are fitting parameters and are found to be $\lambda_k = 19.317$ and $\mu_k = 2.729$, respectively. The fitting curve of kurtosis is also drawn in Fig. 1(d).

We also found that the histogram of the entropy feature in Fig. 1(e) can be well fitted by an extreme value probability density function given by

$$p_e = \frac{1}{\sigma_e} \exp\left[\frac{e-\mu_e}{\sigma_e} - \exp\left[\frac{e-\mu_e}{\sigma_e}\right]\right], \tag{10}$$

where p_e indicates the likelihood of an image being natural given its entropy value e. σ_e and μ_e are fitting parameters and are found to be $\sigma_e = 0.258$ and $\mu_e = 7.540$, respectively. The fitting curve of entropy is also plotted in Fig. 1(e).

B. Quality Prediction Model

For any given image, the set of five likelihood features $\{p_m, p_d, p_s, p_k, p_e\}$ is calculated based on NSS models of mean, standard deviation, skewness, kurtosis and entropy. The quality of the image is predicted based on this feature set. This is a regression problem and we adopt SVR to find the mapping function between the feature set and perceptual quality score.

Assume that $\mathbf{x}_i$ is the feature vector of the ith image in the training image set and y_i is its corresponding desired output (which is the mean opinion score obtained in subjective experiment). We aim to find a function $\hat{\mathrm{y}}_i = f(\mathbf{x}_i)$ that predicts y_i with an acceptable margin of ε. The function to be determined is as follows:

$$\hat{\mathrm{y}} = f(\mathbf{x}) = \mathbf{w}^T \psi(\mathbf{x}) + \gamma \tag{11}$$

where $\psi(\mathbf{x})$ is a kernel function of the feature vector $\mathbf{x}$ and the default setting of Radial Basis Function in [22] is adopted; $\mathbf{w}$ is a weighting vector and γ is the bias term.

In the training stage, the training set is given by $\{\mathbf{x}_i, \mathrm{y}_i\}$ for $i = 1, 2, \ldots, N$, for which more details are given in Section III. The SVR system is employed to estimate ψ, $\mathbf{w}$ and γ in (11). In the test stage, the test feature vector $\mathbf{x}_j$ of the jth test image is the input to the system which creates the objective score $\hat{\mathrm{y}}_j$.

III. Experimental Evaluation

We use three databases to validate the performance of the proposed NR-CDIQA metric: CID2013 [16], TID2013 [17], and CSIQ [18]. CID2013 is built as a benchmark specifically for contrast distortion and includes 15 reference and 400 distorted images in total [16]. Twenty-two subjects were involved in the subjective test to provide a quality score with scale from 1 to 5 for each contrast-distorted image, and the mean opinion score (MOS) of each image is computed. TID2013 contains 25 reference images and 1700 distorted versions with different types of distortions [17], among which the MOS scores in the scale of 0 to 9 of 200 contrast-distorted images are employed for performance evaluation. CSIQ was built from 30 reference images with different types of distortions. It contains 5000 subjective ratings from 35 subjects. The subjective results are reported in the form of Difference of MOS (DMOS) between the reference and distorted images. The 116 contrast-distorted images from CSIQ are selected in our study. The subjective tests on all three databases were conducted following the recommendations of

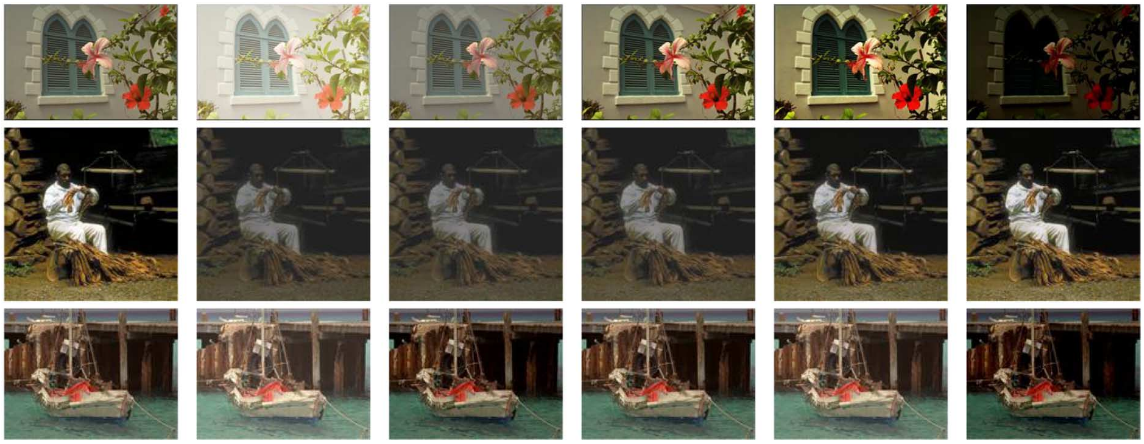

Fig. 2. Visual samples of contrast-distorted images. First column: the reference images; all other columns: contrast-distorted images. Top row: images from CID2013 database [16]; middle row: images from CSIQ database [18]; bottom row: images from TID2013 database [17].

