Journal of Network Intelligence
Taiwan Ubiquitous Information
 ISSN 2414-8105 (Online)
Volume **2**, Number **1**, February **2017**

Blind Quality Assessment of Dehazed Images by Analyzing Information, Contrast, and Luminance

Wei Shen, Shuman Hao*, Jiansheng Qian, Leida Li

School of Information and Electrical Engineering
China University of Mining and Technology
Xuzhou, 221116, China
*Corresponding author: haoshuman@163.com

Abstract. *Image dehazing is widely used to improve the visual quality of dehazed images. While image dehazing has been extensively investigated, the relevant quality assessment of dehazed images remains an open problem, which may hinder further improvement of dehazing techniques. In this paper, a blind quality metric for dehazed images is proposed. Three groups of features are extracted for characterizing the quality of dehazed images, including information, contrast, and luminance. Specifically, a total of 11 perceptual features are extracted and used to train a support vector regression (SVR) model. Then, the trained SVR model is used for predicting the quality of dehazed images. A dehazed image database is built to evaluate the performance of the proposed method. The experimental results demonstrate the efficiency and advantage of the proposed metric.*
Keywords: Image dehazing, Image quality assessment, No reference, Dehazed image database, Information entropy, SVR

1. **Introduction.** Images of outdoor scenes are usually degraded due to the presence of haze and particles. Dehazing, also referred to as haze removal, is highly desirable for both consumer/computational photography and computer vision tasks. Many computer vision algorithms can only work well in haze-free environments. With the increasing prevalence of mobile devices, which have been equipped with compact cameras, image dehazing functionality has been embedded into the image processing software for providing better Quality of Experience (QoE). In the current image dehazing researches, the quality of dehazed images is mainly judged by subjective test, which is laborious and not applicable to real-time automatous systems. Objective quality models are useful for benchmarking and optimizing image dehazing algorithms. The current image quality models are mainly designed for degraded images, so they are very limited in predicting the quality of dehazed images. Quality evaluation of dehazed images is still an open problem.

So far, very little work has been dedicated to the quality evaluation of dehazed images. Hautière *et al.* [1] proposed the blind evaluation of contrast enhancement algorithm based on the visible edge, then Tarel *et al.* [2] directly used this metric for dehazed image quality evaluation. Wang *et al.* [3] directly used the DIIVINE [4] indicator to evaluate the quality of dehazed images. More recently, Li *et al.* [5] devised a full-reference (FR) framework for enhanced image quality evaluation, which consisted of a structure module and a color module.

The aforementioned quality metrics are not specifically designed for dehazed images, so image dehazing quality metrics are still lacking. This paper proposes a new quality model for image dehazing. We achieve the goal by investigating three aspects of quality features in dehazed images, including information, contrast, and luminance. A set of 11 features are extracted and used to train a support vector regression (SVR) model. The quality of a query image is predicted based on SVR regression. A DEHazed Image Database (DEHID) is built to evaluate the performance of the proposed method. Experimental results on DEHID demonstrate the effectiveness and advantages of the proposed method.

2. **Proposed Quality Model for Dehazed Images.** In the past few years, extensive image quality metrics have been proposed, which are mostly based on the measurement of structural distortions [6]. However, in real world applications, the perceived quality is generally determined by many different aspects, which also holds for dehazed images. As a result, measuring non-structural distortions is also important for dehazed images, which will be demonstrated in the experiment section. Non-structural factors like information, contrast, and luminance are needed [7, 8].

With these observations, we propose a new NR quality metric for dehazed images by measuring information, contrast, and luminance. For information, we use entropy. Contrast and luminance are also included in the overall quality of dehazed images. Finally, SVR is used to train dehazed image quality model.

2.1. **Information.** Entropy measures the amount of information in an image. Typically, a high-quality image has high entropy value, which can be changed by the presence of distortions. In [9], image entropy is utilized to represent the information of tone-mapped images. Similar to [9], we calculate the information entropy of a set of intermediate images by brightening/darkening the brightness of the original input dehazed image. The produced intermediate images are defined by

$$I_i = min(max(M_i \cdot I, 0), 255), \tag{1}$$

where I is an input dehazed image, and M_i denotes the ith multiplier. The max and min operators are applied to clip the intermediate images into the range of $0 \sim 255$.

