EXPLOITING GLOBAL AND LOCAL INFORMATION FOR IMAGE QUALITY ASSESSMENT WITH CONTRAST CHANGE

Haining Gu[†], *Guangtao Zhai*[‡], *Min Liu*[‡], *and Ke Gu*[‡]

[†]Dept. of Info. Engineering and Art Design, Shandong Labor Vocational and Technical College
[‡]Insti. of Image Commu. & Infor. Proce., Shanghai Jiao Tong Univ., Shanghai, China

ABSTRACT

Image quality assessment (IQA) has undergone a booming period during the last decade. Contrast change, being an important type of visual alteration for images, has not been seriously treated yet in the current IQA research. To address this problem, in this paper we propose a new reduced-reference (RR) IQA metric for contrast-changed images by exploiting local similarity information and global statistics information (LAGSI). In order to verify the effectiveness of the proposed algorithm, we test LAGSI with a large number of competitors on the recently introduced CID2013 database as well as contrast change related subsets from TID2008 and CSIQ databases. Experimental results demonstrate the superiority of our approach over classical and state-of-the-art IQA metrics.

***Index Terms*—** Contrast-changed images, image quality assessment (IQA), reduced-reference (RR), local similarity information (LSI), global statistics information (GSI)

I. INTRODUCTION

Image quality assessment (IQA) has been an intensive research topic in recent years due to the fact that perceptual visual quality is of crucial importance for the design and optimization of image/video processing systems, e.g. [1]-[4]. Typically, images undergo various process stages such as acquisition, compression, transmission and presentation before reaching the final observers. As a result, various image distortions (e.g. compression artifacts, blurring, noise injection) occur during the course. To assess the quality of images contaminated by these distortions, researchers have put forward a great many of IQA algorithms, some of which have quite well performances matching to human subjective ratings. On the other hand, images may also subject to contrast changes, presumably because of the unfavorable lighting conditions during capture/representation, the gamma correction process of the display device, or deliberate enhancement/alterations [5]-[6]. Most existing IQA methods, however, work invalidly for those contrast-changed images.

To promote the study of IQA on contrast-changed images, we have lately proposed a dedicated contrast-changed image database (CID2013) [7]. This database contains four hundred images derived from fifteen original pictures from the Kodark database [8] by gray scale transfer using concave, convex, cubic and logistic transfer curves and mean shift. We examine the performance of existing classical and state-of-the-art IQA metrics [9]-[18], including classical peak signal-to-noise ratio (PSNR), structural similarity index (SSIM) [9], and state-of-the-art feature similarity index (FSIM) [13], gradient magnitude similarity deviation (GMSD) [17], cannot work well in assessing the quality of contrast-changed images. This phenomenon is partially because these IQA approaches are largely point-wise based and do not use statistical information about the gray scales and contrast that is usually embedded in the image histogram.

Image contrast is mainly controlled by two factors [19]. The former refers to the local waveform on a pixel by pixel basis, while the latter refers to the distribution of the input image histogram. Accordingly, we in this paper combine the local similarity information (LSI) in pixel-based manner and the global statistics information (GSI) of the whole histogram to derive the RR LAGSI for IQA of contrast change. In the computation of LSI, we use the mutual information, which enjoys the advantage of exploiting information of both pixel values and positions, to measure the statistical dependence between the original and distorted image signals that one variable contains about the other.

In the evaluation of GSI, we explore the first to four order statistics of the image histogram. The first order statistic, i.e. the mean, is representation of the average brightness of an image. According to the optimal contrast-tone mapping (OCTM) formulation [19], the second order statistic, or the variance, of the histogram measures the prospective context-free contrast of an image. Skewness, the third order statistic, was shown to be a deterministic feature for human perception of surface quality [20]. Researches on natural image statistics show that kurtosis, the fourth order statistic, captures some intrinsic scale invariant feature of natural images [21]. In the meantime, entropy of the histogram [22], as a measure of the total amount of information in images, is also adopted as a feature in our metric.

