Automatic Contrast Enhancement Technology With Saliency Preservation

Ke Gu, Guangtao Zhai, *Member, IEEE*, Xiaokang Yang, *Senior Member, IEEE*, Wenjun Zhang, *Fellow, IEEE*, and Chang Wen Chen, *Fellow, IEEE*

***Abstract*— In this paper, we investigate the problem of image contrast enhancement. Most existing relevant technologies often suffer from the drawback of excessive enhancement, thereby introducing noise/artifacts and changing visual attention regions. One frequently used solution is manual parameter tuning, which is, however, impractical for most applications since it is labor intensive and time consuming. In this research, we find that saliency preservation can help produce appropriately enhanced images, i.e., improved contrast without annoying artifacts. We therefore design an automatic contrast enhancement technology with a complete histogram modification framework and an automatic parameter selector. This framework combines the original image, its histogram equalized product, and its visually pleasing version created by a sigmoid transfer function that was developed in our recent work. Then, a visual quality judging criterion is developed based on the concept of saliency preservation, which assists the automatic parameters selection, and finally properly enhanced image can be generated accordingly. We test the proposed scheme on Kodak and Video Quality Experts Group databases, and compare with the classical histogram equalization technique and its variations as well as state-of-the-art contrast enhancement approaches. The experimental results demonstrate that our technique has superior saliency preservation ability and outstanding enhancement effect.**

***Index Terms*— Contrast enhancement, histogram modification framework (HMF), quality assessment (QA), saliency preservation, sigmoid transfer mapping.**

I. Introduction

DUE to user's operational error, poor illumination condition, and unideal device functionality, a raw image can sometimes have limited contrast and low visual quality. To solve the problem, various postprocessing algorithms have been proposed, such as contrast enhancement, white balance adjustment, dynamic range expansion, and edge sharpening or high boosting. Contrast enhancement is usually a preferable option because it aims at directly improving the image contrast, and thereby enhancing users' experiences.

Manuscript received April 11, 2014; revised September 16, 2014; accepted November 11, 2014. Date of publication November 20, 2014; date of current version September 1, 2015. This work was supported in part by the National Science Foundation of China under Grant 61025005, Grant 61371146, Grant 61221001, and Grant 61390514; in part by the Foundation for the Author of National Excellent Doctoral Dissertation of China under Grant 201339; and in part by the Major State Basic Research Development Program (973 Program) of China under Grant 2010CB731401. This paper was recommended by Associate Editor M. H. Hayes.

K. Gu, G. Zhai, X. Yang, and W. Zhang are with the Shanghai Key Laboratory of Digital Media Processing and Transmissions, Institute of Image Communication and Information Processing, Shanghai Jiao Tong University, Shanghai 200240, China (e-mail: guke.doctor@gmail.com; zhaiguangtao@sjtu.edu.cn; xkyang@sjtu.edu.cn; zhangwenjun@sjtu.edu.cn).

C. W. Chen is with the Department of Computer Science and Engineering, State University of New York at Buffalo, Buffalo, NY 14260 USA (e-mail: chencw@buffalo.edu).

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/TCSVT.2014.2372392

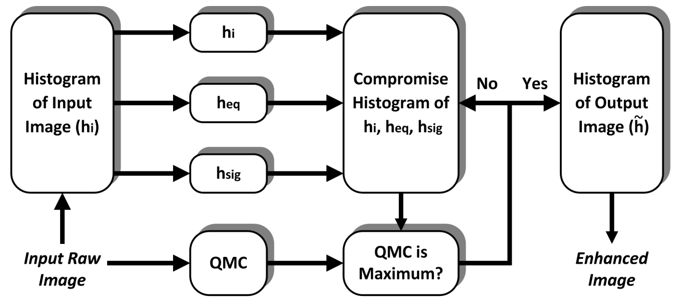


Fig. 1. Flowchart of the proposed RICE contrast enhancement algorithm. $\mathbf{h_i}$, $\mathbf{h_{eq}}$, and $\mathbf{h_{sig}}$ separately represent the histograms of the input image, and associated HE and STBP processed versions.

Contrast enhancement has been an important research topic in image processing and computer vision for a long history. Generally speaking, contrast enhancement targets to generate a perceptually more pleasing or visually more informative image or both. By judiciously reassigning pixel values in an image, the contrast can be drastically improved, as practiced by histogram equalization (HE) [1]. The fundamental objective of HE is to maximize the entropy of the image histogram, so as to reveal image details as much as possible. Owing to its simplicity and quickness, HE has nowadays been widely used in many image postprocessing systems and is the *de facto* synonym for contrast enhancement. However, HE is often questioned for excessive enhancement that can cause serious visible deterioration, such as contouring or ringing. Researchers now tend to agree that HE is far from the ideal contrast enhancement technology, and many attempted to improve HE for better performance.

An important type of solutions to overcome the drawback of overenhancement of HE is to preserve the input image brightness when using HE. Early methods, such as brightness preserving bi-HE (BBHE) [2] and dualistic subimage HE (DSIHE) [3], decompose the input image histogram into dualistic subhistograms, and then apply HE in each subhistogram. Their major difference is that the decomposition step of BBHE relies on mean image brightness, while DSIHE uses median value. Recursive mean-separate HE [4] and recursive subimage HE (RSIHE) [5] adopt similar recursive

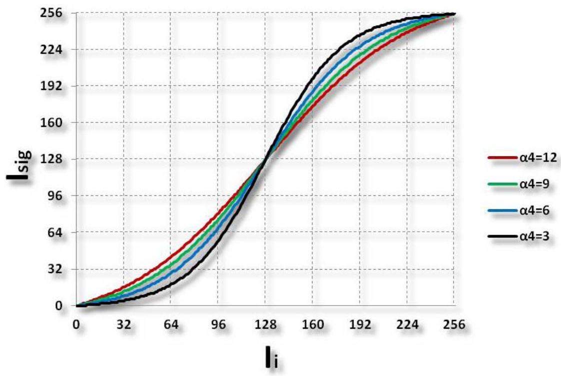
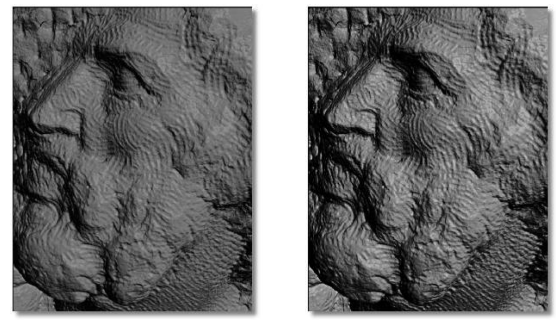


Fig. 2. Sigmoid curves with four different α_4 values.



(a) (b)

Fig. 3. Matthew sculpture (skewness: -0.62) and its processed version (skewness: -0.13) using the proposed sigmoid transfer mapping. (a) Original sculpture. (b) Processed sculpture.

operations to improve BBHE and DSIHE, to maintain the brightness better. Another example is weighted thresholded HE (WTHE) [6], which modifies image histogram by weighting and thresholding before HE. Thereafter, the concept of dynamic range was introduced into contrast enhancement in [7] and [8] using HE in every subhistogram toward a new dynamic range. Ibrahim and Kong [9], [10] proposed two algorithms for enhancing gray and color images with a normalization stage to keep the original brightness. Recently, a modified Laplacian pyramid framework was proposed to partition the input image into bandpass images, followed by a novel robust HE for global contrast enhancement with noise control and local information preservation [11].

Another direction to improve HE is to pose contrast enhancement as an optimization problem that can be solved by minimizing a cost function. Lai *et al.* [12] adopted a quality measure based preprocessing step to reassign the probability distribution. Then, they applied an improved plateau HE to adaptive contrast enhancement in terms of the regulated probability distribution function. In [13], a histogram modification framework (HMF) was designed in two steps: first, to search for an intermediate histogram $\mathbf{h}$ between the input histogram $\mathbf{h}_i$ and the uniform histogram $\mathbf{u}$ by minimizing a weighted distance $||\mathbf{h} - \mathbf{h}_i|| + \lambda||\mathbf{h} - \mathbf{u}||$; second, to perform HE of $\mathbf{h}$. The HMF is able to indirectly restrain undesirable side effects of HE via the proper selection of the Lagrangian multiplier λ.

Other contrast enhancement methods, similar to [12] and [13], were proposed with optimization processes as well. Majumder and Irani [14] improved the local image contrast by controlling the local gradient with a single parameter. Without segmentation, this method tends to maximize the average local contrast of the input image strictly subject to a Weber Law-induced perceptual constraint. Finally, the optimal contrast-tone mapping (OCTM) was designed using linear programming to solve an optimization problem posed by a formal definition of image contrast and tone distortion [15]. OCTM has succeeded in seeking the compromise of two conflicting quality criteria (tone subtlety and contrast enhancement), which was overlooked in previous work, and it permits users to add and fine tune the constraints to obtain desirable visual effects.

Despite of the surge of contrast enhancement approaches, the automatic quality judging or parameter selection criterion is rarely seen. As a consequence, overenhancement and underenhancement are still a big challenge for most existing technologies, such as [2]–[6]. To obtain better results, people have to resort to manual parameter tuning, which is often a labor intensive and time-consuming job. We, in this paper, focus on addressing the mentioned difficulty. It is believed that a good contrast enhancement algorithm should highlight indiscernible image details and suppress visual artifacts simultaneously. In our research, we noticed that image saliency is sensitive to noise injection whereas immune to contrast enhancement. Therefore, it is reasonable to use saliency preservation as an effective judging criterion to ensure a properly enhanced image. On this base, we propose an automatic robust image contrast enhancement (RICE) model with saliency preservation.

In the design of RICE, it is assumed that the ideal histogram of an appropriately enhanced image should be: 1) close to the uniformly distributed histogram to enhance image informativeness; 2) keeping its 2-norm distance from the original image histogram small to reduce the visual deterioration; and 3) of positively skewed statistics to improve the surface quality [16], e.g., using the recently proposed S-shaped transfer function based brightness preserving (STBP) algorithm [17]. To comprehensively meet the three requirements above, an optimization problem is formulated using the sum of weighted histograms of the input image, and its HE and STBP processed ones. Then, a criterion for parameter selection is established by the idea of saliency preservation, which is measured by a quality assessment (QA) metric of contrast (dubbed as QMC) using the image signature model for saliency detection [18].

