FULL-REFERENCE QUALITY ASSESSMENT OF CONTRAST CHANGED IMAGES BASED ON LOCAL LINEAR MODEL

Wen Sun, Wenming Yang, Fei Zhou, Qingmin Liao*

Shenzhen Key Lab. of Info. Sci&Tech/Shenzhen Engineering Lab. of IS&DCP, Tsinghua University, China
Department of Electronic Engineering/Graduate School at Shenzhen, Tsinghua University, China

ABSTRACT

This paper presents a new full-reference method to assess the quality of contrast changed images. In this method, we employ a linear model to describe the relationship between local patches of reference images and contrast changed images. With parameters of this model, three quality measures considering contrast comparison, structure variation, and luminance change are defined. Among them, the first measure produces larger quality scores for higher contrast, which is different from traditional forms of quality measures used in most existing full-reference methods. Experiments on four benchmark databases show that the proposed method is superior to state-of-the-art methods in assessing the quality of contrast changed images.

***Index Terms*—** Image quality assessment, contrast change, local linear model, contrast comparison, structure variation

1. INTRODUCTION

With the popularity of personal phones and cameras, numerous digital images are captured and shared in our daily life. However, due to limitations of imaging devices and environment illumination, images obtained may be low contrast ones, which can not clearly present textures and structures [1]. In view of this, a great number of contrast enhancement algorithms are proposed [2]. Evaluating these algorithms (i.e., assessing the quality of contrast changed images produced by them) is meaningful and necessary. This work focuses on the full-reference (FR) quality assessment of contrast changed images.

FR image quality assessment (IQA) has attracted lots of attention in the last few years, and numerous FR IQA methods are presented. These methods can be roughly categorized into three types: numerical comparison based methods, human visual system (HVS) based methods, and structure/gradient based methods. Peak signal to noise ratio (PSNR) is a method of the first type, it directly computes the intensity differences between reference and test images, which pays on attention to human visual mechanism. HVS based methods usually model some characteristics of HVS, such as near-threshold and supra-threshold properties in visual signal to noise ratio (VSNR [3]) and distortion visibility in most apparent distortion (MAD [4]). Methods of the third type assume that structure/gradient features are more important in human visual perception, and image quality is decided by structure/gradient similarities between reference and test images. Typical examples of the third type include structure similarity (SSIM [5]), feature similarity (FSIM [6]), and gradient magnitude similarity deviation (GMSD [7]). Generally, in these traditional FR IQA methods, the quality of a test image is determined by its feature deviations from the corresponding reference image. Larger deviations indicate worse image quality. However, this is not exactly true for contrast changed images. Images with proper contrast changes may present better clarity and brightness [8] while feature deviations are becoming larger. This problem is also noticed in [9], and a patch-based contrast quality index (PCQI) is then proposed. Nevertheless, PCQI only focuses on achromatic variations while chromatic variations are ignored.

In this paper, a linear model is applied to each pair of patches from reference and test images. By minimizing the total squared error of this linear representation, we can obtain the linear coefficient and intensity deviation. The linear coefficient is used as an index of contrast change. Then, the quality measure regarding contrast is defined in a way that larger quality scores are assigned to higher contrast, which is totally different from traditional FR metrics. In addition, the root mean squared error and intensity deviation are respectively employed to measure structure variation and luminance change. For better performance, color components comparison is also performed. By combining above image quality measures, we can obtain a pixel-wise quality map and then the overall quality score is computed by averaging scores of each pixel. Finally, we validate the proposed method, named as QCCI, by comparing it with state-of-the-art FR methods on four publicly IQA databases.

*Corresponding author: flying.zhou@163.com.

This work was supported by the Natural Science Foundation of China (No. 61771276) and the Special Foundation for the Development of Strategic Emerging Industries of Shenzhen (No. JCYJ20170307153940960).