Fig. 1 (a) and (b) show two dehazed images in DEHID database. A group of intermediate images with $M = \{1, n, \frac{1}{n} | n = 1.5, 2.0, \cdots, 10\}$ for each of this pair of images is created before measuring the associated entropy values. Fig. 1 (c) demonstrates how the entropy E varies with the changes of the multiplier M. Referring to subjective human ratings in the DEHID database, (a) indeed has a higher quality score than (b).

To find a good balance between efficiency and efficacy, we take advantage of only nine entropy values that are measured using $M = \{1, n, \frac{1}{n} | n = 3.5, 5, 6.5, 8\}$ as features. It should be noted that we also testify and compare the performance of numerous other choices for n, and results show that our final selection can bring better performance boosts relative to the majority of choices.

Notice that the above-mentioned nine numbers are global based entropy. Broadly speaking, the perception of the human brain to visual signals inclines to a local-and-global manner [10]. Hence the novel defined image entropy of each of nine intermediate images state above is defined as [9]:

$$E_t(I_i) = \omega E_g(I_i) + (1 - \omega) E_l(I_i)^{\nu}, \tag{2}$$

where $E_g(I_i)$ and $E_l(I_i)$ indicate the global and local entropy respectively, ω and ν are positive constants for manipulating the relative importance of the two components above. Referring to the scheme in [11], the local entropy is estimated as the mean of block-based entropy, which is defined by

$$E_l(I_i) = \frac{1}{L} \sum_{j=1}^{L} E(B_{i,j}), \tag{3}$$

where $B_{i,j}$ represents the j th block of size 64×64 in the i th intermediate image, L is the number of the blocks in the image.

2.2. **Contrast.** Contrast plays a vital role in image dehazing. The goodness of image contrast can be a very important factor of image quality. Image contrast can be approximately described using image histogram. An image with good contrast is expected to have uniform histogram distribution. Therefore, the distance/similarity between a real histogram and the ideal uniform histogram can be used as a measurement of goodness of image contrast. In this work, the Kullback-Leibler divergence is adopted to compute the contrast feature [12]

$$D_{KL}(P, U) = \sum_{x} P(x) ln \frac{P(x)}{U(x)}, \tag{4}$$

where P and U represent the histogram distributions of a query image and the ideal uniform distribution.

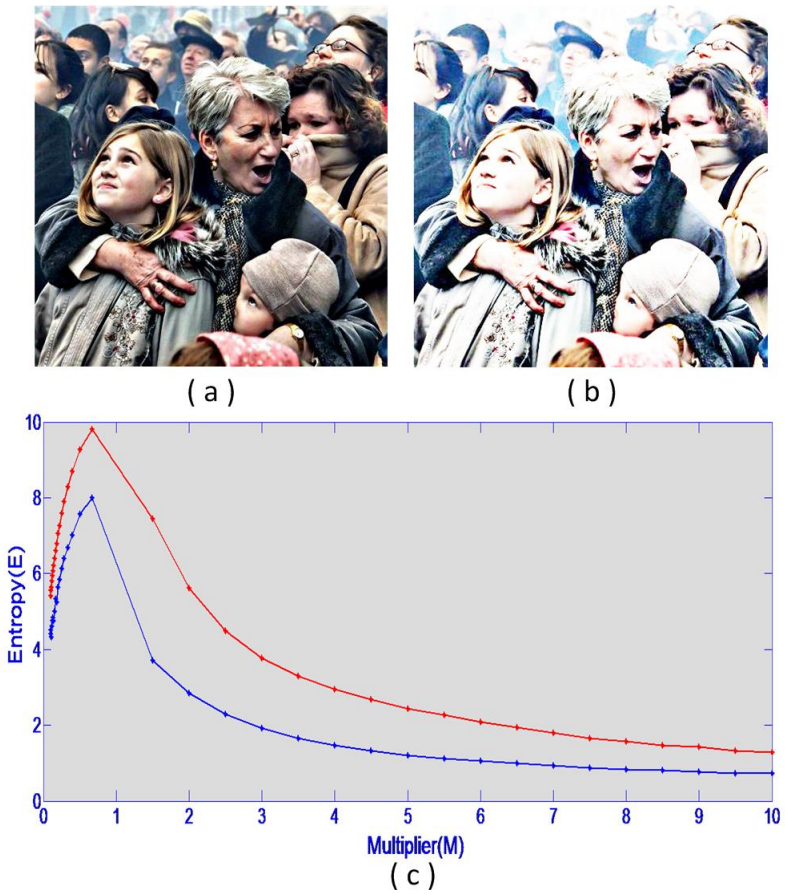

FIGURE 1. (a) A comparatively high-quality dehazed image. (b) A low quality dehazed image. (c) The relationship between the varied multiplier M and the corresponding entropy E in (a) and (b). The top red curve and bottom blue curve correspond to (a) and (b), respectively.