The rest of this paper is organized as follows. Section II first describes the proposed RICE algorithm. We conduct RICE, and classical and state-of-the-art contrast enhancement approaches on Kodak database [19] in Section III, and the experimental results confirm that RICE works better than the competing methods. In Section IV, an extension of the proposed RICE to video enhancement is given. Section V concludes this paper.

II. Automatic Enhancement Technology

The proposed algorithm works in two stages: 1) to pose the cost function regarding the ideal histogram and

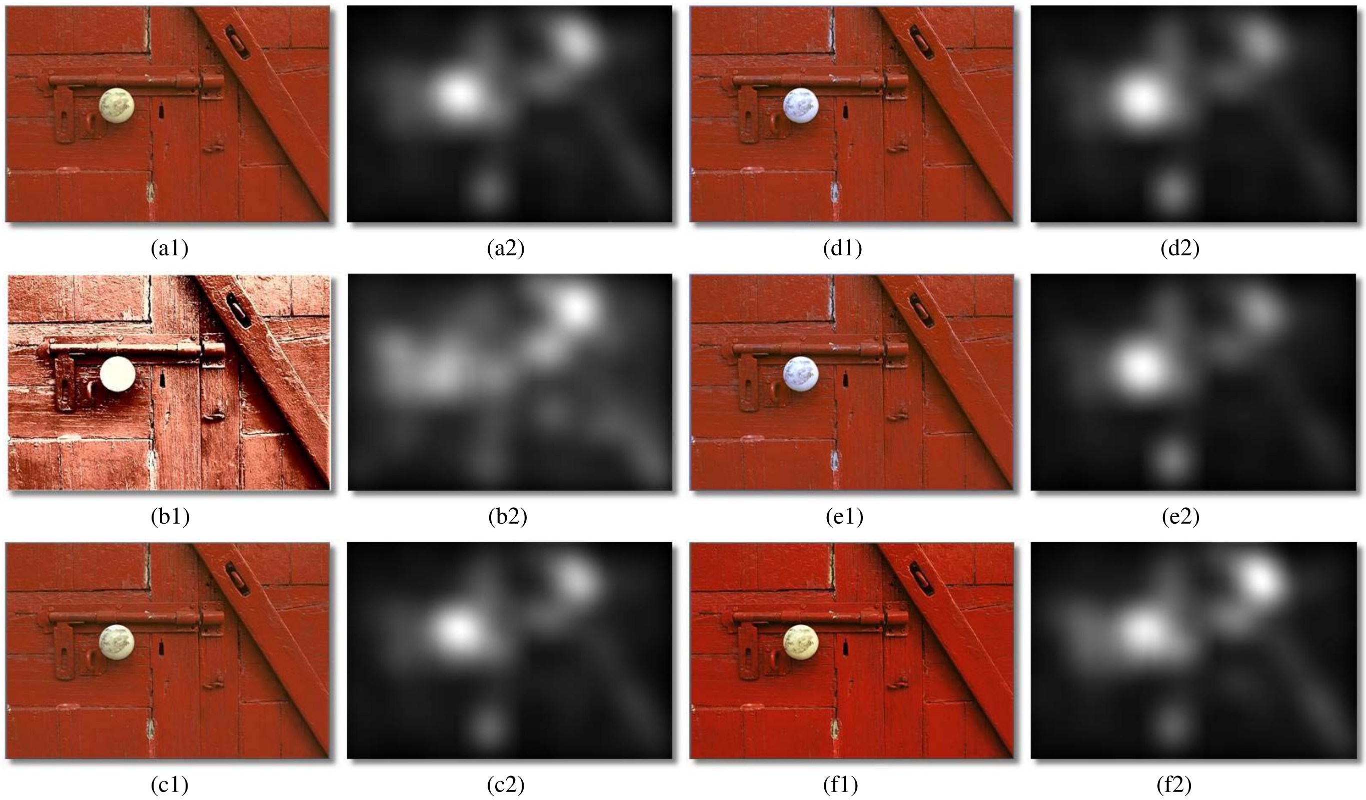


Fig. 4. (a1) Natural image red door in the Kodak database. (b1) Output of HE. (c1) Output of STBP. (d1)–(f1) Outputs of (8) and (9) with $\{\phi, \psi\} = \{1e-4, 0.02\}, \{1e-4, 0.2\}$, and $\{1e-3, 0.2\}$. (a2)–(f2) Saliency maps of (a1)–(f1) using (10)–(12).

2) to automatically obtain the ideal histogram following the instruction of QMC, and then enhance image contrast by histogram matching. We show the flowchart of RICE in Fig. 1 for easy understanding of the proposed framework.

A. Ideal Histogram for Contrast Enhancement

In earlier studies, researchers seek to fully exploit the available dynamic range for contrast enhancement. HE is such a classical method that aims at generating a uniformly distributed histogram with a cumulated histogram as its mapping function. HE can increase image informativeness, and sometimes produce output images with fairly well result. However, HE also suffers from many criticisms, because it tends to easily cause visible deterioration due to overenhancement. As a consequence, various improved HE-type of methods have been proposed up to date. These approaches, however, do not always guarantee satisfactory outputs. In this paper, we define a general HMF inspired by a recent work in [13].

For an input image I_i, we first denote by $\mathbf{h_i}$ the histogram of I_i and by $\mathbf{h_u}$ a uniformly distributed histogram. We then pose a bicriteria optimization problem based on the supposition that the target histogram $\tilde{\mathbf{h}}$ should be closer to $\mathbf{h_u}$ as required by the task of enhancement, but also keep the distance $\tilde{\mathbf{h}} - \mathbf{h_i}$ small as a fidelity constraint. That is to say, the goal histogram is expected to be more visually informative yet with minimum perceptual deterioration. In practice, we find that $\mathbf{h_u}$ is not a good choice since most image histograms cannot be distributed uniformly after HE on account of various kinds of image scenes. This inspires us to replace $\mathbf{h_u}$ with the equalized histogram $\mathbf{h_{eq}}$ that is computed from $\mathbf{h_i}$ using HE. We therefore formulate the optimization problem as a weighted sum of the following two objectives:

$$\tilde{\mathbf{h}} = \arg\min_{\mathbf{h}} \|\mathbf{h} - \mathbf{h_i}\| + \phi\|\mathbf{h} - \mathbf{h_{eq}}\| \tag{1}$$

where $\tilde{\mathbf{h}}$, $\mathbf{h}$, $\mathbf{h_i}$, $\mathbf{h_{eq}} \in \mathbb{R}^{256\times1}$, and ϕ is a control parameter varying over $[0, \infty)$. Note that the solution of (1) finds the optimal tradeoff between two histograms of the original image and its histogram equalized version. The standard HE can be acquired as ϕ goes to infinity, while (1) converges to the input image when ϕ is close to zero.

It is easy to find that (1) does not involve any perceptual quality related term. In [17], a properly defined sigmoid transfer function was shown to produce perceptually pleasing images, resulting in substantial visual quality improvement. More precisely, we use a four-parameter logistic function to define the sigmoid transfer mapping $T_{\text{sig}}(\cdot)$ and its associated enhanced image I_{sig} as

$$I_{\text{sig}} = T_{\text{sig}}(I_i, \boldsymbol{\pi}) = \frac{\pi_1 - \pi_2}{1 + \exp\left(-\frac{(I_i - \pi_3)}{\pi_4}\right)} + \pi_2 \tag{2}$$

where $\boldsymbol{\pi} = \{\pi_1, \pi_2, \pi_3, \pi_4\}$ are free parameters required to be solved. We assume that the transfer curve passes four points (β_i, α_i), $i = \{1, 2, 3, 4\}$. Motoyoshi *et al.* [16] found that an image of a long positive tail in histogram (namely a positively

TABLE I
PERFORMANCE MEASURES (SROCC) AND COMPUTATIONAL LOAD (AVERAGE RUN TIME) OF STATE-OF-THE-ART FR FSIM, GSIM, AND IGM, AND RR FEDM, SDM, RIQMC, AND THE PROPOSED RR QMC. WE BOLD THE METRIC WITH THE BEST PERFORMANCE AND THE LEAST COMPUTATIONAL TIME

FR QA metrics	FSIM	GSIM	IGM	SW-SSIM
SROCC	0.8486	0.8372	0.8244	0.8344
Run Time (second)	0.6799	0.0469	18.884	18.342
RR QA metrics	FEDM	SDM	RIQMC	QMC
SROCC	0.7271	0.6145	0.9133	**0.9335**
Run Time (second)	85.698	0.6419	0.1030	**0.0184**

TABLE II
PERFORMANCE EVALUATIONS (SROCC) OF THE TESTING QA METRICS ON EACH PRISTINE IMAGE AND CORRESPONDING CONTRAST-CHANGE IMAGES. WE EMPHASIZE THE BEST PERFORMED METRIC AND LABEL THE LOWEST SCORE WITH BRACKETS FOR EACH QA METRIC

j-th	FSIM	GSIM	IGM	FEDM	SDM	RIQMC	QMC
01	0.797	0.841	(0.750)	0.652	0.666	0.956	**0.968**
02	0.894	0.870	0.871	0.826	0.733	0.900	(**0.927**)
03	0.879	0.858	0.841	(0.591)	0.373	0.905	**0.944**
04	0.833	0.847	0.809	0.706	0.719	0.890	**0.929**
05	0.799	0.866	0.764	0.594	0.632	0.931	**0.942**
06	**0.943**	0.910	0.908	0.836	(0.273)	0.890	(0.927)
07	0.830	0.843	0.798	0.710	0.603	(0.868)	**0.930**
08	0.910	0.890	0.908	0.840	0.385	0.945	**0.975**
09	0.922	0.938	0.905	0.816	0.592	0.926	**0.952**
10	0.935	0.900	0.945	**0.983**	0.769	0.918	(0.927)
11	0.915	0.878	0.902	0.803	0.768	0.898	**0.934**
12	0.924	0.932	0.864	0.778	0.647	0.906	**0.938**
13	(0.796)	0.797	0.751	0.647	0.764	0.949	**0.964**
14	0.822	(0.772)	0.813	0.743	0.799	0.933	**0.952**
15	0.833	0.845	0.795	0.721	0.592	0.940	**0.958**

skewed statistics) always tends to appear darker and glossier and has better surface quality than a similar image with lower skewness. Furthermore, the authors also provided a possible neural mechanism in human brains, which includes ON-center and OFF-center cells and an accelerating nonlinearity to compute the subband skewness. This motivates the usage of the sigmoid mapping for advancing surface quality, which is rolling symmetry with respect to the straight line $y = x$. We fix seven parameters: $(\beta_1, \alpha_1) = (0, 0)$, $(\beta_2, \alpha_2) = (255, 255)$, $(\beta_3, \alpha_3) = (x, y)$, where $x = y = \lceil \text{mean}(I_i)/32 \rceil * 32$, $\beta_4 = 25$, and let α_4 to be the unique free parameter. We then search for the optimal control parameters $\boldsymbol{\pi} = \{\pi_1, \pi_2, \pi_3, \pi_4\}$ by minimizing the following objective function:

$$\boldsymbol{\pi}_{\text{opt}} = \arg\min_{\boldsymbol{\pi}} \sum_{i=1}^{4} |\alpha_i - T_{\text{sig}}(\beta_i, \boldsymbol{\pi})|. \tag{3}$$