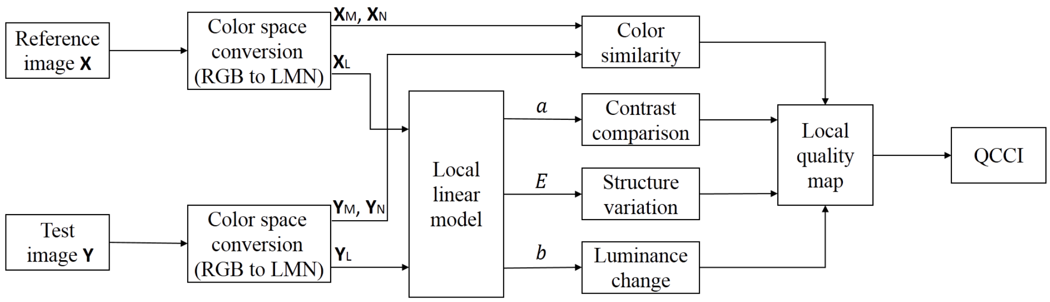

Fig. 1. Framework of the proposed method.

2. LOCAL LINEAR MODEL AND PROPOSED IMAGE QUALITY INDEX

In this section, we describe how the local linear model is applied and how quality measures are calculated. Fig.1 illustrates the main process of the proposed method. The reference image and contrast changed image are firstly converted to LMN color space [10]. Then, their L components are input to a local linear model whose outputs are employed to measure contrast comparison, structure variation, and luminance change. Their chrominance components are compared to obtain color similarity measure. The final quality on each pixel is the product of above measures and the overall image quality is the mean of scores on each pixel.

2.1. Local Linear Model

Though an enhanced (contrast changed) image may have quiet a different global appearance with its reference image, the relationship between local patches of them can be still approximated as linear [11]. Therefore, we can utilize a local linear model [12] to analyze this relationship. Specifically, for two image patches $\mathbf{x}$ (from the reference image) and $\mathbf{y}$ (from the contrast changed image), we can use a linear representation of $\mathbf{x}$

$$\mathbf{z} = a \cdot \mathbf{x} + b \tag{1}$$

to approximate $\mathbf{y}$, where a is the linear coefficient and b is the linear deviation. To calculate a and b, an error function is defined as

$$\begin{aligned} E(a,b) &= \sum_{(i,j)\in\omega} (\mathbf{y}(i,j) - \mathbf{z}(i,j))^2 \\ &= \sum_{(i,j)\in\omega} (\mathbf{y}(i,j) - a \cdot \mathbf{x}(i,j) - b)^2, \end{aligned} \tag{2}$$

where ω denotes the local window covering $\mathbf{x}$, i and j are coordinates of pixels inside ω. Using a similar approach to the one in [11], we can solve this problem by minimizing $E(a,b)$ and then obtain a and b as

$$a = \frac{\frac{1}{|\omega|}\sum_{(i,j)\in\omega} \mathbf{x}(i,j) \cdot \mathbf{y}(i,j) - u_{\mathbf{x}} \cdot u_{\mathbf{y}}}{\sigma_{\mathbf{x}}^2} \tag{3}$$

$$b = u_{\mathbf{y}} - a \cdot u_{\mathbf{x}}, \tag{4}$$

where $|\omega|$ means the number of pixels inside ω, $u_{\mathbf{x}}$ and $u_{\mathbf{y}}$ denote mean values, $\sigma_{\mathbf{x}}$ represents standard deviation. To avoid instability when $\sigma_{\mathbf{x}}$ is too small, we revise (3) as

$$a = \frac{\frac{1}{|\omega|}\sum_{(i,j)\in\omega} \mathbf{x}(i,j) \cdot \mathbf{y}(i,j) - u_{\mathbf{x}} \cdot u_{\mathbf{y}} + \delta}{\sigma_{\mathbf{x}}^2 + \delta}, \tag{5}$$

where δ stands for a small positive number. Here, δ is an important parameter. For larger structures, δ makes no difference to a because the numerator and denominator are larger. For small structures, δ is comparable with the numerator and denominator, and thus it helps to make a closer to 1. Therefore, the usage of δ makes a more sensitive to variations of patches with dominant structures than those with minor structures, which is consistent with characteristics of human visual perception [13].