TABLE I
PERFORMANCE EVALUATION BASED ON THREE IMAGE DATABASES

Models	CID2013 Database			TID2013 Database			CSIQ Database		
	PLCC	SRCC	RMSE	PLCC	SRCC	RMSE	PLCC	SRCC	RMSE
PSNR	0.6504	0.6649	0.4733	0.5071	0.5434	0.8453	0.8987	0.8621	0.0739
SSIM	0.8072	0.8132	0.3678	0.6870	0.5510	0.7127	0.7437	0.7397	0.1126
MAD	0.8151	0.8079	0.3610	0.6028	0.5515	0.8738	0.9320	0.9207	0.0611
RIQMC	0.9080	**0.9133**	0.2611	**0.7848**	**0.7239**	**0.6079**	**0.9593**	**0.9576**	**0.0476**
NIQE	0.4778	0.3824	0.8193	0.0234	0.0661	1.7063	0.3572	0.2292	1.0498
Proposed	**0.9096**	0.8874	**0.2387**	0.6588	0.5015	0.6690	0.8508	0.8044	0.0823

ITU-R BT 500-12 [26]. The MOS scores were computed after the subjective experiments by taking the mean of subjective opinion scores. The DMOS scores in CSIQ were obtained by calculating the differences of MOS scores between the reference and distorted images. Some visual samples of contrast-distroted images in these databases are shown in Fig. 2.

To the best of our knowledge, there is no other NR measure that has been developed specifically for contrast distortion. We compare the performance of the proposed method with existing image quality metrics including PSNR, SSIM [3], MAD [19], RIQMC [16] and NIQE [21], among which, PSNR, SSIM and MAD are widely used FR quality metrics; RIQMC is a recent RR metric designed specifically for contrast-distorted images; NIQE is a NR quality metric for general-purpose quality assessment. Three commonly used performance metrics are employed to compare the subjective and objective quality evaluation results: Pearson Linear Correlation Coefficient (PLCC), Spearman Rank-Order Correlation Coefficient (SRCC) and Root Mean-Square Error (RMSE). A good objective quality is expected to achieve high values in PLCC and SRCC, while low values in RMSE.

We divide each image database into 10 subsets and use ten-fold leave-one-out cross-validation to test the proposed metric. Specifically, at each time, nine of the subsets are used for training and the remaining one for testing. The procedure repeats 10 times, each with a different testing set. The average results of the ten-fold evaluation for each database are shown in Table I, and the average performance over three databases is summarized in Table II. For CID2013, a dedicated database for contrast distortion, the proposed method achieves significantly better performance than popular FR methods and is comparable

TABLE II
AVERAGE PERFORMANCE OVER THREE DATABASES

Models	PLCC	SRCC
PSNR	0.6854	0.6901
SSIM	0.7460	0.7031
MAD	0.7833	0.7600
RIQMC	0.8840	0.8649
NIQE	0.2861	0.2259
Proposed	0.8064	0.7311

to RIQMC, an RR method specifically designed for contrast distortion [16]. The results of the proposed method on contrast-distorted images in TID2013 and CSIQ databases are also promising when compared against FR and NR metrics, given the fact that no information about the reference image is available to the proposed method. Note that although state-of-the-art NR metric NIQE performs quite well in other tests [21], it has major difficulty on contrast-distorted images.

IV. CONCLUSION

We propose an NR quality metric for contrast-distorted images based on NSS models built upon moment and entropy features. Promising quality prediction performance is achieved based on our test on three public databases. Our results suggest that NSS models are promising at handling contrast-distorted images. On the other hand, NSS modeling is a highly challenging problem by itself that needs future research to resolve. The current research is limited at capturing some specific aspects of NSS modeling that are likely to be relevant to the perceptual quality of contrast-distorted images. Future work will be focused on better NSS models and the extension of the application scope of the general approach to other types of distortion.

REFERENCES

[1] W. Lin and C. J. Kuo, "Perceptual visual quality metrics: A survey," *J. Vis. Commun. Image Represent.*, vol. 22, no. 4, pp. 297–312, 2011.