With the known parameters $\boldsymbol{\pi}_{\text{opt}}$, we can finally get

$$I_{\text{sig}} = \max(\min(T_{\text{sig}}(I_i, \boldsymbol{\pi}_{\text{opt}}), 255), 0) \tag{4}$$

where max and min operations are used to limit I_{sig}'s pixel values in the bound of 0–255. Note that α_4 is the only control

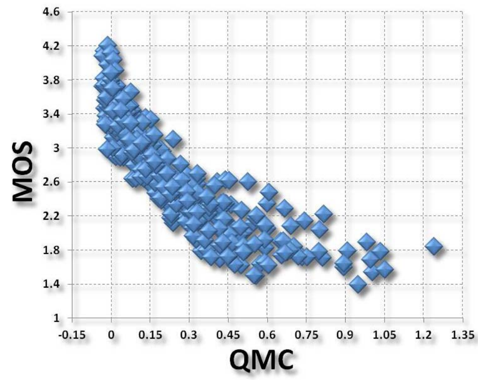


Fig. 5. Scatter plot of MOS versus the proposed QMC on the overall CID2013 database.

parameter to alter curvature of the transfer function. In this paper, we set $\alpha_4 = 12$. To visualize the sigmoid curve, we plot four exemplary curves with the same $(\beta_3, \alpha_3) = (128, 128)$ but different α_4 in Fig. 2.

The proposed sigmoid transfer mapping is used to process the Matthew sculpture image, as shown in Fig. 3, and this clearly increases the surface quality in comparison to the original counterpart. Furthermore, we present a classical natural image red door as well as its histogram equalized and sigmoid curve transferred versions in Fig. 4(a1), (b1), and (c1). We can readily find that the sigmoid mapping produces perceptually pleasing images (c1) with respect to the other two (a1) and (b1). It is natural to combine the histogram $\mathbf{h_{sig}}$ that is computed from I_{sig} into (1), thus making the optimization objective function more complete

$$\tilde{\mathbf{h}} = \arg\min_{\mathbf{h}} \|\mathbf{h} - \mathbf{h_i}\| + \phi\|\mathbf{h} - \mathbf{h_{eq}}\| + \psi\|\mathbf{h} - \mathbf{h_{sig}}\| \tag{5}$$

where $\mathbf{h_{sig}} \in \mathbb{R}^{256\times1}$, and ψ is the second control parameter similar to ϕ. Note that, with different choices of $\{\phi, \psi\}$, the solution of (5) will create the original input image, or its histogram equalized output, or the sigmoid transferred copy. Of course, a proper selection of $\{\phi, \psi\}$ will lead to the best tradeoff and generate optimally enhanced images.

To simplify the optimization equation stated above, we use the squared sum of the Euclidean norm to obtain an analytical solution to (5)

$$\tilde{\mathbf{h}} = \arg\min_{\mathbf{h}} \|\mathbf{h} - \mathbf{h_i}\|_2^2 + \phi\|\mathbf{h} - \mathbf{h_{eq}}\|_2^2 + \psi\|\mathbf{h} - \mathbf{h_{sig}}\|_2^2 \tag{6}$$

which results in the quadratic optimization problem

$$\tilde{\mathbf{h}} = \arg\min_{\mathbf{h}} \big[(\mathbf{h} - \mathbf{h_i})^T(\mathbf{h} - \mathbf{h_i}) + \phi(\mathbf{h} - \mathbf{h_{eq}})^T(\mathbf{h} - \mathbf{h_{eq}}) + \psi(\mathbf{h} - \mathbf{h_{sig}})^T(\mathbf{h} - \mathbf{h_{sig}})\big]. \tag{7}$$

By derivation, we can derive the solution of (7) as

$$\tilde{\mathbf{h}} = \frac{\mathbf{h_i} + \phi\mathbf{h_{eq}} + \psi\mathbf{h_{sig}}}{1 + \phi + \psi}. \tag{8}$$

Given $\tilde{\mathbf{h}}$, the histogram matching function $T_{\text{hm}}(\cdot)$ given in [13] is used to produce the enhanced image

$$\tilde{I} = T_{\text{hm}}(I_i, \tilde{\mathbf{h}}(\phi, \psi)). \tag{9}$$

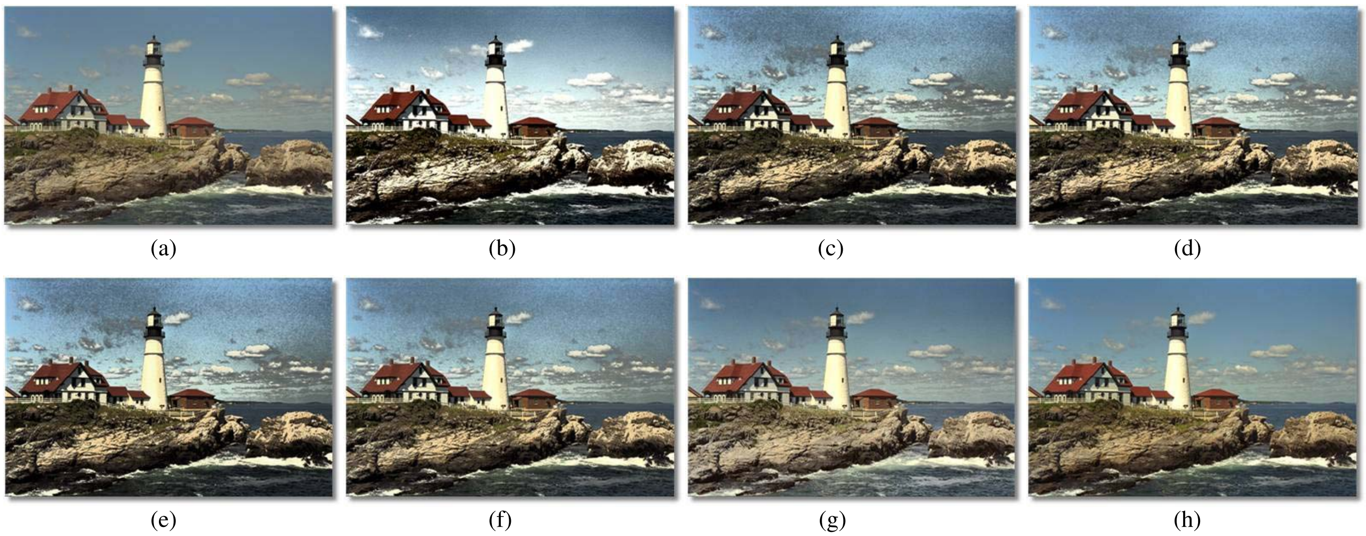


Fig. 6. Natural image Portland Head Light in Kodak database and the outputs. (a) Original image. (b) Output of HE. (c) Output of DSIHE [3]. (d) Output of RSIHE [5]. (e) Output of WTHE [6]. (f) Output of HMF [13]. (g) Output of OCTM [15]. (h) Output of our RICE.

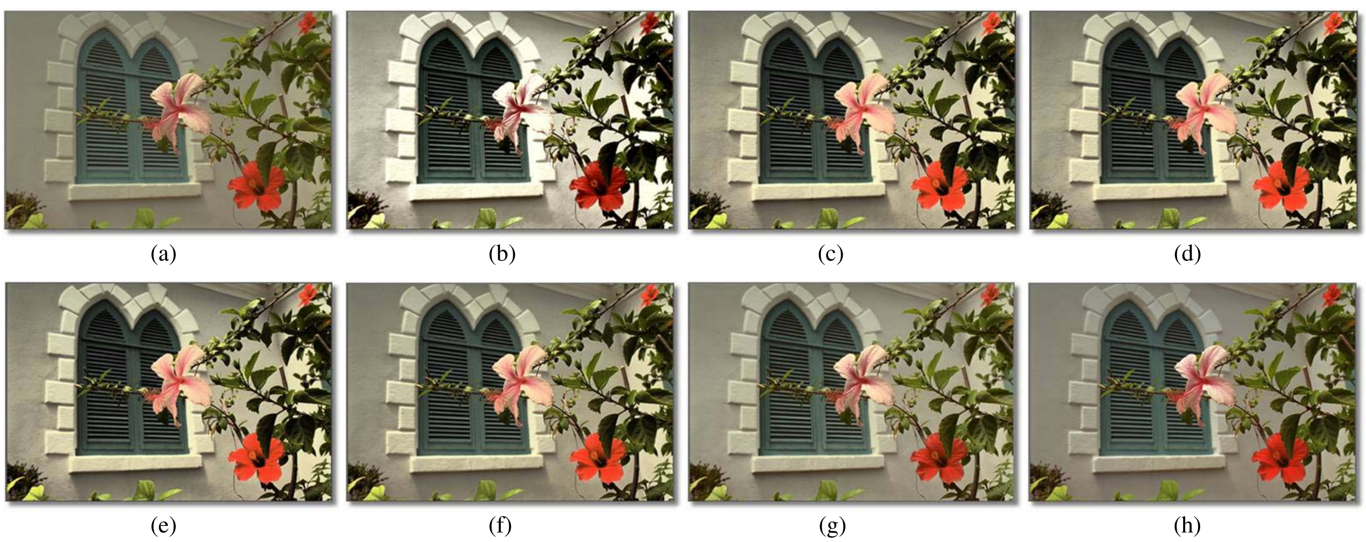


Fig. 7. Natural image *Shuttered windows* in Kodak database and the outputs. (a) Original image. (b) Output of HE. (c) Output of DSIHE [3]. (d) Output of RSIHE [5]. (e) Output of WTHE [6]. (f) Output of HMF [13]. (g) Output of OCTM [15]. (h) Output of our RICE.