2.2. Proposed Image Quality Index

By applying the gradient operator ∇ to (1), we have

$$\nabla\mathbf{z} = a \cdot \nabla\mathbf{x}, \tag{6}$$

which describe the relationship between edges in $\mathbf{x}$ and $\mathbf{z}$ as gradient operator can extract local edges. Generally speaking, image contrast is small in plain areas while large in areas with strong edges. In other words, stronger edges usually correspond to larger contrast. Therefore, a in (6) actually conveys the information of contrast comparison between $\mathbf{x}$ and $\mathbf{z}$. Since $\mathbf{z}$ approximates $\mathbf{y}$, we can employ a to measure contrast comparison between $\mathbf{x}$ and $\mathbf{y}$. Here, we define the quality measure of contrast comparison as

$$CC(\mathbf{x}, \mathbf{y}) = f(\lambda a)/f(\lambda), \tag{7}$$

where $f(x)$ is the hyperbolic tangent function [14], λ is a parameter to adjust the influence of a. From the quality measure in (7), we can find that patches showing larger contrast are assigned with larger CC, i.e., better quality. This is because larger contrast usually makes image much clearer. It is worthwhile to note that the quality measure in (7) is totally different from traditional FR methods. Quality measures used in traditional FR methods can be roughly summarized as

$$q(\mathbf{x},\mathbf{y}) = h(|f_{\mathbf{y}} - f_{\mathbf{x}}|) \quad or \quad g(f_{\mathbf{y}}/f_{\mathbf{x}}), \tag{8}$$

where $f_{\mathbf{x}}$ and $f_{\mathbf{y}}$ represent features extracted from patches **x** and **y**, $h(x)$ is a monotonically increasing function, $g(x) = (2x)/(x^2+1)$. For example, image intensities and the mean squared function are respectively used as features and $h(x)$ in PSNR; SSIM, FSIM, and GMSD all adopt $g(x)$ while their features include luminance, standard deviations and gradients. For quality measures in (8), large contrast may result in quality degradations, which is not consistent with human visual perception. To validate this point, we conduct an experiment using CC in (7) (the size of ω is 5×5, $\lambda = 0.7$) and $g(x)$ (the feature is gradient magnitude) respectively as quality measures. They both have the value of '1' when **x=y**. Results are shown in Fig. 2, where Fig. 2(a) and Fig. 2(b) are the reference and the test images, Fig. 2(c) presents the result of CC, and Fig. 2(d) shows the performance of $g(x)$. In Fig. 2(c) and Fig. 2(d), a brighter pixel means better local quality. It can be observed from Fig. 2 that many structures and objects become much clearer as marked by red squares, which show better perceptual appearance. CC well captures these variations and performs positive evaluations (brighter pixels), while $g(x)$ holds negative opinions (darker pixels) about them. This comparison demonstrates that CC is more reasonable and accurate than traditional FR quality measures in predicting the quality of contrast changed images.

In the process of contrast change, local structure patterns may be more or less altered. We can use the minimized error function $E(a,b)$ in (2) to reflect this difference. A larger $E(a,b)$ corresponds to a more different structure pattern. Therefore, we define the quality measure of structure variation as

$$SV(\mathbf{x},\mathbf{y}) = e^{-\alpha\sqrt{E(a,b)/|\omega|}}, \tag{9}$$

where α is a positive parameter.