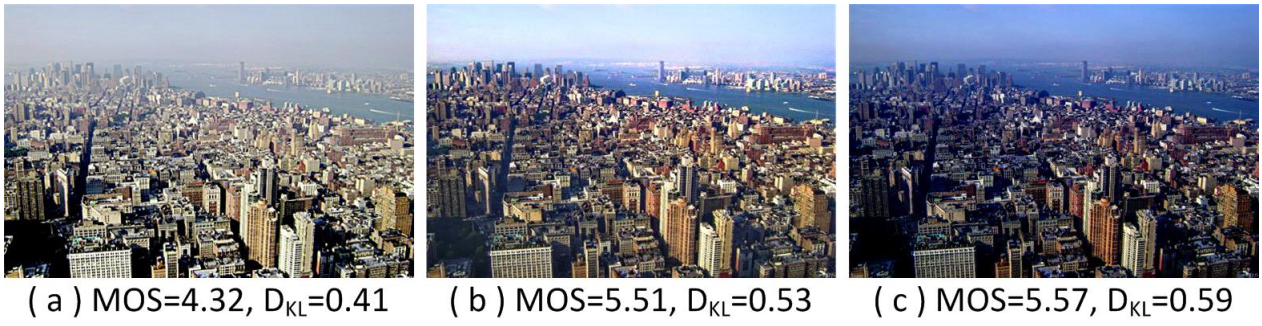

FIGURE 2. Example images selected from the DEHID database, their subjective scores (MOS), and contrast values (D_{KL}).

2.3. **Luminance.** In image dehazing, both under- and over- dehazings cause image luminance to be uncomfortable. So a measurement of image luminance is needed. In this work, we employ the locally weighted luminance value as the measurement of luminance quality [12]

$$Q_L = I_{avg} \times C_L, \tag{5}$$

where I_{avg} denotes the average pixel value, and C_L denotes the Michelson contrast:

$$C_L = \frac{I_{max} - I_{min}}{I_{max} + I_{min}}, \tag{6}$$

where I_{max} and I_{min} denote the maximum and minimum intensity values. In implementation, the luminance feature C_L is computed block-wisely. Specifically, the original image is first partitioned into non-overlapping blocks, and the luminance features are obtained. The mean of the values of all blocks is calculated as the final feature.

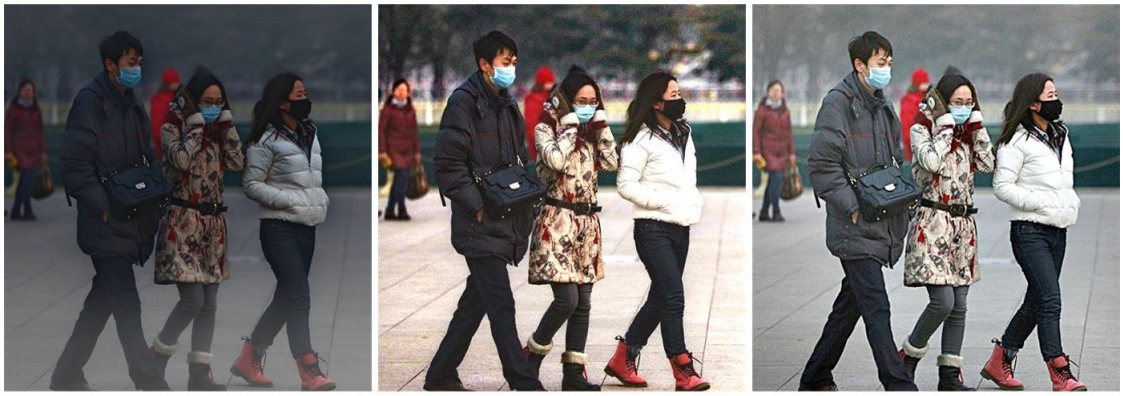

FIGURE 3. Example images selected from the DEHID database, their subjective scores (MOS), and luminance values (Q_L).