The structure of this paper is arranged as follows. Section 2 introduces contrast related databases/subsets. In Section 3, the features used in our metric are outlined before presenting the proposed LAGIS algorithm. A comparison of our tech-

(a) One original image in CID2013 database and four logistic function changed images.

(b) One original image in TID2008 database and four contrast changed images.

(c) One original image in CSIQ database and four mean shift changed images.

Fig. 1. The example of three pristine images (the leftmost column) and associated contrast-changed images.

nique with a large quantity of IQA metrics are validated on CID2003, TID2008 and CSIQ in Section 4. Finally, Section 5 concludes this paper.

II. CONTRAST-CHANGED IMAGE DATABASES

In order to comprehensively explore the impact of contrast change, we recently proposed a dedicated CID2013 database. This database consists of a total number of 400 images which are generated by using five kinds of transfer functions on fifteen "clear" images of size 720×576 in the Kodak database [8]. These transfer functions include: 1) concave curve; 2) convex curve; 3) cubic curve; 4) logistic curve; 5) mean shifts of thirteen levels [0, ±20, ±40, ±60, ±80, ±100, ±120]. A suggestion given by ITU-R BT.500-12 [23] was provided for twenty-two inexperienced testers to grade every image, and the raw data were collected by an automatic interaction software followed by a series of postprocessing operations according to [7].

Similar to the CID2013 database, there also exist relevant subsets in TID2008 and CSIQ. The former one contains 200 images with mean shift and contrast change, while the latter one is only composed of images with contrast change. Several examples of contrast-changed images in aforementioned three image databases are exhibited in Fig. 1. When looking carefully, it can be found that contrast-changed images are quite distinct from intensively researched image compression, Gaussian blur and white noise, and furthermore, some contrast-changed images have better quality than the original version (labeled with red rectangles), which leads most existing IQA metrics to work ineffectively.

III. THE PROPOSED LAGSI METRIC

III-A. Local similarity information

Our LAGSI metric is constituted by two classes of features, one local LSI feature and five global GSI features. In the computation of LSI, the mutual information is adopted to measure the statistical dependence between the original and distorted images. Specifically, for an input contrast-changed image $\mathbf{y}$, we use four steps as follows.

1) For an image $\mathbf{y}$ with the dimension of $m \times n$, and a window with the size of $(2r+1) \times (2r+1)$, we have the window shift through image $\mathbf{y}$ pixel by pixel horizontally and vertically, then place pixels from the current window into a column vector q_i. Where the window goes, we could get a column vector. By putting these columns together, the final metric Q is generated with $N = (m-2r) \times (n-2r)$ columns, and each column contains $d = (2r+1) \times (2r+1)$ elements.

2) After step one, we get the metric $Q = [q_1, q_2, ...q_N]$. To make every column centered at the origin, we minus the mean from these points:

$$Q_0 = Q - \frac{1}{N} \sum_{i=1}^{N} q_i. \tag{1}$$

3) We calculate the covariance of the points:

$$C_{\mathbf{y}} = \frac{1}{N} Q_0 Q_0^T. \quad (2)$$

4) Assuming Gaussian distribution of the pixels as in [22], we compute entropy of image $\mathbf{y}$ as:

$$Hg(C_{\mathbf{y}}) = \log((2\pi e)^{\frac{d}{2}} \det(C_{\mathbf{y}})^{\frac{1}{2}}). \quad (3)$$

The original image $\mathbf{x}$ can be treated with the same process and get $Hg(C_{\mathbf{x}})$. Then, the LSI feature is defined as

$$\mathrm{LSI} = Hg(C_{\mathbf{x}}) - Hg(C_{\mathbf{y}}). \quad (4)$$

Note that the LSI feature is not totally imagined, but inspired by modifying regional mutual information, a robust similarity measure [24].