Luminance represents the brightness perceived by HVS, it is necessary to measure luminance changes. With b in (4), we compute luminance change as

$$LC(\mathbf{x},\mathbf{y}) = e^{-\beta|b|}, \tag{10}$$

where β is a positive parameter. Further, color assessment is also important in IQA [15]. So we convert images to LMN color space [10], and then color similarity is defined as

$$\begin{aligned} MS(\mathbf{x},\mathbf{y}) &= (2\overline{M_{\mathbf{x}}}\cdot\overline{M_{\mathbf{y}}}+T)/(\overline{M_{\mathbf{x}}}^2+\overline{M_{\mathbf{y}}}^2+T) \\ NS(\mathbf{x},\mathbf{y}) &= (2\overline{N_{\mathbf{x}}}\cdot\overline{N_{\mathbf{y}}}+T)/(\overline{N_{\mathbf{x}}}^2+\overline{N_{\mathbf{y}}}^2+T) \\ CS(\mathbf{x},\mathbf{y}) &= MS(\mathbf{x},\mathbf{y})\cdot NS(\mathbf{x},\mathbf{y}), \end{aligned} \tag{11}$$

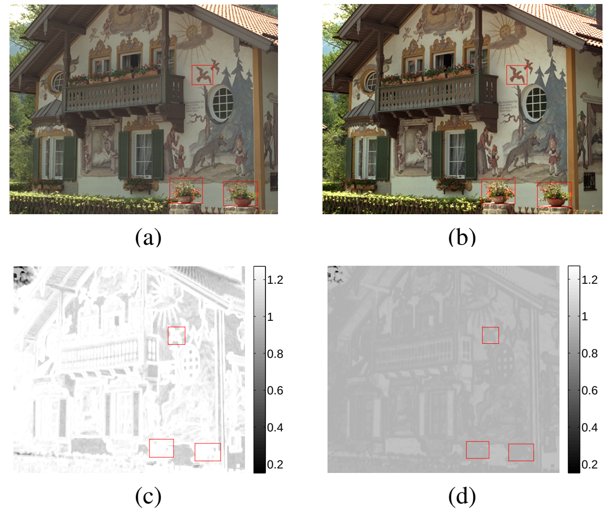

Fig. 2. Comparison of quality measures CC and $g(x)$. (a) Reference image. (b) Contrast changed image. (c) Quality map of CC. (d) Quality map of $g(x)$.

where T is a positive number, $M_{\mathbf{x}}$ and $N_{\mathbf{x}}$ denote M component and N component of patch **x** respectively, $\overline{M_{\mathbf{x}}}$ is the mean of $M_{\mathbf{x}}$, similar meanings for variables reltated to patch **y**.

With four quality measures above, we can obtain the quality of contrast changed image (QCCI) as

$$QCCI(\mathbf{x},\mathbf{y}) = CC(\mathbf{x},\mathbf{y})\cdot SV(\mathbf{x},\mathbf{y})\cdot LC(\mathbf{x},\mathbf{y})\cdot CS(\mathbf{x},\mathbf{y}).$$

Then the overall quality score of the contrast changed image is

$$QCCI = \frac{\sum_{i=1}^{N} QCCI(\mathbf{x}_i,\mathbf{y}_i)}{N}, \tag{12}$$

where N denotes the number of pixels, $\mathbf{x}_i$ and $\mathbf{y}_i$ represent patches centered at the i^{th} pixel in the reference image and test image respectively.

3. EXPERIMENTAL RESULTS AND ANALYSIS

In this section, experiments are performed on four benchmark databases to compare the performance of QCCI with several state-of-the-art FR IQA methods. The four databases, including CSIQ [4], TID2013 [16], CID2013 [17], and CCID2014 [8], are all publicly available and contain contrast changed images. Among them, the first two databases contain various distortion types, we select those images with contrast change to test FR methods. Therefore, 116 images with global contrast decrement in CSIQ and 250 images with mean shift or contrast change in TID2013 are chosen as test images. CID2013 contains 400 test images that are generated by applying different transfer curves and mean shifting to each reference image. It is a database specifically designed for quality