2.4. **SVR-based quality prediction.** With the aforementioned quality-aware perceptual features, we employ the support vector regression (SVR) [13] to train a quality model. For a query image, the quality can be predicted using the trained SVR model. In our implementation, the Radial Basis Function (RBF) is used as the SVR kernel.

3. Experimental Results.

3.1. **Dehazed Image Database (DEHID).** In order to evaluate the performance of the proposed method, a dehazed image quality database(DEHID) is needed, which should contain dehazed images and their subjective scores. Since such an image database is not publicly available, we first build one and then use it to conduct the experiments. Specifically, a number of degraded images are first selected. Then each image is processed using the state-of-the-art image dehazing methods, producing the dehazed images. Finally, a subjective test is conducted to collect the human scores, which are used as ground truth.

We select 40 hazy images to cover diverse outdoor scenes, which is divided into mist, moderate and thick. The contents of hazy images include humans, animals, plants, landscapes, traffics, architectures, statics and morn scenes, as shown in Fig. 4.

Ten dehazing algorithms are selected to generate the dehazed images. These algorithms include advanced algorithms 1) fattal08 [14], 2) tarel09 [2], 3) meng13 [15], 4)he09 [16], 5) GDCP [17], 6) gao14 [18], 7) chen15 [19], 8) DEFADE [20], 9) MSRCR [21] and popular image processing software 10) Kolor photoshop auto-contrast [22]. In processing the haze images, default parameter settings are adopted. To be specific, each image is processed using the ten approaches, producing ten dehazed versions of the image. Finally, a total of 400 dehazed images are produced, which constitute the DEHID database. It should be noted that all images in the DEHID database are represented in color format. Fig. 5 shows some example images in the database.

In order to obtain the ground truth of image quality, subjective test is performed using the single stimulus (SS) method on an Absolute Category Rating-Hidden Reference (ACR-HR) scale $1 \sim 10$, corresponding to the worst-best quality. In the test, 30 inexperienced volunteers are involved, including 14 males and 16 females, all aged between 19 and 38. After obtaining the rating scores of all subjects, outliers are removed using the method in [23]. For each image, an average of six outliers are removed. Then the mean of the remaining scores are computed and used as the ground truth, which is also known as the Mean Opinion Score (MOS).

FIGURE 4. Ten of the original images used to build the DEHID database.

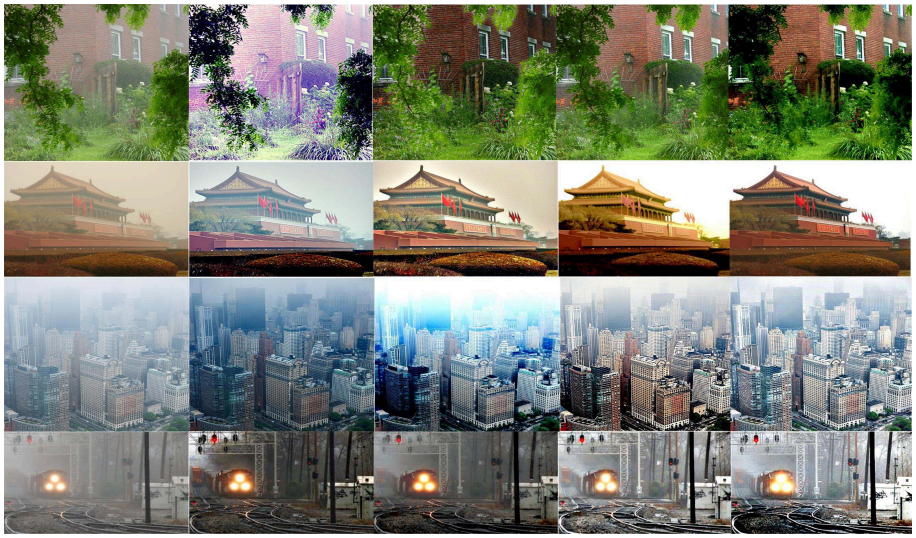

FIGURE 5. Example images and their dehazed versions using different dehazing methods. The leftmost column shows the original images.

3.2. **Evaluation protocols.** Three widely adopted criteria are adopted to evaluate metric performances, including Pearson Linear Correlation Coefficient (PLCC), Spearman Rank order Correlation Coefficient (SRCC) and Root Mean Squared Error (RMSE). PLCC and RMSE are used to measure prediction accuracy, and SRCC is used to measure prediction monotonicity. Before computing them, a five-parameter logistic mapping is performed between the subjective and objective scores

$$q(r) = \tau_1(\frac{1}{2} - \frac{1}{1 + e^{\tau_2(r-\tau_3)}}) + \tau_4 r + \tau_5, \tag{7}$$

where $\tau_i, i = 1, 2, ..., 5$ are the parameters to be fitted.

