SUBJECTIVE AND OBJECTIVE QUALITY ASSESSMENT FOR IMAGES WITH CONTRAST CHANGE

Ke Gu, Guangtao Zhai, Xiaokang Yang, Wenjun Zhang, and Min Liu

Institute of Image Communication and Information Processing, Shanghai Jiao Tong University, Shanghai, China
Shanghai Key Laboratory of Digital Media Processing and Transmissions

ABSTRACT

It is widely known that, for most natural images, appropriate contrast enhancement can usually lead to improved subjective quality. Despite of its importance to image processing, contrast change has largely been overlooked in the current research of image quality assessment (IQA). To fill this void, in this paper we first report a new and dedicated contrast-changed image database (CID2013). The CID2013 database is composed of four hundred contrast-changed images of fifteen original natural images and the mean opinion scores (MOSs) recorded from twenty-two inexperienced viewers. We then proposed a novel reduced-reference image quality metric for contrast-changed images (RIQMC) using entropies and order statistics of the image histograms. Experimental results on the CID2013, TID2008, and CSIQ databases demonstrate that the proposed RIQMC metric outperforms some mainstream image quality assessment methods.

***Index Terms*—** Contrast-changed image database, mean opinion score (MOS), image quality assessment (IQA), entropy, order statistics

1. INTRODUCTION

The research of image quality assessment (IQA) aims to develop an image quality metric that is well correlated with the human subjective scores. Limited by the dependence on subjective quality databases, most of existing IQA research focuses on compression artifacts, noise injection and blurring. Although contrast enhancement is an active topic in image processing [1]-[4], contrast as an important feature of image, has been largely overlooked in the currently research of IQA. And this may be partially because there lacks of a dedicated and large IQA database of images with contrast-change.

We introduce in this paper a new contrast-changed image database (CID2013)[1] to facilitate the research of contrast related IQA. The CID2013 database includes four hundred contrast-changed versions of fifteen natural images chosen from the Kodak database [5]. Twenty-two inexperienced observers participated the subjective viewing tests designed according to ITU-R BT.500-12 [6] and the mean opinion scores (MOSs) are recorded.

Since the subjective quality test is always laborious and impractical for real-world image processing systems, we further proposed a reduced-reference image quality metric for contrast-changed images (RIQMC) using entropies and order statistics of the image histograms. Entropy [7] measures the average unpredictability of random variables and so we use image entropy as an indicator of image contents.

Order statistics on histograms have widely been used in many contrast related image processing tasks. Obviously, the mean, or the first order statistic of an image, determines the overall brightness of the image histogram [1]. In the recently proposed optimal contrast-tone mapping (OCTM) algorithm [4], the concept of expected context-free contrast is defined as a function of variance (second order statistic) of the histogram. A surface perception model [8] suggests that there exists connection between human perception of surface glossiness and skewness (third order statistic). Some recent work on natural image analysis reveals that kurtosis (fourth order statistic) captures some intrinsic properties of natural images [9]. As inspired by the entropy and order statistics based methods, we design the RIQMC algorithm through linearly combining the entropy and the first four order statistics of the image histogram. The RIQMC method is compared with some well-known benchmark full-reference IQA methods such as PSNR, SSIM [10], MS-SSIM [11], IW-SSIM [12], and MAD [13]. The SSIM algorithm and its classical improved versions are considered quite relevant here because

Fig. 1. The fifteen lossless natural color images in Kodak image database [5].

[1]CID2013 database will be open to the public soon.