Table 1. Performance comparison of FR IQA methods on four databases

		PSNR	SSIM	VIF	IW-SSIM	FSIMc	GMSD	PCQI	SCQI	QCCI
CCID2014	SROCC	0.6902	0.8136	0.8350	0.7811	0.7657	0.8177	0.8721	0.7812	**0.8880**
	PLCC	0.6874	0.8256	0.8588	0.8342	0.8204	0.8521	0.8869	0.8200	**0.8957**
	RMSE	0.4749	0.3689	0.3350	0.3606	0.3739	0.3422	0.2995	0.3743	**0.2908**
CID2013	SROCC	0.6696	0.8212	0.8687	0.8632	0.8488	0.8803	0.9246	0.8467	**0.9345**
	PLCC	0.6561	0.8219	0.8771	0.8756	0.8577	0.8818	0.9232	0.8489	**0.9293**
	RMSE	0.4703	0.3538	0.2993	0.3010	0.3204	0.2939	0.2394	0.3294	**0.2301**
TID2013	SROCC	0.5230	0.5040	0.7716	0.4528	0.4398	0.4451	**0.8738**	0.4786	0.8733
	PLCC	0.5387	0.5747	0.8458	0.6919	0.6813	0.6753	**0.9175**	0.6633	0.9126
	RMSE	0.8263	0.8027	0.5233	0.7081	0.7180	0.7235	**0.3993**	0.7340	0.4009
CSIQ	SROCC	0.8860	0.7926	0.9345	**0.9539**	0.9438	0.9039	0.9482	0.9329	0.9466
	PLCC	0.9085	0.7901	0.9439	**0.9614**	0.9452	0.9232	0.9488	0.9351	0.9512
	RMSE	0.0704	0.1032	0.0556	**0.0423**	0.0550	0.0647	0.0451	0.0597	0.0444
Weighted Average	SROCC	0.6709	0.7595	0.8414	0.7625	0.7463	0.7768	0.8934	0.7587	**0.9033**
	PLCC	0.6704	0.7775	0.8686	0.8312	0.8166	0.8351	0.9075	0.8099	**0.9127**

assessment of contrast changed images. CCID2014 is an extension of CID2013, it uses more types of transform curves and totally 655 test images are obtained. With the results of FR IQA metrics and human subjective ratings on above images, we can calculate Spearman rank order correlation coefficient (SROCC), Pearson linear correlation coefficient (PLCC), and root mean squared error (RMSE) as performance indices. Thereinto, SROCC measures prediction monotonicity, PLCC and RMSE show the prediction accuracy [18]. Commonly, an excellent IQA method should achieve large SROCC and PLCC while RMSE is small.

We compare QCCI with eight well-known FR algorithms, including PSNR, SSIM [5], VIF [19], IW-SSIM [18], FSIMc [6], GMSD [7], PCQI [9], and SCQI [10]. Among them, PCQI is specifically designed for contrast changed images. Parameters involved in the computation of QCCI are set as: ω is a 5×5 window, $\delta = 5$, $\lambda = 0.7$, $\alpha = -0.005$, $\beta = 1/480$, and $T = 100$. Experimental results are shown in Table 1, where the algorithm obtaining the best performance in each row are highlighted in boldface.

From Table 1, we can find that QCCI has the largest SROCC and PLCC as well as smallest RMSE on CID2013 and CCID2014, which indicates that QCCI achieves the best outcomes on these two specialized databases. On the other two databases, the proposed method is also among the top three methods that correlate most consistently with human subjective ratings. In the bottom of Table 1, the weighted average indices are provided, where the weight is defined as the number of test images in each database. The weighted results show that our method obtains the best overall performance. To further compare these FR algorithms, we performed statistical significance tests [7] on the largest database CCID2014. The results are shown in Fig. 3, where a value of ‘1’ (highlighted in green) indicates that the method in the row is statistically better than the one in the column and ‘0’ (highlighted in red) otherwise. It can be observed that QCCI is statistically superior to almost all other methods.

CCID 2014	PSNR	SSIM	VIF	IWSSIM	FSIMc	GMSD	PCQI	SCQI	QCCI
PSNR	0	0	0	0	0	0	0	0	0
SSIM	1	0	0	0	0	0	0	0	0
VIF	1	1	0	1	1	0	0	1	0
IWSSIM	1	0	0	0	1	0	0	0	0
FSIMc	1	0	0	0	0	0	0	0	0
GMSD	1	1	0	1	1	0	0	1	0
PCQI	1	1	1	1	1	1	0	1	0
SCQI	1	0	0	0	0	0	0	0	0
QCCI	1	1	1	1	1	1	0	1	0

Fig. 3. Results of statistical significance tests of the competing IQA approaches on CCID2014.