We choose three couples of representative $\{\phi, \psi\}$ $(= \{1e-4, 0.02\}, \{1e-4, 0.2\}, \{1e-3, 0.2\})$, and illustrate the three enhanced images in Fig. 4(d1), (e1), and (f1). As expected, the enhanced output, as an optimal tradeoff between minimized perceptual deterioration and maximized visual informativeness and perceptual pleasure, achieves considerable improvement in terms of visual quality.

B. Automatic Realization of Ideal Histogram

The major shortage of most existing contrast enhancement technologies is overenhancement or underenhancement, which introduces unpleasant noise and deteriorates users' experience. In most cases, manual parameter tuning is used for proper enhancement. This method, however, largely reduces the applicability since it is a quite labor intensive and time-consuming job that is impractical for real-time systems. It is usually required that a good contrast enhancement approach should help to reveal indiscernible image details without introducing noticeable distortions. However, most existing QA metrics are not applicable to the contrast enhanced images [20]–[25], because the newly revealed image details and the newly introduced visual distortions are often difficult to tell apart.

In this paper, we find that the proper contrast enhancement generally reveals indiscernible image details while keeping the image saliency unchanged [Fig. 4(c2)–(f2)], whereas poor enhancement creates artifacts and thus alters visual saliency profile, as shown in Fig. 4(b2).[1] This phenomenon might be easily justified because: 1) the artifacts caused by contrast enhancement are extremely incoherent with its surroundings,

[1]Of course, this finding is not applicable to those images without prominent salient points.

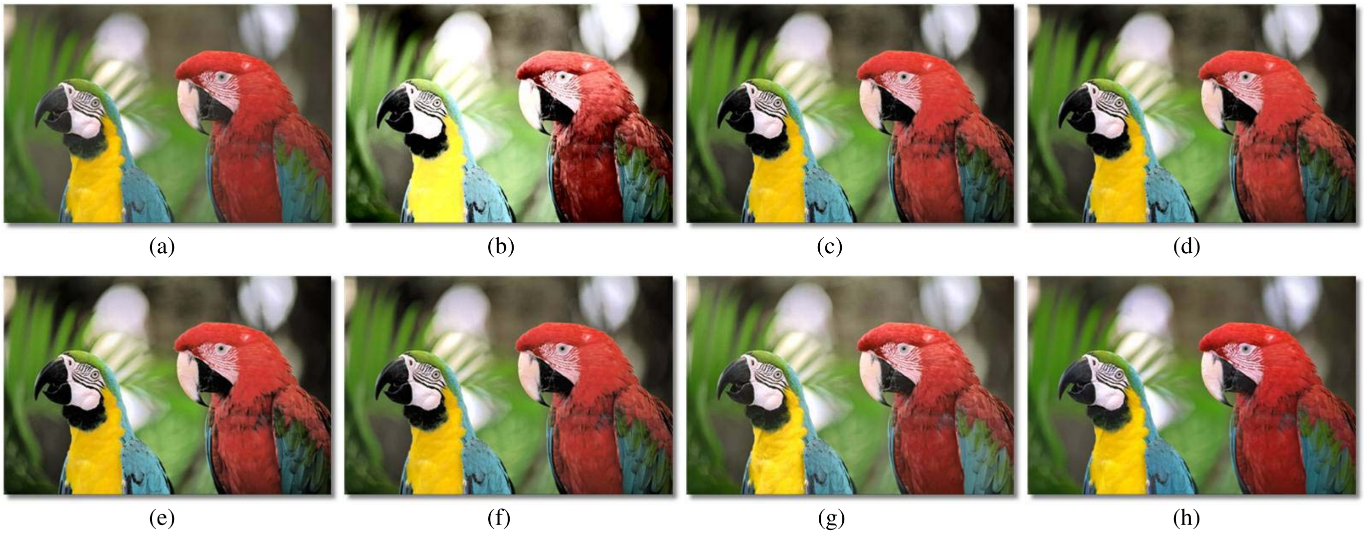


Fig. 8. Natural image *Two macaws* in Kodak database and the outputs. (a) Original image. (b) Output of HE. (c) Output of DSIHE [3]. (d) Output of RSIHE [5]. (e) Output of WTHE [6]. (f) Output of HMF [13]. (g) Output of OCTM [15]. (h) Output of our RICE.

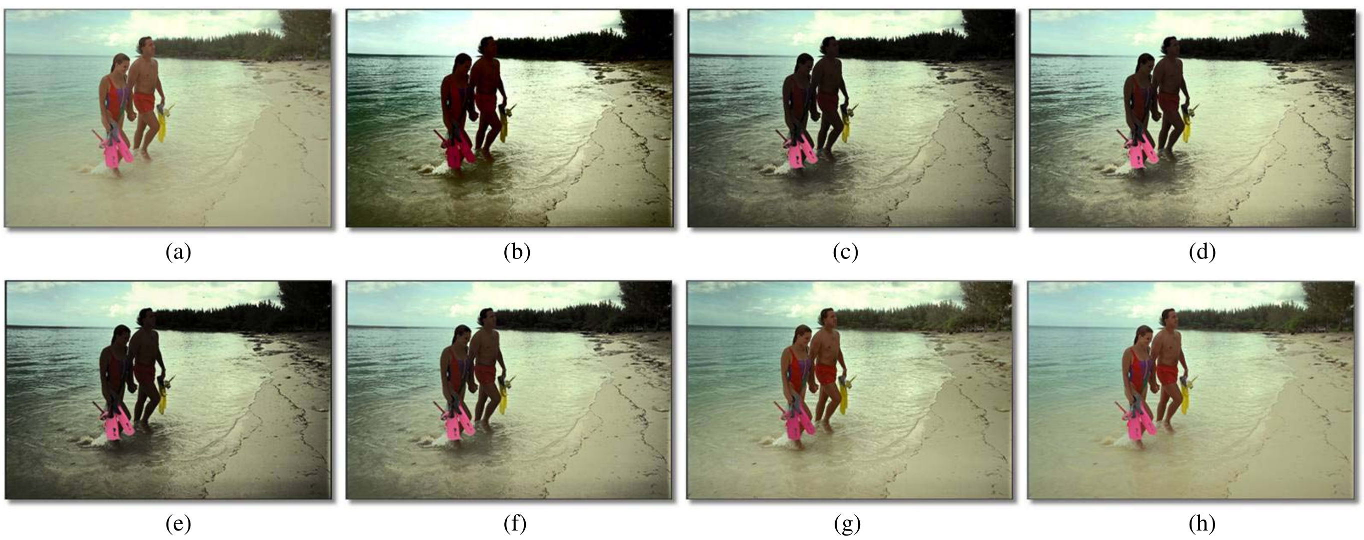


Fig. 9. Natural image *Couple on beach* in Kodak database and the outputs. (a) Original image. (b) Output of HE. (c) Output of DSIHE [3]. (d) Output of RSIHE [5]. (e) Output of WTHE [6]. (f) Output of HMF [13]. (g) Output of OCTM [15]. (h) Output of our RICE.

thereby changing saliency distribution and 2) the subtle details highlighted by contrast enhancement also exist in the original image, thus do not change the saliency distribution. Accordingly, saliency preservation can be used as an optimal rule for fine tuning the performance of contrast enhancement algorithms.

The tool of visual saliency has been successfully applied in various kinds of research topics, e.g., QA metrics [26]–[28] and video compression [29], [30]. Hayes *et al.* [31] and Oppenheim and Lim [32] pointed out that more high-frequency information are stored in the residual. Hou and Zhang [33] follow the idea and found that residual Fourier amplitude spectrum, namely the difference between the original Fourier amplitude spectrum and its smoothed version, can be used to form a saliency map. In comparison, the recently proposed image signature model discards the whole amplitude information and just retains the sign of each discrete cosine transform component. In other words, this model just requires a single bit per component, making it very compact. Specifically, the image signature is defined as

$$\text{Image Signature}(I_i) = \text{sign}(\text{DCT2}(I_i)) \tag{10}$$

where sign $(\cdot)$ is used to obtain the sign, and then the reconstructed image is derived by

$$\bar{I} = \text{IDCT2}(\text{Image Signature}(I_i)) \tag{11}$$

where DCT2 and IDCT2, respectively, stand for discrete cosine and inverse discrete cosine transforms for 2-D signals. In the end, we can get the saliency map by smoothing the squared reconstructed image

$$\text{Saliency Map} = g * (\bar{I} \circ \bar{I}) \tag{12}$$

where g is a Gaussian kernel, and $\circ$ and $*$ are the entrywise and convolution product operators, respectively. An example in Fig. 4(a2) shows the high accuracy of the image signature model in saliency detection.