III-B. Global statistics information

The GSI features come from the reduced-reference image quality metric for contrast changed images (RIQMC) [7] that has two building blocks: the computation order statistics of histogram and the evaluation of entropy, and their effective aggregation. The first order statistic is mean of an image, and is defined as $E(\mathbf{y})$ for distorted image $\mathbf{y}$. To eliminate outliers, we introduce a Gaussian kernel mapping for the first order statistic, and the information term is computed as:

$$\mathrm{GSI}_1 = \alpha \cdot \exp[-(\frac{E(\mathbf{y}) - \beta}{\gamma})^2] \quad (5)$$

where α, β, γ are model parameters. The second order statistic term for our algorithm is denoted as:

$$\mathrm{GSI}_2 = \sigma^2(\mathbf{y}_h) = E(\mathbf{y}_h^2) - E(\mathbf{y}_h)^2 \quad (6)$$

where $\mathbf{y}_h$ is the histogram of image $\mathbf{y}$. A neural mechanism proposed in [20] suggests that the perceptual surface gloss is related to skewness of the luminance histogram. The calculation of third order statistic, i.e. skewness, is:

$$\mathrm{GSI}_3 = \frac{E[(\mathbf{y} - E(\mathbf{y}))^3]}{\sigma^3(\mathbf{y})}. \quad (7)$$

The kurtosis, or the fourth order statistic of the histogram, is defined as:

$$\mathrm{GSI}_4 = \frac{E[(\mathbf{y} - E(\mathbf{y}))^4]}{\sigma^4(\mathbf{y})} - 3. \quad (8)$$

Finally, as a measure of "business" of an image, the entropy (in bits) of an 8 bit distorted image $\mathbf{y}$ can be defined as $H(\mathbf{y})$:

$$H_{\mathbf{y}} = -\sum_{k=0}^{255} P_k(\mathbf{y}) \cdot \log_2\, P_k(\mathbf{y}) \quad (9)$$

where $P_k(\mathbf{y})$ is the probability distributions of the pixel k in the image $\mathbf{y}$. Identically, the entropy of original image $\mathbf{x}$ is denoted as $H_{\mathbf{x}}$. As a result, we derive $\mathrm{GSI}_5 = H_{\mathbf{x}} - H_{\mathbf{y}}$.

III-C. The LAGSI metric

We finally incorporate the above-mentioned two groups of features together to derive the final LAGSI metric as follows:

$$\mathrm{LAGSI} = \sum_{i=1}^{5} C_i \cdot \mathrm{GSI}_i + \mathrm{LSI} \quad (10)$$

where C_1 to C_5 are model parameters of controlling the relative importance of different information terms. All of these coefficients will be trained on the CID2013 database.

IV. EXPERIMENTAL RESULTS

In this paper we chose some representative IQA metrics for comparison. One group of them is composed of classical PSNR and SSIM [9], and their modified multi-scale SSIM (MS-SSIM) [10], information content weighted SSIM (IW-SSIM) [11], internal generative model (IGM) [15]. The methods in the other group focus on natural scene statistics, gradient, phase and etc. They include most apparent distortion (MAD) [12], feature similarity index (FSIM, FSIMC) [13], gradient similarity metric (GSM) [14] and gradient magnitude similarity deviation (GMSD) [17]. Finally, our previous reduced-reference image quality metric for contrast change (RIQMC) [7] is also used in the comparison.

Four commonly used performance metrics as suggested by VQEG [25] are adopted in our experiment to evaluate the degree of similarity between the subjective and objective results. They are Pearson linear correlation coefficient (PLCC), Spearman rank-order correlation coefficient (SROCC), Kendall's rank-order correlation coefficient (KROCC) and root mean-squared error (RMSE). A four-parameter logistic function is chosen to fit the scores of our method to subjective MOS/DMOS scores

$$Quality(z) = \frac{\eta_1 - \eta_2}{1 + \exp\left(-(z - \eta_3)/\eta_4\right)} + \eta_2 \quad (11)$$

where z is the input score, $Quality(z)$ the mapped score, and η_1 to η_4 are free parameters to be determined during the curve fitting process.