3.3. **Results and comparisons.** In our experiments, 80% of the images are randomly selected for training, and the remaining 20% images are used for test. The training-test process is repeated by 1000 times, and the median values are used as the performance results. The performance of the proposed method is also compared with the state-of-the-art FR and NR image quality metrics, including, FR: SSIM [5], MS-SSIM [24], PSNR [25], FSIM, FSIMc [26], VIF [27], GSM [28], GMSD [29], SFF [30], PC-QI [31], RIQMC [32], and NR: BIQI [33], BLIINDS2 [34], BRISQUE [35], CORNIA [36], DESIQUE [37], DIIVINE [4], NIQE [38], QAC [11], SSEQ [39], Fang [40], Hautière [1]. The experimental results are summarized in Tables 1, where the best results are marked in boldface.

It can be seen from Tables 1 and 2 that the proposed method achieves the best performances in DEHID database in terms of both prediction accuracy and monotonicity. Furthermore, it significantly outperforms the other compared metrics. These results are not surprising, because almost all the current quality metrics are based on the measurement of structural distortions, which are not the dominated

TABLE 1. Performances of FR IQA metrics in DEHID database.

Metric	**Type**	**PLCC**	**SRCC**	**RMSE**
SSIM	FR	0.1831	0.1046	1.9041
MS-SSIM	FR	0.2409	0.1506	1.8799
PSNR	FR	0.1948	0.1590	1.8997
FSIM	FR	0.2224	0.2008	1.8883
FSIMc	FR	0.2271	0.2042	1.8862
VIF	FR	**0.2636**	**0.2370**	**1.8683**
GSM	FR	0.1902	0.1556	1.9014
GMSD	FR	0.2034	0.1531	1.8963
SFF	FR	0.2213	0.1840	1.8888
PCQI	FR	0.2594	0.2429	1.8705
RIQMC	FR	0.2555	0.2553	1.8725
Proposed metric	NR	**0.5635**	**0.5502**	**1.5832**

TABLE 2. Performances of NR IQA metrics in DEHID database.

Metric	**Type**	**PLCC**	**SRCC**	**RMSE**
BIQI	NR	0.1415	0.1223	1.9176
BLIINDS2	NR	0.3144	0.2737	1.8386
BRISQUE	NR	**0.3301**	**0.3014**	**1.8282**
CORNIA	NR	0.1543	0.1432	1.9136
DESIQUE	NR	0.3291	0.3200	1.8289
DIIVINE	NR	0.1685	0.1426	1.9208
NIQE	NR	0.2096	0.1762	1.8937
QAC	NR	0.1669	0.1023	1.9096
SSEQ	NR	0.3174	0.3118	1.8372
Fang	NR	0.2711	0.2722	1.8642
Hautière	NR	0.1385	0.1047	1.9197
Proposed metric	NR	**0.5635**	**0.5502**	**1.5832**

distortions in image dehazing. In other words, non-structural distortions play a more important role in the quality assessment of dehazed images.

4. **Conclusions.** In this paper, we have addressed the quality evaluation of digitally dehazed images, an important yet much less investigated problem. We have built a dehazed image database based on nine image dehazing algorithms and one image processing software. Subjective test has been conducted to collect the ground truth of human scores. With the consideration that the current distortion-based quality metrics are very limited in the quality evaluation of dehazed images, we have proposed a new quality metric for dehazed images by simultaneously measuring information, contrast, and luminance. We have evaluated the performance of the proposed method on the image dehazing databases. The experimental results have confirmed the effectiveness and advantages of the proposed quality model. As future work, we will use the proposed model for benchmarking and perceptual optimization of image dehazing algorithms.

Acknowledgment. This work was supported in part by the National Natural Science Foundation of China under Grant 61379143, in part by National Key R&D Program under Grant 2016YFC0801800, in part by the Fundamental Research Funds for the Central Universities under Grant 2015QNA66, and in part by the Qing Lan Project of Jiangsu Province.