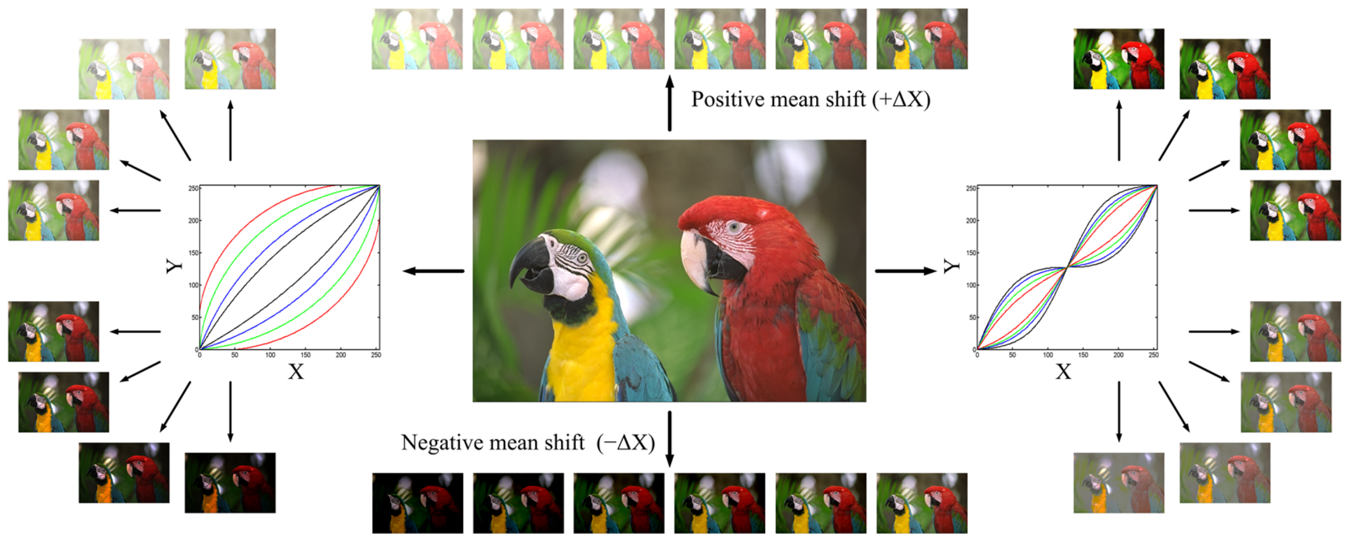

Fig. 2. One example of the "parrot" image and its derived contrast-changed images.

the luminance and contrast similarity components involved in SSIM is believed to quantify contrast change of images.

The remainder of this paper is organized as follows: Section 2 introduces the CID2013 database. Then, the RIQMC metric is outlined in Section 3. In Section 4, experimental results are reported and analyzed. Finally, concluding remarks are given in Section 5.

2. SUBJECTIVE QUALITY ASSESSMENT: THE CID2013 DATABASE

We first chose fifteen natural images of size 720×576 from Kodak image database [5], as illustrated in Fig. 1. Then, a group of more than twenty-five contrast-changed images are generated for each of the fifteen original images, making a total number of 400 images. For each group of contrast-changed images, about sixteen are processed by using the two transfer curves shown in Fig. 2. The other eight to twelve contrast-changed images are produced by mean shifting the original image X, i.e. positive mean shift ($+\triangle X$) or negative mean shift ($-\triangle X$). The shifts $\triangle X$ have six levels of $[20, 40, 60, 80, 100, 120]$. The out of bound values in mean shift are clipped into the range of $0 \sim 255$. To stabilize the test, we also include one or two original images in each group. Fig. 2 presents an example of the "parrot" image and its derived contrast-changed images.

Single-stimulus (SS) method was employed in our subjective experiments. Twenty-two inexperienced subjects (15 males and 7 females) were involved. Most of viewers were college students. We designed an interactive system, as illustrated in Fig. 3, to automatically display the test images and collect the subjective scores. The viewing distance is fixed at three times the image height. The subjects were asked to provide their overall perception of quality on a continuous quality scale from 1 to 5 with the precision up to 1/10000 in a test conditions specified by ITU-R BT.500-12 [6]. The presentation order of the test images for each observer was randomized according to the standard procedure. The subjective scores were then computed for each contrast-changed image after the screening of outliers. Table 1 summarizes the sub-

Table 1. Subjective test conditions and parameters.

Method	Single-stimulus (SS)
Evaluation scales	Continuous quality scale from 1 to 5
Color depth	24-bits/pixel color images
Image coder	Bitmap
Subjects	Twenty-two inexperienced viewers
Image resolution	720×576
Viewing distance	Three times the image height
Room illuminance	Dark

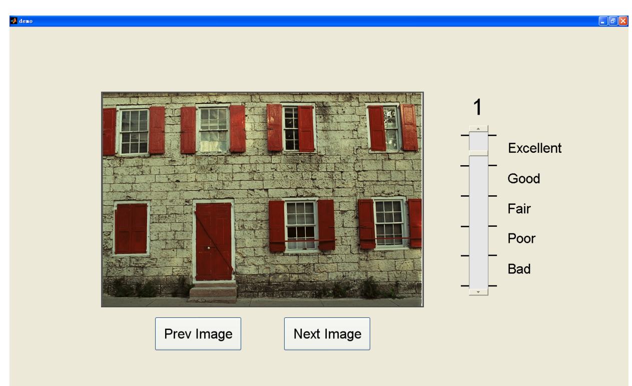

Fig. 3. Illustration of the interactive system used in our subjective viewing experiments.

jective test conditions and parameters stated above, and Fig. 4 shows a histogram of MOS scores.