We present the performance evaluations (PLCC, SROCC, KROCC and RMSE) of those above testing IQA algorithms in Table I. It can be easily observed that our LAGSI achieves significantly superior performances than those competitors used. Furthermore, we also notice that, with respect to our recently developed RIQMC than only using the global information, the proposed LAGSI metric has made a noticeable improvement. For a more straightforward comparison, we also provide the scatter plots of MOS vs. the classical SSIM, the state-of-the-art FSIM, GMSD and RIQMC, as well as our proposed LAGSI in Figure. 2. Better fitting effect of the developed metric compared to others can be noticed from these figures.

Table I. PLCC, SROCC, KROCC and RMSE results (after nonlinear regression) of PSNR, SSIM, MS-SSIM, IW-SSIM, MAD, FSIM, FSIMC, GSM, IGM, GMSD, RIQMC and the newly proposed LAGSI algorithm on the CID2013 database (400 images), TID2008 subset (200 images), CSIQ subset (116 images), along with their direct average scores. We bold the best performance results.

	CID2013 database (400 images) [7]				TID2008 database (200 images) [26]			
IQA metrics	PLCC	SROCC	KROCC	RMSE	PLCC	SROCC	KROCC	RMSE
PSNR	0.6504	0.6649	0.4847	0.4733	0.4751	0.5207	0.3640	0.8466
SSIM [9]	0.8072	0.8132	0.6140	0.3678	0.5056	0.4890	0.3417	0.8301
MS-SSIM [10]	0.8494	0.8554	0.6593	0.3289	0.6607	0.5877	0.4303	0.7222
IW-SSIM [11]	0.8756	0.8632	0.6896	0.3010	0.6981	0.4503	0.3475	0.6888
MAD [12]	0.8151	0.8079	0.6174	0.3610	0.3427	0.2828	0.2047	0.9038
FSIM [13]	0.8573	0.8486	0.6662	0.3208	0.6466	0.4403	0.3348	0.7339
FSIMC [13]	0.8575	0.8488	0.6671	0.3206	0.6837	0.4370	0.3326	0.7021
GSM [14]	0.8342	0.8372	0.6371	0.3436	0.6544	0.5126	0.3946	0.7275
IGM [15]	0.8467	0.8244	0.6461	0.3316	0.6924	0.5825	0.4781	0.6941
GMSD [17]	0.8818	0.8803	0.6896	0.2939	0.5685	0.4511	0.3136	0.7915
RIQMC [7]	0.9080	0.9133	0.7343	0.2611	0.7733	0.7316	0.5641	0.6100
LAGSI	**0.9224**	**0.9274**	**0.7572**	**0.2407**	**0.8313**	**0.7866**	**0.6060**	**0.5347**

	CSIQ database (116 images) [27]				Direct average (716 images)			
IQA metrics	PLCC	SROCC	KROCC	RMSE	PLCC	SROCC	KROCC	RMSE
PSNR	0.8987	0.8621	0.6449	0.0739	0.6747	0.6826	0.4979	0.4646
SSIM [9]	0.7437	0.7397	0.5323	0.1126	0.6855	0.6806	0.4960	0.4368
MS-SSIM [10]	0.8956	0.8833	0.6899	0.0749	0.8019	0.7755	0.5932	0.3753
IW-SSIM [11]	0.9497	0.9539	0.8168	0.0527	0.8412	0.7558	0.6180	0.3475
MAD [12]	0.9320	0.9207	0.7460	0.0611	0.6966	0.6704	0.5227	0.4420
FSIM [13]	0.9366	0.9420	0.7883	0.0590	0.8135	0.7436	0.5964	0.3712
FSIMC [13]	0.9385	0.9438	0.7937	0.0582	0.8266	0.7432	0.5978	0.3603
GSM [14]	0.9315	0.9354	0.7721	0.0613	0.8067	0.7617	0.6013	0.3775
IGM [15]	0.9466	0.9547	0.8174	0.0543	0.8286	0.7872	0.6472	0.3600
GMSD [17]	0.9175	0.9039	0.7184	0.0670	0.7893	0.7451	0.5739	0.3841
RIQMC [7]	0.9593	0.9576	0.8258	0.0476	0.8802	0.8675	0.7081	0.3062
LAGSI	**0.9601**	**0.9578**	**0.8273**	**0.0471**	**0.9046**	**0.8906**	**0.7302**	**0.2742**