[2] Z. Wang and A. C. Bovik*, Modern Image Quality Assessment (Syntheses Lectures on Image, Video and Multimedia Processing)*. San Mateo, CA, USA: Morgan & Claypool, 2006.

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

[4] 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.

[5] A. Rehman and Z. Wang, "Reduced-reference image quality assessment by structural similarity estimation," *IEEE Trans. Image Process.*, vol. 21, no. 8, pp. 3378–3389, 2012.

[6] J. Wu, W. Lin, G. Shi, and A. Liu, "Reduced-reference image quality assessment with visual information fidelity," *IEEE Trans. Multimedia*, vol. 15, no. 7, pp. 1700–1705, 2013.

[7] G. Zhai, W. Zhang, X. Yang, W. Lin, and Y. Xu, "No-reference noticeable blockiness estimation in images," *Signal Process: Image Commun.*, vol. 23, no. 6, pp. 417–432, 2008.

[8] R. Hassen, Z. Wang, and M. M. A. Salama, "No-reference image sharpness assessment based on local phase coherence measurement," in *IEEE ICASSP*, 2010.

[9] T. Arici, S. Dikbas, and Y. Altunbasak, "A histogram modification framework and its application for image contrast enhancement," *IEEE Trans. Image Process.*, vol. 18, no. 9, pp. 1921–1935, Sep. 2009.

[10] R. C. Gonzalez and R. E. Woods, "Image enhancement in the spatial domain," in *Digital Image Processing.*, 3rd Ed. ed. Reading, MA, USA: Addison-Wesley, 2008.

[11] X. Wu, "A linear programming approach for optimal contrast-tone mapping," *IEEE Trans. Image Process.*, vol. 20, no. 5, pp. 1262–1272, May 2011.

[12] I. Motoyoshi, S. Nishida, L. Sharan, and E. H. Adelson, "Image statistics and the perception of surface qualities," *Nature*, vol. 447, pp. 206–209, May 2007.

[13] D. Zoran and Y. Weiss, "Scale invariance and noise in natural images," in *Proc. the IEEE Int. Conf. Computer Vision*, 2009, pp. 2209–2216.

[14] E. P. Simoncelli and B. A. Olshausen, "Natural image statistics and neural representation," *Annu. Rev. Neurosci*, vol. 24, pp. 1193–1216, May 2001.

[15] J. Xiao, J. Hays, K. Ehinger, A. Oliva, and A. Torralba, "SUN database: Large-scale scene recognition from abbey to zoo," in *IEEE Conf. Computer Vision and Pattern Recognition*, 2010.

[16] K. Gu, G. Zhai, X. Yang, W. Zhang, and M. Liu, "Subjective and objective quality assessment for images with contrast change," in *IEEE Int. Conf. Image Processing*, 2013.

[17] N. Ponomarenko, O. Ieremeiev, V. Lukin, K. Egiazarian, L. Jin, J. Astola, B. Vozel, K. Chehdi, M. Carli, F. Battisti, and C.-C. Jay Kuo, "Color image database TID2013: Peculiarities, and preliminary results," in *Proc. 4th Europian Workshop on Visual Information Processing*, 2013.

[18] E. C. Larson and D. M. Chandler*, Categorical image quality (CSIQ) database* [Online]. Available: http://vision.okstate.edu/csiq

[19] E. C. Larson and D. M. Chandler, "Most appearant distortion: Full-reference image quality assessment and the role of strategy," *J. Electron. Imag.*, vol. 19, no. 1, 2010.

[20] W. Geisler, "Visual perception and the statistical properties of natural scenes," *Annu. Rev. Neurosci.*, 2007.

[21] A. Mittal, R. Soundararajan, and A. C. Bovik, "Making a completely blind image quality analyzer," *IEEE Signal Process. Lett.*, vol. 22, no. 3, pp. 209–212, 2013.

[22] [Online]. Available: http://www.csie.ntu.edu.tw/~cjlin/libsvm/

[23] S. Triantaphillidou, J. Jarvis, and G. Gupta, "Contrast sensitivity and discrimination in pictorial images," *Proc. SPIE 9016, Image Qual, Syst, Perf. XI*, 2014.

[24] A. K. Tripathi, S. Mukhopadhyay, and A. K. Dhara, "Performance metrics for image contrast," in *IEEE Int. Conf. Image Information Processing*, 2011.

[25] 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, 2012.

[26] ITU, ITU-R Recommendation BT 500-12: Methodology for the subjective assessment of the quality of television pictures ITU - Radiocom. Sector, Tech. Rep., Sep. 2009.