4. CONCLUSION

Based on the observation that contrast changed images are patch-wise linear with reference images, we propose a FR method specialized for evaluating the quality of contrast changed images based on local linear model. With this model, we assess image quality from three aspects: contrast comparison, structure variation, and luminance change. Thereinto, the quality measure of contrast comparison gives larger quality scores to patches with higher contrast, which is consistent with human visual perception about image quality. Further, color component similarity is also considered and incorporated into the proposed method. Experimental results on four databases demonstrate that the proposed method outperforms most existing method in assessing the quality of contrast changed images.

5. REFERENCES

[1] T. Celik, “Spatial entropy-based global and local image contrast enhancement.,” *IEEE Transactions on Image Processing*, vol. 23, no. 12, pp. 5298–308, 2014.

[2] D. Raj and P. Mamoria, “Comparative analysis of contrast enhancement techniques on different images,” in *International Conference on Green Computing and Internet of Things*, 2016, pp. 27–31.

[3] D. M. Chandler and S. S. Hemami, “VSNR: a wavelet-based visual signal-to-noise ratio for natural images,” *IEEE Transactions on Image Processing*, vol. 16, no. 9, pp. 2284–2298, 2007.

[4] 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, pp. 011006, 2010.

[5] 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.

[6] L. Zhang, L. Zhang, X. 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.

[7] W. Xue, L. Zhang, X. 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.

[8] K. Gu, G. Zhai, W. 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.

[9] S. Wang, K. Ma, H. Yeganeh, Z. Wang, and W. 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.

[10] S. H. Bae and M. Kim, “A novel image quality assessment with globally and locally consilient visual quality perception,” *IEEE Transactions on Image Processing*, vol. 25, no. 5, pp. 2392–2406, 2016.

[11] H. Wang, F. Jie, W. Lin, S. Hu, C. C. Jay Kuo, and L. Zuo, “Image quality assessment based on local linear information and distortion-specific compensation,” *IEEE Transactions on Image Processing*, vol. 26, no. 2, pp. 915–926, 2017.

[12] S. Roweis, L. K. Saul, and G. E. Hinton, “Global coordination of local linear models,” *Advances in Neural Information Processing Systems*, vol. 14, pp. 889–896, 2001.

[13] D. Li, H. Huang, and Z. Yu, “Image quality assessment using directional anisotropy structure measurement,” *IEEE Transactions on Image Processing*, vol. 26, no. 4, pp. 1799–1809, 2017.

[14] F. Xiao, Y. Honma, and T. Kono, “A simple algebraic interface capturing scheme using hyperbolic tangent function,” *International Journal for Numerical Methods in Fluids*, vol. 48, no. 9, pp. 1023–1040, 2005.

[15] D. Liu, F. Li, and H. Song, “Image quality assessment using regularity of color distribution,” *IEEE Access*, vol. 4, pp. 4478–4483, 2016.

[16] N. Ponomarenko, O. Ieremeiev, V. Lukin, K. Egiazarian, L. Jin, J. Astola, B. Vozel, K. Chehdi, M. Carli, and F. Battisti, “Color image database TID2013: Peculiarities and preliminary results,” in *European Workshop on Visual Information Processing*, 2013, pp. 106–111.

[17] K. Gu, G. Zhai, X. Yang, W. Zhang, and M. Liu, “Subjective and objective quality assessment for images with contrast change,” in *IEEE International Conference on Image Processing*, 2014, pp. 383–387.

[18] Z. Wang and Q. Li, “Information content weighting for perceptual image quality assessment,” *IEEE Transactions on Image Processing*, vol. 20, no. 5, pp. 1185–1198, 2011.

[19] 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.