V. CONCLUSION

In this paper, we propose a novel reduced-reference quality metric LAGSI for contrast-changed images. Our algorithm works by systematically combining the local similarity information in pixel-based manner and the global statistics information of the entire image histogram, which respectively correspond to a pair of important controlling factors in adjusting image contrast. Experimental results show that the metric outperforms several classical and state-of-the-arts on the contrast related CID2013, TID2008 and CSIQ databases.

ACKNOWLEDGMENT

This work was supported in part by NSFC (61025005, 61371146, 61221001, 61390514), FANEDD (201339) and SHEITC (140310).

VI. REFERENCES

[1] G. Zhai, J. Cai, W. Lin, X. Yang, and W. Zhang, “Three dimensional scalable video adaptation via user-end perceptual quality assessment,” *IEEE Trans. Broadcasting*, vol. 54, no. 3, pp. 719-727, September 2008.

[2] G. Zhai, J. Cai, W. Lin, X. Yang, W. Zhang, and M. Etoh, “Cross-dimensional perceptual quality assessment for low bitrate videos,” *IEEE Trans. Multimedia*, vol. 10, no. 7, pp. 1316-1324, November 2008.

[3] K. Gu, G. Zhai, X. Yang, W. Zhang, and C. W. Chen, “Automatic contrast enhancement technology with saliency preservation,” *IEEE Trans. Circuits Syst. Video Technol.*, vol. 25, 2015.

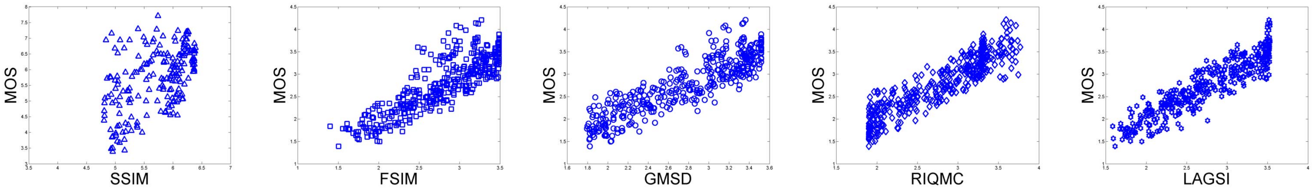
[4] K. Gu, G. Zhai, W. Lin, and M. Liu, “The analysis of image contrast: From quality assessment to automatic enhancement,” *IEEE Trans. Cybernetics*, vol. 45, 2015.

[5] 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.

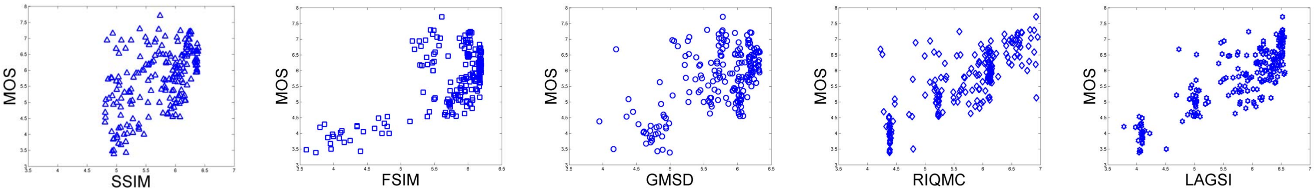
[6] J. A. Stark, “Adaptive image contrast enhancement using generalizations of histogram equalization,” *IEEE Trans. Image Process.*, vol. 9, no. 5, pp. 889-896, May 2000.