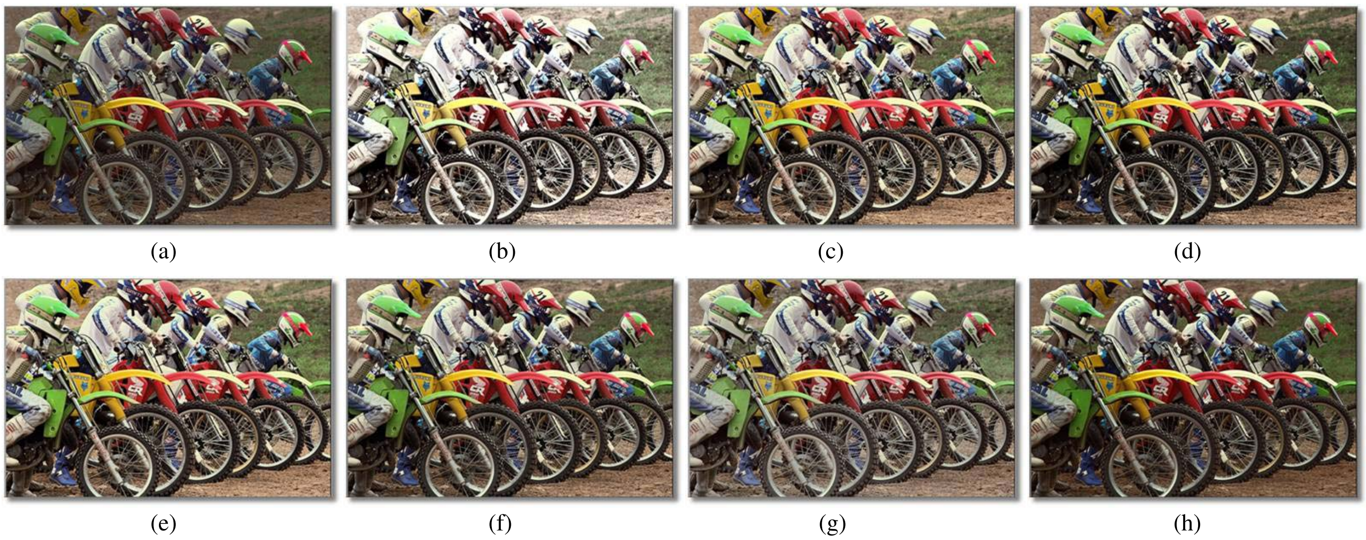


Fig. 10. Natural image *Motocross bikes* in Kodak database and the outputs. (a) Original image. (b) Output of HE. (c) Output of DSIHE [3]. (d) Output of RSIHE [5]. (e) Output of WTHE [6]. (f) Output of HMF [13]. (g) Output of OCTM [15]. (h) Output of our RICE.

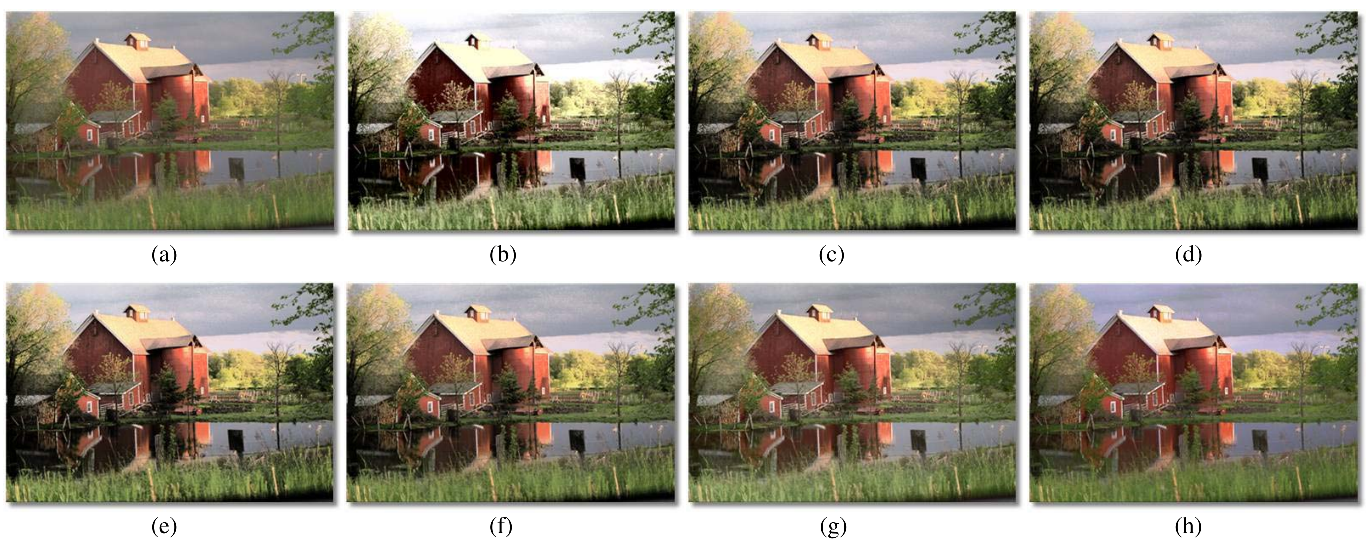


Fig. 11. Natural image *Mountain chalet* in Kodak database and the outputs. (a) Original image. (b) Output of HE. (c) Output of DSIHE [3]. (d) Output of RSIHE [5]. (e) Output of WTHE [6]. (f) Output of HMF [13]. (g) Output of OCTM [15]. (h) Output of our RICE.

Based on the aforementioned image signature model, we define a distance metric of the input image I_i and its contrast-changed version I_c using the ℓ^0 distance (i.e., the Hamming distance) of their image signatures as the first term of QMC

$$\triangle D = \|\text{sign}(\text{DCT2}(\dot{I}_i)), \text{sign}(\text{DCT2}(\dot{I}_c))\|_0 \tag{13}$$

where $\dot{I}_i$ and $\dot{I}_c$ are downsampled images of I_i and I_c by a factor of 4 using the bilateral method. This term means that the smaller the difference of saliency maps between I_i and I_c is, the higher the quality score of I_c will be.

The second term of QMC comes from [34]. Information entropy is an important concept in statistics [35]. It measures the information amount for a random image signal by quantifying its average unpredictability. In most cases, a high-contrast image is of large entropy, which enlightens us to define the second term as $\triangle E = E(I_i) - E(I_c)$. Of course, other advanced metrics, such as the Kullback–Leibler divergence and its modified symmetric Jensen–Shannon divergence [36], can be also considered, e.g., QA models [37], [38], while it was found that they do not lead to performance improvement and yet cause much higher complexity.

In this paper, we combine saliency preservation and entropy increment together with a simple linear function to derive the QMC as

$$\text{QMC}(I_i, I_c) = \triangle D + \gamma \triangle E \tag{14}$$

where γ is a fixed parameter to adjust the relative importance of two components. We find the optimal value of γ to be 0.2, which represents that saliency preservation has a more important role in the QA of contrast enhancement. It is noted that our QMC is a reduced-reference (RR) QA metric, because it only needs one single number $E(I_i)$ and a small binary map $\text{sign}(\text{DCT2}(\dot{I}_i))$ of 1/16th original image resolution.

TABLE III
OVERALL SUBJECTIVE QUALITY SCORES OF THE ORIGINAL IMAGE AND EACH OF ENHANCED VERSIONS WITH CONTRAST ENHANCEMENT ALGORITHMS TESTED ON THE KODAK IMAGE DATABASE. WE BOLD THE TOP ENHANCEMENT METHOD IN EACH IMAGE SET

j-th	ORG	HE	DSIHE	RSIHE	WTHE	HMF	OCTM	RICE
01	72	48	34	50	75	90	**96**	95
02	82	48	17	22	78	86	112	**115**
03	61	48	39	35	70	93	105	**109**
04	56	42	47	58	97	88	71	**101**
05	63	50	28	37	75	72	111	**124**
06	59	43	20	48	82	85	100	**123**
07	79	69	50	35	**93**	89	62	83
08	72	45	28	20	71	86	113	**125**
09	66	57	16	45	48	92	106	**130**
10	76	50	34	35	86	86	84	**109**
11	87	54	60	57	91	**94**	66	51
12	85	54	53	59	**107**	94	39	69
13	56	49	48	55	79	93	84	**96**
14	98	56	22	40	71	94	59	**120**
15	86	64	34	38	41	69	96	**132**
16	75	47	24	33	82	86	91	**122**
17	52	38	48	63	**100**	94	70	95
18	58	34	34	46	81	95	**106**	**106**
19	75	53	19	35	53	84	114	**127**
20	60	48	76	**80**	72	**80**	78	66
21	76	48	19	31	77	89	96	**124**
22	76	54	32	32	55	72	112	**127**
23	84	66	32	36	26	80	**122**	114
24	71	49	21	23	70	91	109	**126**
Mean	72	51	35	42	74	87	92	**108**

TABLE IV
SIMILARITY EVALUATIONS BETWEEN SALIENCY MAPS IN EACH OF TESTING IMAGE SUBSETS. WE HIGHLIGHT THE BEST PERFORMED ENHANCEMENT METHOD IN EACH IMAGE SET, AND LABEL THE LOWEST SCORE WITH BRACKETS IN EACH ALGORITHM

j-th	HE	DSIHE	RSIHE	WTHE	HMF	OCTM	RICE
01	0.934	0.917	0.911	0.919	0.934	**0.977**	0.967
02	0.869	(0.848)	(0.862)	(0.847)	0.908	(0.930)	(**0.932**)
03	0.867	0.893	0.902	0.891	0.928	0.954	**0.962**
04	0.939	0.958	0.961	0.958	0.968	**0.975**	0.974
05	0.916	0.932	0.932	0.930	0.946	0.958	**0.970**
06	0.932	0.937	0.936	0.936	0.941	0.980	**0.980**
07	0.951	0.961	0.961	0.961	**0.968**	0.964	0.965
08	0.952	0.958	0.956	0.959	0.961	**0.971**	0.970
09	0.935	0.891	0.892	0.891	0.914	0.969	**0.977**
10	0.900	0.928	0.941	0.926	0.945	0.964	**0.971**
11	0.948	0.971	0.975	0.970	**0.981**	0.976	**0.981**
12	0.926	0.963	0.974	0.957	0.970	0.954	**0.973**
13	(0.824)	0.912	0.923	0.911	0.934	**0.952**	0.948
14	0.946	0.919	0.931	0.917	0.944	0.967	**0.981**
15	0.932	0.959	0.958	0.959	0.962	**0.974**	0.973
16	0.877	0.906	0.907	0.905	0.933	**0.962**	0.958
17	0.914	0.932	0.934	0.933	0.947	0.964	**0.978**
18	0.896	0.919	0.926	0.915	0.944	0.949	**0.958**
19	0.919	0.912	0.919	0.912	0.940	0.958	**0.973**
20	0.917	0.936	0.948	0.933	0.958	0.951	**0.960**
21	0.872	0.888	0.891	0.886	0.920	**0.944**	0.942
22	0.892	0.931	0.935	0.933	0.943	0.970	**0.972**
23	0.889	0.904	0.904	0.903	0.920	0.944	**0.963**
24	0.888	0.875	0.879	0.880	(0.903)	0.950	**0.964**
Mean	0.910	0.923	0.927	0.922	0.942	0.961	**0.966**

For the QMC to be a practical method, it must have high accuracy and low computational cost. We test the performance of QMC and compare it with state-of-the-art QA metrics on the recent contrast-changed image database (CID2013) [34], which is composed of 15 natural images of size 768×512 from the Kodak database [19] and totally 400 contrast-changed version and associated mean opinion scores (MOSs) obtained from 22 inexperienced observers. Most of the viewers were college students with different majors. They include 15 males and 7 females. All testing images can be classified into two groups. The first one is generated by mean shifting natural images with positive or negative numbers that have six levels of $\{20, 40, 60, 80, 100, 120\}$, and the second group of contrast-changed images is created with four kinds of transfer mapping curves, i.e., concave arc, convex arc, cubic function, and logistic function. The six testing QA methods are given below.