REFERENCES

[1] N. Hautière, J. P. Tarel, D Aubert, and èRIC DUMONT, "Blind Contrast Enhancement Assessment by Gradient Ratioing at Visible Edges," *Image Analysis and Stereology* , vol. 27, no. 2, pp. 87-95, 2008.

[2] J. P. Tarel, and N. Hautière, "Fast visibility restoration from a single color or gray level image," *IEEE International Conference on Computer Vision*, Kyoto, pp. 2201-2208, 2009.

[3] Y. K Wang, and C. T Fan, "Single image defogging by multiscale depth fusion," *IEEE Transactions on Image Processing* , vol. 23, no. 11, pp. 4826-37, 2014.

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

[5] L. D. Li, Y. Zhou, J. J. Wu, J. S. Qian, and B. J. Chen, "Color-enriched gradient similarity for retouched image quality evaluation," *IEICE Transactions on Information and Systems*, vol. E99-D, no. 3, pp. 773-776, 2016.

[6] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *IEEE Transactions on Image Processing*, vol. 13, no. 4, pp. 600-612, 2004.

[7] A. K. Moorthy, K. Seshadrinathan, R. Soundararajan, and A. C. Bovik, "Wireless Video Quality Assessment: A Study of Subjective Scores and Objective Algorithms," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 20, no. 4, pp. 587-599, 2010.

[8] L. D. Li, W. H. Xia, Y. M. Fang, K. Gu, J. J. Wu, W. S. Lin, and J. S. Qian, "Color image quality assessment based on sparse representation and reconstruction residual," *Journal of Visual Communication and Image Representation*, vol. 38, pp. 550-560, 2016.

[9] K. Gu, S. Q. Wang, G. T. Zhai, S. W. Ma, X. K. Yang, W. S. Lin, W. J. Zhang, and W. Gao, "Blind Quality Assessment of Tone-Mapped Images Via Analysis of Information, Naturalness, and Structure," *IEEE Transactions on Multimedia*, vol. 18, no. 3, pp. 432-443, 2016.

[10] K. Gu, G. Zhai, X. Yang, and W. Zhang, "An efficient color image quality metric with local-tuned-global model," IEEE International Conference on Image Processing , Paris, pp. 506-510, 2014.

[11] W. Xue, L. Zhang, and X. Mou, "Learning without Human Scores for Blind Image Quality Assessment," *IEEE Conference on Computer Vision and Pattern Recognition*, Portland, OR, pp. 995-1002, 2013.

[12] M. A. Saad, P. Corriveau, and R. Jaladi, "Consumer content framework for blind photo quality evaluation," *International Workshop on Video Processing and Quality Metrics for Consumer Electronics*, 2015.

[13] C. C. Chang, C. J. Lin, LIBSVM: A library for support vector machines. *ACM Transactions on Intelligent Systems and Technology*, vol. 2, no. 3, pp. 389-396, 2011.

[14] R. Fattal, "Single image dehazing," *ACM Transactions on Graphics (TOG)*, vol. 27, no. 3, pp. 1-9, 2008.

[15] G. Meng, Y. Wang, J. Duan, S. Xiang, and C. Pan, "Efficient image dehazing with boundary constraint and contextual regularization," *IEEE International Conference on Computer Vision*, pp. 617-624, 2013.

[16] K. M. He, J. Sun, and X. Tang, "Single image haze removal using dark channel prior," *IEEE Conference on Computer Vision and Pattern Recognition*, Miami, FL, pp. 1956-1963, 2009.

[17] K. M. He, J. Sun, and X. Tang, "Guided Image Filtering," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 35, no. 6, pp. 1397-1409, 2013.

[18] Y. Y. Gao, H. M. Hu, S. H. Wang, and B. Li, "A fast image dehazing algorithm based on negative correction," *Signal Processing*, vol. 130, no. 10, pp. 380-398, 2014.

[19] Y. H. Lai, Y. L. Chen, C. J. Chiou, and C. T. Hsu, "Single-Image Dehazing via Optimal Transmission Map Under Scene Priors," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 25, no. 1, pp. 1-14, 2015.

[20] L. K. Choi, J. You, and A. C. Bovik, "Referenceless Prediction of Perceptual Fog Density and Perceptual Image Defogging," *IEEE Transactions on Image Processing*, vol. 24, no. 11, pp. 3888-3901, 2015.

[21] A. B. Petro, C. Sbert, and J. M. Morel, "Multiscale Retinex," *Image Processing On Line*, pp. 71-88, 2014.