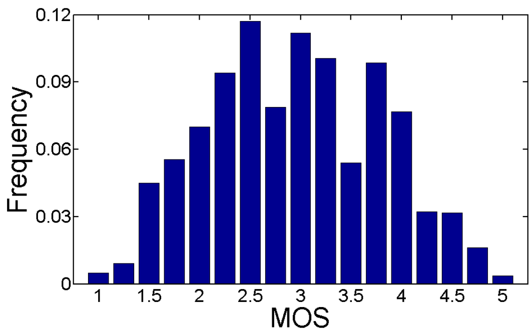

Fig. 4. Histogram of MOS scores.

3. OBJECTIVE QUALITY ASSESSMENT: THE RIQMC METRIC

The proposed RIQMC algorithm is a two-step method: the evaluation of entropy and order statistics, and their linear combination. We first denote the entropy of an image Y as $H(Y)$:

$$H(Y) = -\sum_{i=0}^{255} p_i(Y) \cdot \log_2 p_i(Y) \tag{1}$$

where $p_i(Y)$ indicates the probability density of grayscale i in the image Y. And the entropy of original image X is denoted as $H(X)$ following the same definition. Then, the first order statistic or the mean of an image Y is defined as $E(Y)$. Since we want to punish images with very large or small means, we introduce a Gaussian kernel and design the first order statistic related term as:

$$F_1(Y) = \alpha \cdot exp[-(\frac{E(Y) - \beta}{\gamma})^2] \tag{2}$$

where α, β, γ are model parameters. Besides, as inspired by the concept of expected contrast in OCTM [4], we define the second order statistic term for the RIQMC algorithm as:

$$\begin{aligned} F_2(Y) &= \sigma^2(\mathbf{y}) \\ &= E(\mathbf{y}^2) - E(\mathbf{y})^2 \end{aligned} \tag{3}$$

where $\mathbf{y}$ is the image histogram computed from Y using the "imhist" function in MATLAB. According to the neural mechanism proposed in [8], the on-center and off-center cells and an accelerating nonlinearity in the human visual system compute the subband skewness to estimate the perceptual quality of surface. So, we include the third order statistic (skewness) in the RIQMC metric that is computed as

$$\begin{aligned} F_3(Y) &= skewness(Y) \\ &= \frac{E[(Y - E(Y))^3]}{\sigma^3(Y)}. \end{aligned} \tag{4}$$

Finally, the last fourth order statistic (kurtosis) of the histogram used in the RIQMC algorithm can be evaluated by

$$\begin{aligned} F_4(Y) &= kurtosis(Y) \\ &= \frac{E[(Y - E(Y))^4]}{\sigma^4(Y)} - 3. \end{aligned} \tag{5}$$

In the second step, we adopt a linear combination to integrate the aforementioned entropies and order statistics as

$$RIQMC = F_1 + \mu F_2 + \nu F_3 + \omega F_4 + \kappa[H(Y) - H(X)] \tag{6}$$

where μ, ν, ω, κ, and α, β, γ in Eq. (2) are model parameters. All of the parameters are optimized on a training set including 322 images (about 80%) derived from 12 original natural images in the CID2013 database. The rest 78 images (about 20%) will be used as the test data set.

4. EXPERIMENTAL RESULTS

Mappings of the scores of the six image quality metrics, including PSNR, SSIM [10], MS-SSIM [11], IW-SSIM [12], MAD [13] and the proposed RIQMC, to subjective scores are achieved through using nonlinear regression with a four-parameter logistic function as suggested by VQEG [16]:

$$q(z) = \frac{\lambda_1 - \lambda_2}{1 + exp(\frac{-(z-\lambda_3)}{\lambda_4})} + \lambda_2 \tag{7}$$

with z and $q(z)$ being the input score and the mapped score, respectively. The free parameters λ_1 to λ_4 are to be determined during the curve fitting process.

Three commonly used performance metrics, Pearson Linear Correlation Coefficient (PLCC), Spearman Rank-Order Correlation Coefficient (SROCC) and Root Mean-Squared Error (RMSE) as suggested by VQEG [16], are applied to evaluate the performance of the quality metrics on the CID2013 database, and contrast related subsets in TID2008 and CSIQ. Table 2 tabulates and compares the detailed results. We can easily find that the proposed reduced-reference RIQMC method achieves much better results than other full-reference IQA algorithms.

Fig. 5 shows the scatter plots of the MOS versus the proposed RIQMC metric on the CID2013 database. Among the overall scatter plots, red and blue plots represent the original natural images and their derived contrast-changed images, respectively. It is interesting to find that some "blue" contrast-altered images has higher quality than the "red" natural images. As expected, the quality of natural images may be improved by contrast-alteration, most likely contrast enhancement.

Eventually, it deserves to clarify that since entropy of original image is required as prior information, the proposed RIQMC algorithm is essentially a reduced reference approach. However, it should also be noted that the original image's entropy is just a number, which is usually negligible as compared to the image file's size and can be encoded precisely with only a few bits in the file headers.