1) Full-reference (FR) QA algorithms assuming that both original and distorted images are wholly known: a) feature similarity (FSIM) index [20], which is inspired by the fact that the human visual system understands an image mainly relying on low-level features, and uses the complementary phase congruency [39] and gradient magnitude [40] to characterize the image quality; b) gradient similarity (GSIM) index [21], measuring the changes of GSIM in contrast and structure in images; c) internal generative mechanism (IGM) [22], fusing modified PSNR and structural similarity [41] values computed on predicted and disorderly regions with psychophysical parameters [42]; and d) SSIM weighted [23], estimating weights in a block-based manner for improving SSIM.
2) RR QA metrics: a) free energy based distortion metric (FEDM) [24], quantifying the psychovisual quality as the agreement between an input image and its output of the internal generative model based on the recent free-energy theory [43], which explains some brain theories in biological and physical sciences about human action, perception, and learning; b) structural degradation model (SDM) [25], which succeeded in improving the FR SSIM [41] into RR QA according to an observation that, for most images with various distortion types and quality levels, their low-pass filtered versions have different spatial frequency decrease; and c) RR image quality metric for contrast change (RIQMC) [34], which depends on the information residual between the input and distorted images as well as the first four-order statistics of the distorted image histogram.

As suggested by the Video Quality Experts Group (VQEG) [44], we compare the proposed QMC with the testing QA metrics via Spearman rank-order correlation coefficient (SROCC), which is one of the most popular performance measures and has been widely used to find the suitable parameters in several existing QA metrics such as [45]–[47]. The SROCC is defined by

$$\mathrm{SROCC} = 1 - \frac{6 \sum_{i=1}^{R} r_i^2}{R(R^2 - 1)} \tag{15}$$

where r_i represents the distinction between the ith image's ranks in subjective and quality objective scores, and R stands for the image number in the testing database. SROCC is a nonparametric rank-based correlation measure, independent of any monotonic linear/nonlinear mapping between subjective and objective evaluations. A value close to 1 for SROCC indicates superior performance of the QA model.

We provide the performance indices of the competing QA metrics in Table I. It is clear that the proposed QMC model has achieved substantially high performance, much better than state-of-the-art FR and RR QA algorithms. Furthermore, our QMC needs very little average run time as compared with the testing methods, as listed in Table I. We then calculate SROCC on each original image (the CID2013 database includes 15 natural images) and associated contrast-changed versions, and report those performance measures in Table II. Our approach also obtains very high and stable results: all of SROCC values are higher than 0.927. In Fig. 5, we display the scatter plot of QMC on the overall CID2013 database to show the good monotonicity.

As QMC turns out to be an effective quality metric for contrast-changed images, we in this paper utilize QMC to optimize the parameters $\{\phi, \psi\}$ for the contrast enhancement algorithm as

$$\begin{aligned}\{\phi_{\text{opt}}, \psi_{\text{opt}}\} &= \arg\min_{\{\phi,\psi\}} \text{QMC}(I_i, \tilde{I}) \\ &= \arg\min_{\{\phi,\psi\}} \text{QMC}\left(I_i, T_{\text{hm}}\left(I_i, \frac{\mathbf{h_i} + \phi\mathbf{h_{eq}} + \psi\mathbf{h_{sig}}}{1+\phi+\psi}\right)\right).\end{aligned} \tag{16}$$

In this way, the algorithm can automatically obtain the properly enhanced image I_{opt} with $\{\phi_{\text{opt}}, \psi_{\text{opt}}\}$ and histogram matching in (9). Note that, as indicated in Fig. 5, the smaller the QMC value, the better the visual quality.

III. Experimental Results and Discussion

Based on the analysis in previous sections, we can quantify the performance of an image contrast enhancement algorithm in the following three aspects: 1) subjective quality; 2) saliency preservation; and 3) computational complexity. We first select 24 natural images from the Kodak image database [19]. The testing images have a wide range of contents, such as humans and animals, and indoor and outdoor scenes. We then choose six contrast enhancement techniques for comparison, which include the classical HE and its modified DSIHE [3], RSIHE [5], WTHE [6], as well as state-of-the-art HMF [13] and OCTM [15].

A. Subjective Quality

As shown in Fig. 4(d1)–(f1), we have found a way to generate a more visually informative and perceptually pleasing as well as less visually disturbing image via finding a good compromise among the original image and its histogram equalized and STBP transferred copies. The enhanced output by STBP indeed stretches the original image histogram to both sides, and thus increases the image contrast. Next, we propose the QMC (a high-performance quality metric for contrast change) and use it to automatically acquire appropriate parameters and create associated optimal enhanced products. As shown in Figs. 6–11(h), the output images are of suitable luminance, hue, and tone, and do not introduce artifacts and noise. Furthermore, in these enhanced outputs, the foggy appearance has been removed and the images are more vivid and clear.

HE and its variants DSIHE and RSIHE work ineffectively because they usually generate too-bright or too-dark regions [Figs. 7, 9, and 10(b)–(d)]. In addition, they also sometimes introduce disturbing artifacts/noise [Fig. 6(b)–(d)]. Due to the weighting and thresholding on HE, WTHE reduces the unfavorable effect of HE. However, WTHE still encounters the problem of overbrightness, e.g., in Fig. 10(e) or overdark, e.g., in Fig. 9(e), and may even cause noticeable artifacts, e.g., in Fig. 6(e). The recently developed HMF searches for the good tradeoff of the input image and its HE product, and this lessens the disadvantage of HE to some extent. As shown in Fig. 6(f), HMF cannot always guarantee a balance between visual quality and artifacts prevention. Another state-of-the-art algorithm, OCTM, although solves the problem of overenhancement or less enhancement, as shown in Figs. 6–11(g), the enhanced images look pale and somewhat unnatural.

In addition, we also conduct a subjective experiment for quantitative perceptual quality measurements. In this experiment, we invited a total of 20 viewers to score the overall enhanced images. The subjects participating in this test includes 15 males and 5 females. To make the subjective ratings more faithful, the popular paired comparison method is used to rank each pair of the wholly 192 images, which consists of 24 original images and associated 168 images generated by seven contrast enhancement technologies. An elaborately designed interactive system reduces the process of scoring to alternatively pressing two adjacent keys (left is better or right is better). We tabulate the overall score for each enhanced image in Table III. Note that higher score indicates better performance. According to the mean subjective rating score, we find that our RICE model is much superior to other algorithms tested. For each image set, the proposed RICE technique has also obtained outstanding results by winning the first place on 17 image sets, far beyond the WTHE and OCTM, which only win three times, respectively.

B. Saliency Preservation

We have argued that saliency preservation can effectively be used to avoid overenhancement and underenhancement. To further validate this assumption, a couple of objective and subjective experiments are implemented to compare contrast enhancement results; that is to evaluate how well the enhanced output preserves the original visual saliency.

In the objective test, we apply a new fast similarity metric [48]. We denote by SM_i and SM_e the normalized saliency maps of the input image and the enhanced output that is computed by the state-of-the-art saliency detection model in [18]. The similarity is the sum of the minimum values at each point in the two normalized maps, and is mathematically

TABLE V

SIMILARITY EVALUATION OF SALIENCY MAPS OF THE ORIGINAL AND ENHANCED IMAGES IN EACH TESTING DATA SET. WE EMPHASIZE THE TOP ENHANCEMENT METHOD IN EACH IMAGE SET WITH BOLDFACE, AND LABEL THE LOWEST SCORE WITH BRACKETS IN EACH ALGORITHM

j-th	HE	DSIHE	RSIHE	WTHE	HMF	OCTM	RICE
01	0.732	0.762	0.822	0.751	0.752	**0.826**	0.795
02	(0.592)	(0.691)	0.662	0.662	(0.427)	0.687	**(0.713)**
03	0.770	0.746	0.725	0.724	0.805	0.804	**0.820**
04	**0.829**	0.765	0.802	0.750	0.806	0.746	0.773
05	0.704	0.805	0.791	0.805	0.779	0.812	**0.827**
06	0.793	0.756	**0.805**	0.790	0.784	0.797	0.774
07	0.699	0.764	0.741	0.770	0.698	0.732	**0.841**
08	0.833	0.807	0.833	0.828	0.843	0.819	**0.867**
09	0.812	0.730	0.819	0.778	0.852	0.826	**0.855**
10	0.813	0.781	0.781	0.790	0.800	0.754	**0.824**
11	0.740	0.765	0.811	0.767	0.800	0.730	**0.852**
12	0.759	0.790	0.782	0.757	0.707	0.778	**0.829**
13	0.769	0.760	0.789	**0.822**	0.735	0.804	0.821
14	0.755	0.779	0.859	0.734	0.713	0.853	**0.866**
15	0.815	0.813	0.739	(0.649)	0.765	0.829	**0.831**
16	0.775	0.739	0.755	0.769	0.784	0.798	**0.805**
17	0.802	0.817	0.826	0.831	0.759	0.823	**0.839**
18	0.805	0.770	0.796	0.776	0.787	0.796	**0.816**
19	0.779	0.766	0.790	0.740	0.739	0.805	**0.829**
20	0.804	0.786	0.822	0.794	0.816	0.800	**0.827**
21	0.823	0.785	0.790	**0.846**	0.759	(0.653)	0.833
22	0.809	0.756	0.813	0.775	0.809	0.803	**0.826**
23	0.695	0.808	(0.625)	0.758	0.832	0.830	**0.850**
24	0.731	0.763	0.698	0.755	0.734	0.815	**0.821**
Mean	0.768	0.771	0.778	0.767	0.762	0.788	**0.822**

defined as

$$\text{Similarity} = \sum_{l=1}^{L} \min(\text{SM}_i(l), \text{SM}_e(l)) \tag{17}$$

where

$$\sum_{l=1}^{L} \text{SM}_i(l) = \sum_{l=1}^{L} \text{SM}_e(l) = 1 \tag{18}$$

with L being the image pixel number. Note that a similarity score of one indicates that the two saliency maps are the same, whereas that of zero indicates that they do not overlap at all, namely totally opposite. In other words, a value close to 1 indicates high performance.