[22] Kolor neutralhazer plugin for photoshop, http://www.kolor.com/.

[23] ITU Recommendation, Methodology for the subjective assessment of the quality of television pictures, International Telecommunication Union/ITU Radiocommunication Sector, 2009.

[24] Z. Wang, E. P. Simoncelli, and A. C. Bovik, "Multiscale structural similarity for image quality assessment," *Conference Record of Asilomar on Signals, Systems and Computers*, vol. 2, pp. 1398-1402, 2003.
[25] C. Li, and A. C. Bovik, "Content-weighted video quality assessment using a three-component image model," *Journal of Electronic Imaging*, vol. 29, no. 1, pp. 143-153, 2010.
[26] L. Zhang, L. Zhang, X. Q. Mou, and D. Zhang, "FSIM: a feature similarity index for image quality assessment," *IEEE Transactions on Image Processing*, vol. 20, no. 8, pp. 2378-2386, 2011.
[27] H. R. Sheikh, and A. C. Bovik, "Image information and visual quality," *IEEE Transactions on Image Processing*, vol. 15, no. 2, pp. 430-444, 2006.
[28] A. M. Liu, W. S. Lin, and M. Narwaria, "Image quality assessment based on gradient similarity," *IEEE Transactions on Image Processing*, vol. 21, no. 4, pp. 1500-1512, 2012.
[29] W. F. Xue, L. Zhang, X. Q. Mou, and A. C. Bovik, "Gradient magnitude similarity deviation: a highly efficient perceptual image quality index," *IEEE Transactions on Image Processing*, vol. 23, no. 2, pp. 684-695, 2014.
[30] H. W. Chang, H. Yang, Y. Gan, and M. H. Wang, "Sparse feature fidelity for perceptual image quality assessment," *IEEE Transactions on Image Processing*, vol. 22, no. 10, pp. 4007-4018, 2013.
[31] S. Q. Wang, K. D. Ma, H. Yeganeh, Z. Wang, and W. S. Lin, "A patch-structure representation method for quality assessment of contrast changed images," *IEEE Signal Processing Letters*, vol. 22, no. 12, pp. 2387-2390, 2015.
[32] K. Gu, G. T. Zhai, W. S. Lin, and M. Liu, "The analysis of image contrast: From quality assessment to automatic enhancement," *IEEE Transactions on Cybernetics*, vol. 46, no. 1, pp. 284-297, 2016.
[33] A. K. Moorthy, and A. C. Bovik, "A two-step framework for constructing blind image quality indices," *IEEE Signal Processing Letters*, vol. 17, no. 5, pp. 513-516, 2010.
[34] M. A. Saad, and A. C. Bovik, "Blind image quality assessment: A natural scene statistics approach in the DCT domain," *IEEE Transactions on Image Processing*, vol. 21, no. 8, pp. 3339-3352, 2012.
[35] A. Mittal, A. K. Moorthy, and A. C. Bovik, "No-reference image quality assessment in the spatial domain," *IEEE Transactions on Image Processing*, vol. 21, no. 12, pp. 4695-4708, 2012.
[36] P. Ye, J. Kumar, L. Kang, and D. Doermann, "Unsupervised Feature Learning Framework for No-reference Image Quality Assessment," *Computer Vision and Pattern Recognition*, pp. 1098-1105, 2012.
[37] Y. Zhang, D. M. Chandler, "An algorithm for no-reference image quality assessment based on log-derivative statistics of natural scenes," *The International Society for Optical Engineering Proceedings of SPIE*, vol. 22, no. 4, pp. 1-11, 2013.
[38] A. Mittal, R. Soundararajan, and A. C. Bovik, "Making a 'completely blind' image quality analyzer," *IEEE Signal Processing Letters*, vol. 22, no. 3, pp. 209-212, 2013.
[39] L. Liu, B. Liu, H. Huang, and A. C. Bovik, "No-reference image quality assessment based on spatial and spectral entropies," *Image Communication of Signal Processing*, vol. 29, no. 8, pp. 856-863, 2014.
[40] Y. M. Fang, K. D. Ma, Z. Wang, W. S. Lin, Z. J. Fang, and G. T. Zhai, "No-Reference Quality Assessment of Contrast-Distorted Images Based on Natural Scene Statistics," *IEEE Signal Processing Letters*, vol. 22, no. 7, pp. 838-842, 2015.