5. CONCLUSION

In this paper, we propose a new and dedicated database (CID2013) to facilitate the research of image quality assessment of contrast-changed images. The CID2013 database includes four hundred images derived from fifteen original images and corresponding MOS scores given by twenty-two observers. In addition, a reduced-reference image quality metric for contrast-changed images (RIQMC) is proposed by properly combining entropy and order statistics of image histogram. The proposed RIQMC metric only needs very little reduced-reference information, but outperforms mainstream full-reference PSNR, SSIM, MS-SSIM, IW-SSIM, and MAD algorithms on the CID2013 database, and TID2008 and CSIQ subsets.

Table 2. Pearson Linear Correlation Coefficient (PLCC), Spearman Rank-Order Correlation Coefficient (SROCC) and Root Mean-Squared Error (RMSE) results (after nonlinear regression) of PSNR, SSIM, MS-SSIM, IW-SSIM, MAD and the proposed RIQMC algorithms on the CID2013 database (400 images), TID2008 subsets (200 images), and CSIQ subset (116 images), and their directly average results.

CID2013 database (400 images)			
Algorithm	PLCC	SROCC	RMSE
PSNR	0.6504	0.6649	0.4733
SSIM	0.8072	0.8132	0.3678
MS-SSIM	0.8494	0.8554	0.3289
IW-SSIM	0.8756	0.8632	0.3010
MAD	0.8151	0.8079	0.3610
RIQMC	**0.9080**	**0.9133**	**0.2611**

TID2008 subsets (200 images) [14]			
Algorithm	PLCC	SROCC	RMSE
PSNR	0.4751	0.5207	0.8466
SSIM	0.5056	0.4890	0.8301
MS-SSIM	0.6607	0.5877	0.7222
IW-SSIM	0.6981	0.4503	0.6888
MAD	0.3427	0.2828	0.9038
RIQMC	**0.7733**	**0.7316**	**0.6100**

CSIQ subsets (116 images) [15]			
Algorithm	PLCC	SROCC	RMSE
PSNR	0.8987	0.8621	0.0739
SSIM	0.7437	0.7397	0.1126
MS-SSIM	0.8956	0.8833	0.0749
IW-SSIM	0.9497	0.9539	0.0527
MAD	0.9320	0.9207	0.0611
RIQMC	**0.9593**	**0.9576**	**0.0476**

Directly average results			
Algorithm	PLCC	SROCC	RMSE
PSNR	0.6747	0.6826	0.4646
SSIM	0.6855	0.6806	0.4368
MS-SSIM	0.8019	0.7755	0.3753
IW-SSIM	0.8412	0.7558	0.3475
MAD	0.6966	0.6704	0.4420
RIQMC	**0.8802**	**0.8675**	**0.3062**

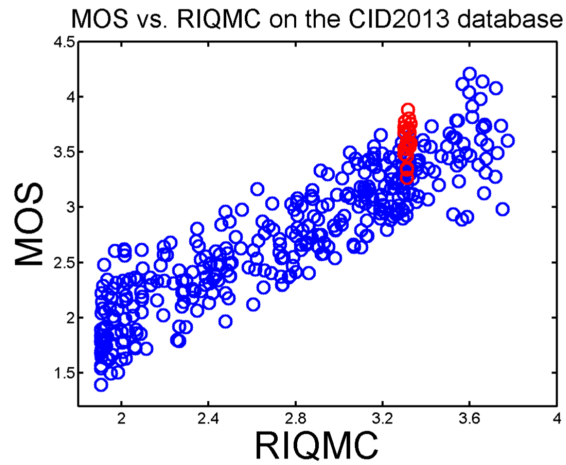

Fig. 5. Scatter plots (after nonlinear regression) of MOS vs. the proposed RIQMC algorithm on our CID2013 database. Red and blue plots represent original natural images and their derived contrast-changed images, respectively.

Acknowledgment

This work was supported in part by NSERC, NSFC (61025005, 60932006, 61001145), SRFDP (20090073110022), postdoctoral foundation of China 20100480603, 201104276, postdoctoral foundation of Shanghai 11R21414200, the 111 Project (B07022), and STCSM (12DZ2272600).

6. REFERENCES

[1] R. C. Gonzalez and R. E. Woods, "Image enhancement in the spatial domain," Digital Image Processing. Addison-Wesley, Reading, Massachusetts, 3rd edition, 2008.

[2] 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.

[3] 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, September 2009.

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

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

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

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

[8] 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.

[9] 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.

[10] 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.

[11] Z. Wang, E. P. Simoncelli, and A. C. Bovik, "Multiscale structural similarity for image quality assessment," *IEEE Asilomar Conference Signals, Systems and Computers*, November 2003.

[12] 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.

[13] 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.

[14] 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.

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

[16] 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/.