In Table IV, we report the similarity measures of the testing contrast enhancement approaches on the overall 24 images in the Kodak database. As expected, the proposed RICE achieves the best result in 16 images, up to 67% of all the test scenarios. Not surprisingly, our RICE also acquires the highest similarity score on average, outperforming other testing methods. In practice, for those contrast enhancement algorithms under comparison, we can roughly come to an agreement of subjective assessment results on average: HE $<$ DSIHE $\leq$ RSHIE $\leq$ WTHE $<$ HMF $<$ OCTM $<$ RICE. This is almost the same with the average objective measure in Table IV, except for DSIHE, RSIHE, and WTHE, which have very close enhancement effects. Furthermore, we want to emphasize that the proposed RICE algorithm is also robust across various image scenes since its similarity evaluation for each image is larger than 0.932, whereas the others have comparatively much lower scores, for instance, HE: 0.824, DSIHE: 0.848, RSIHE: 0.862, WTHE: 0.847, HMF: 0.903, and OCTM: 0.930, as labeled with brackets in Table IV.

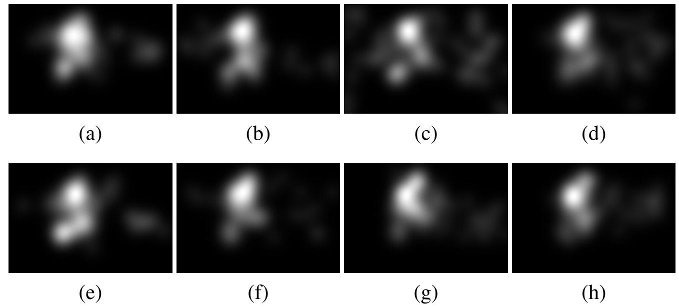


Fig. 12. Exemplary saliency maps obtained from (a) original image and enhanced counterparts that are created by (b) HE, (c) DSIHE, (d) RSIHE, (e) WTHE, (f) HMF, (g) OCTM, and the proposed (h) RICE models shown in Fig. 9 using the subjective eye-tracking test.

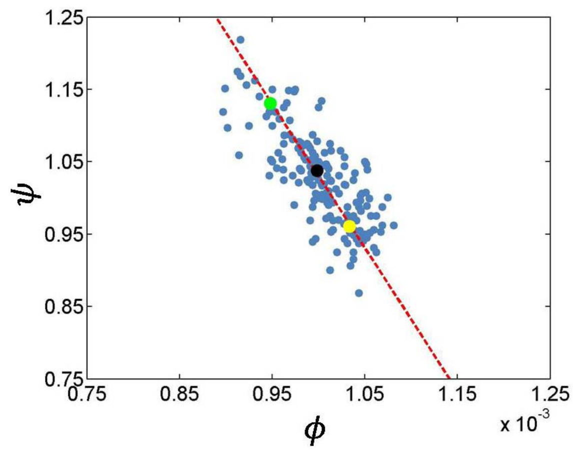


Fig. 13. Scatter plot of the optimal ϕ and ψ computed on 200 images from the Berkeley database [49]. The red dash line is fitted using the least square method on (20). The colored points are three clusters using k-means clustering [50].

An eye-tracking experiment for saliency preservation is also conducted using the Tobbi T120 Eye Tracker. Tobbi T120 is integrated into a 17-in thin-film transistor monitor to make the user experience as natural as possible and it has sample frequency of up to 120 Hz. The monitor has a resolution of 1280×1024 pixels. It has a spatial resolution of 0.3° with a typical accuracy of 0.5°. The head movement box of Tobbi T120 (width $\times$ hight) is 30×22 cm at 70 cm, and the suitable viewing distance is 50–80 cm. During the test, each subject was asked to look freely at the entire 192 images used in the aforementioned subjective QA shown on the monitor. After the acquirement of the fixation data, we generated the saliency maps with the fixation location according to [26]. A 2-D Gaussian mask is used to generate the final saliency map

$$\text{SM}(k, l) = \sum_{i=1}^{T} \exp\left[-\frac{(x_i - k)^2 + (y_i - l)^2}{\sigma^2}\right] \tag{19}$$

where $\text{SM}(k, l)$ indicates the saliency map of the input visual stimulus. $k \in [1, M]$ and $l \in [1, N]$ with M and N being the image height and width. (x_i, y_i) is the spatial coordinate of the

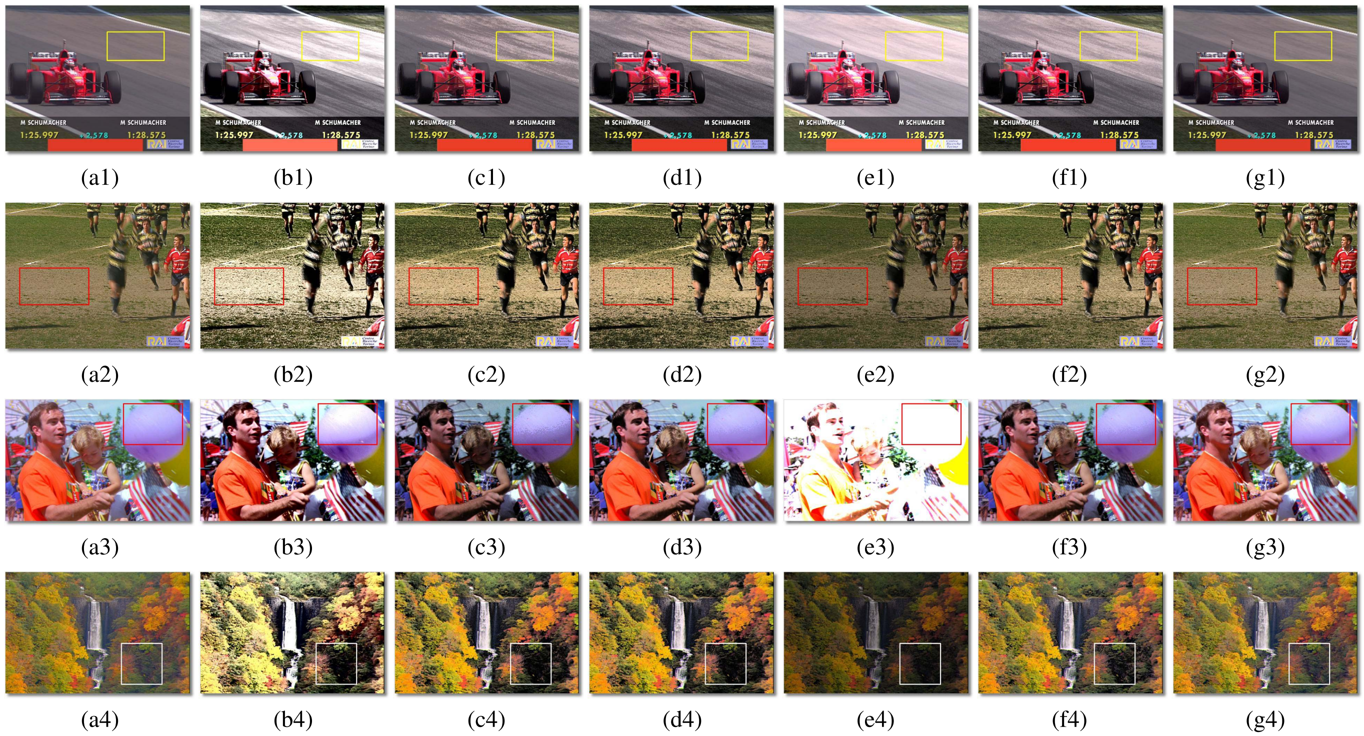


Fig. 14. Enhanced video frames in four representative video sequences in the VQEG database [44] and the corresponding outputs. (a1)–(a4) Original frame. (b1)–(b4) HE. (c1)–(c4) DSIHE [3]. (d1)–(d4) RSIHE [5]. (e1)–(e4) WTHE [6]. (f1)–(f4) HMF [13]. (g1)–(g4) Proposed RICE. We label some remarkable regions with colored rectangles for comparison.

ith fixation ($i = 1, \ldots, T$) with T being the total number of all fixations over all subjects. σ indicates the standard deviation of the Gaussian kernel. We linearly normalize the intensity of the resulting saliency maps to the range [0, 1]. Fig. 12 shows some exemplary saliency maps obtained from this subjective test. In Fig. 12(a)–(h), the saliency maps come from the original image and enhanced versions created by HE, DSIHE, RSIHE, WTHE, HMF, OCTM, and the proposed RICE algorithms. It can be readily viewed that the saliency map of our RICE is more similar to that of the original image, as compared with other contrast enhancement approaches.

Using the fast similarity metric introduced above, we further quantify the similarity degree, as shown in Table V. The proposed model has achieved the best average performance among all the contrast enhancement technologies tested. We find that our RICE wins as much as 19× the first place, up to around 80%. It also needs to stress that the similarity result of RICE on each image subset is higher than 0.7, and even larger than 0.8 on 20 sets, over 83%. The above results and comparisons indicate the good saliency preservation ability of the proposed technique.