[7] K. Gu, G. Zhai, X. Yang, W. Zhang, and Min Liu, “Subjective and objective quality assessment for images with contrast change”, *Proc. IEEE Int. Conf. Image Process.*, 2013.

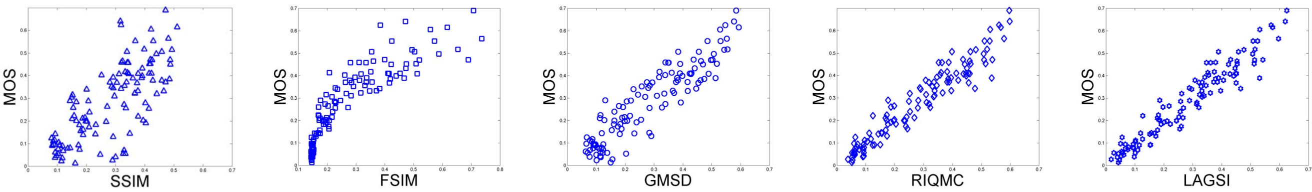
[8] Kodak Lossless True Color Image Suite: http://r0k.us/graphics/kodak/



(a) Five metrics vs MOS values on CID2013 database



(b) Five metrics vs MOS values on contrast subset from TID2008 database



(c) Five metrics vs MOS values on contrast subset from CSIQ database

Fig. 2. The scatter of SSIM, FSIM, GMSD, RIQMC and LAGSI vs. MOS values on CID2013, contrast subsets of TID2008 and CSIQ databases.

[9] 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, April 2004.

[10] Z. Wang, E. P. Simoncelli, and A. C. Bovik, “Multi-scale structural similarity for image quality assessment,” *IEEE Asilomar Conference Signals, Systems and Computers*, November 2003.

[11] Z. Wang and Q. Li, “Information content weighting for perceptual image quality assessment,” *IEEE Trans. Image Process.*, vol. 20, no. 5, pp. 1185-1198, 2011.

[12] E. C. Larson and D. M. Chandler, “Most apparent distortion: full-reference image quality assessment and the role of strategy,” *Journal of Electronic Imaging*, vol. 19, no. 1, 2010.

[13] 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, August 2011.

[14] 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, April 2012.

[15] 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, January 2013.

[16] M. Liu, G. Zhai, K. Gu, Q. Xu, X. Yang, X. Sun, W. Chen, and Y. Zuo, “A new image quality metric based on MIx-Scale transform,” *Proc. IEEE Workshop on Signal Processing Systems*, pp. 266-271, October 2013.

[17] W. Xue, L. Zhang, X. Mou, and A. C. Bovik, “Gradient magnitude similarity deviation: A highly efficient perceptual image quality index,” *IEEE Trans. Image Process.*, vol. 23, no. 2, pp. 684-695, February 2014.

[18] K. Gu, G. Zhai, X. Yang, and W. Zhang, “An efficient color image quality metric with local-tuned-global model,” *in Proc. IEEE Int. Conf. Image Process.*, pp. 506-510, Oct. 2014.

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

[20] 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.

[21] D. Zoran and Y. Weiss, “Scale invariance and noise in natural images,” *in Proceedings of the IEEE 12th International Conference on Computer Vision*, pp. 2209-2216, 2009.

[22] C. E. Shannon, “A mathematical theory of communication,” *Bell Syst. Tech. J.*, vol. 27, pp. 379-423, Oct. 1948.

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

[24] Daniel B. Russakoff, Carlo Tomasi, Torsten Rohlfing, and Calvin R.Maurer, “Image similarity using mutual information of regions,” *Computer Vision-ECCV*, 2004.

[25] VQEG, “Final report from the video quality experts group on the validation of objective models of video quality assessment,” March 2000, http://www.vqeg.org/.

[26] 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,” *Advances of Modern Radioelectronics*, vol. 10, pp. 30-45, 2009.

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