C. Computational Complexity

This section will discuss and compare the computational complexities of the proposed RICE and those testing methods for an image of size $W \times H$ and B bins, following the method in [13]. For HE, the computation of the histogram requires $\mathcal{O}(WH)$ time, calculating the mapping function from the histogram requires $\mathcal{O}(2^B)$ time, and finally obtaining the enhanced image with the mapping function requires $\mathcal{O}(WH)$ time. Hence, its total time complexity is $\mathcal{O}(2WH + 2^B)$.

For DSIHE, RSIHE, WTHE, and HMF, the computation of the histogram requires $\mathcal{O}(WH)$ time, that of the modified histogram for each bin requires $\mathcal{O}(2^B)$ time, and that of the mapping function requires $\mathcal{O}(2^B)$ time. In summary, those methods totally requires $\mathcal{O}(2WH + 2^{B+1})$ time to create an enhanced image. For OCTM, it needs a great amount of run time since linear programming is used to solve a complicated optimization function.

For a fair comparison, this paper estimates the computational complexity of the proposed RICE without the automatic optimization step. Note that the solutions of $\lceil \text{mean}(I_i)/32 \rceil * 32$ are limited, so we can first solve the optimization problem in (3), compute the associated sigmoid transfer mappings offline, and store them in a lookup table to speed up the RICE algorithm. Consequently, the total time complexity of RICE is also $\mathcal{O}(2WH + 2^{B+1})$.

To further reduce the computational load, we first acquire optimal $\{\phi, \psi\}$ values on 200 random images from the Berkeley database [49] in Fig. 13(a). Based on an observation that there exists an approximate linear relationship of the optimal ϕ and ψ values, we then fit the linear regression model

$$\psi = s \cdot \phi + t \tag{20}$$

where s and t are acquired using the least square method, and their estimated values are 1982 and 3.012. In this way, we can remarkably decrease the computational complexity of the proposed technique. Moreover, we apply k-means clustering [50] to find three clusters, as labeled in Fig. 13, and this can further save the computational time of our algorithm by enumerating the three possibilities and choosing the best one. According to the above analysis, the proposed RICE model is shown to

TABLE VI

SIMILARITY EVALUATIONS OF SALIENCY MAPS IN EACH SUBSETS IN THE VQEG DATABASE. WE BOLD THE TOP ENHANCEMENT METHOD IN EACH SET, AND LABEL THE LOWEST SCORE WITH BRACKETS IN EACH MODEL

j-th	HE	DSIHE	RSIHE	WTHE	HMF	RICE
01	0.9488	0.9529	0.9528	**0.9640**	0.9634	(0.9622)
02	0.9375	0.9490	0.9593	0.9633	0.9322	**0.9831**
03	0.9122	0.9424	0.9513	0.9609	0.9558	**0.9785**
04	0.9035	0.9137	0.9279	0.9471	(0.8256)	**0.9795**
05	0.8895	0.9287	0.9405	0.9525	0.9127	**0.9770**
06	0.9269	0.9435	0.9504	0.9607	0.9456	**0.9793**
07	0.9365	0.9543	0.9554	0.9696	0.9455	**0.9771**
08	0.9242	0.9305	0.9423	0.9489	0.9681	**0.9716**
09	0.9089	0.9482	0.9490	0.9646	0.9713	**0.9796**
10	0.9772	0.9561	0.9773	0.9746	0.9834	**0.9837**
11	0.9461	0.9441	0.9597	0.9656	0.8293	**0.9782**
12	0.9055	0.8989	0.9583	0.9423	0.9431	**0.9744**
13	0.9758	0.9458	0.9696	0.9665	0.9327	**0.9826**
14	0.9246	0.9375	0.9483	0.9576	0.9332	**0.9752**
15	0.8749	0.9318	0.9601	0.9532	0.8278	**0.9804**
16	0.9248	0.9374	0.9377	0.9616	0.9456	**0.9825**
17	0.9036	0.9205	0.9210	0.9452	0.9586	**0.9807**
18	0.9699	0.9778	0.9753	0.9830	0.9168	**0.9768**
19	(0.8311)	(0.8737)	(0.8756)	(0.9125)	0.9417	**0.9742**
20	0.9388	0.9486	0.9660	0.9672	0.9034	**0.9734**
Mean	0.9230	0.9368	0.9489	0.9580	0.9268	**0.9775**

TABLE VII

COMPARISON OF THE MEDIAN BRIGHTNESS ON 20 VIDEO SEQUENCES IN THE VQEG DATABASE. WE EMPHASIZE THE METHOD THAT HAS THE CLOSEST MEDIAN BRIGHTNESS TO THE ORIGINAL VERSION

j-th	Original	HE	DSIHE	RSIHE	WTHE	HMF	RICE
01	86	119	89	83	87	107	**86**
02	113	131	59	79	84	124	**112**
03	78	139	67	76	75	110	**78**
04	72	120	**72**	58	74	252	76
05	97	124	101	92	98	115	**97**
06	121	128	103	110	115	136	**121**
07	94	129	109	93	100	88	**94**
08	175	134	106	153	153	154	**173**
09	102	122	101	95	**102**	83	101
10	106	128	72	93	88	84	**106**
11	126	131	89	110	108	237	**126**
12	62	125	102	65	80	74	**62**
13	124	133	72	102	95	236	**124**
14	82	119	77	74	80	102	**82**
15	41	130	90	44	62	142	**41**
16	91	113	72	72	83	35	**87**
17	107	117	84	85	99	96	**106**
18	130	128	111	133	118	238	**129**
19	109	129	109	106	109	94	**109**
20	65	124	76	**65**	70	181	67

be not only of low computational complexity but also of high flexibility in reducing the computational complexity.

IV. EXTENSION TO VIDEO ENHANCEMENT

Besides image enhancement, the proposed technique can also be extended to video enhancement. Early researchers emphasized the significance of brightness preservation [2]–[5]. However, as shown in Figs. 6–11(b)–(d), those enhancement approaches cannot always create delighting outputs due to the introduction of visually disturbing artifacts. Brightness preservation is more important for videos than for images, because the brightness deviation usually generates temporal flickering artifacts, which are commonly seen in enhanced video sequences processed by HE-based techniques. As a result, we adopt the median brightness preservation for video streams owing to its simpleness and acquirement of maximum entropy [3], and then we rewrite (8) by replacing $\mathbf{h_{eq}}$ by $\mathbf{h_{dsihe}}$ that is calculated with DSHIE

$$\tilde{\mathbf{h}} = \frac{\mathbf{h_i} + \phi \mathbf{h_{dsihe}} + \psi \mathbf{h_{sig}}}{1 + \phi + \psi}. \tag{21}$$

Since $\mathbf{h_i}$, $\mathbf{h_{dsihe}}$, and $\mathbf{h_{sig}}$ almost have the same median brightness value, their weighted combination $\tilde{\mathbf{h}}$ has the equivalent median brightness as well. In this way, we succeed in keeping brightness when enhancing video sequences.

Another common problem in video technology is the efficiency. To solve this, an entropy-inspired model is applied. More precisely, at the beginning of the process, we utilize the RICE to generate a mapping curve of the first video frame, and store this mapping curve. For subsequent video frames, the entropy model is used to compute the differences of the information content between two successive frames, which can be approximated as

$$E = -\sum_{i=0}^{255} p_i \log(p_i) \tag{22}$$

where p_i is the probability density at the ith pixel. When the absolute difference of E between the current and previous frames exceeds the threshold T, the transfer mapping curve will be updated. Otherwise, the existing mapping curve stored is immediately applied to transform each intensity level in the incoming video frame.

We exhibit enhanced video frames from four representative video scenes in Fig. 14 and label some important regions with colored rectangles for comparison. The HE and its related DSIHE, RSIHE, and WTHE methods still suffer from those above-mentioned drawbacks, e.g., generating too-bright or too-dark outputs or introducing noise and temporal flickering artifacts. Although HMF performs somewhat well, it sometimes causes artifacts, e.g., in Fig. 14(f1)–(f3), or renders the results unnatural, e.g., in Fig. 14(f4). The RICE technique produces properly enhanced images, not only highlighting indiscernible details but also preventing noticeable artifacts, as can be observed in Fig. 14(g1)–(g4).

We also measure saliency preservation on the VQEG video database, and list the results in Table VI. The RICE algorithm has achieved the best performance in 95% video sequences, and outperforms other contrast enhancement algorithms with seizable margins. The robustness of RICE is good as well, with each of the similarity scores greater than 0.9622. In contrast, as labeled with brackets in Table VI, other methods have poor similarity scores, e.g., HE: 0.8311, DSIHE: 0.8737, RSIHE: 0.8756, WTHE: 0.9125, and HMF: 0.8256.

Video contrast enhancement requires keeping brightness, because a small amount of luminance fluctuations will produce intensively flickering artifacts, and thus seriously degrades the perceptual quality. In addition, a video sequence perhaps involves different scenes, e.g., daylight seaside and dark-night seabed. In these conditions, most existing contrast enhancement technologies tend to generate gloomy seaside and bright seabed, violating the common sense. The proposed RICE can guarantee the original brightness well preserved and thus avoid temporal artifacts. We report in Table VII the median brightness of the competing approaches, and highlight the method that has the closest median brightness to the original one. Clearly, the proposed RICE achieves outstanding results, and outperforms all other methods for 85% video sequences.

V. Conclusion

In this paper, we have proposed a new RICE technology. We comprehensively consider the properties of visual informativeness, perceptual deterioration, and visual pleasantness. We then design a general framework to combine the constraints from the original image and its histogram equalized and sigmoid mapping transferred versions to get the properly enhanced images. To address the problems of overenhancement and underenhancement, which are faced by most existing contrast enhancement methods, we design an efficient and effective QMC for image contrast based on the concept of saliency preservation. QMC also helps to optimize the model parameters used in the RICE algorithm so as to guarantee the optimal outputs. We have tested the performance of the RICE model with many existing enhancement algorithms, such as HE and its variants DSIHE, RSIHE, WTHE, and state-of-the-art HMF and OCTM, in terms of subjective quality, saliency preservation, and computational complexity. The experimental results prove the superiority of the proposed model. MATLAB codes will be released online at http://multimedia.sjtu.edu.cn/.

References

[1] R. C. Gonzalez and R. E. Woods, *Digital Image Processing*. Reading, MA, USA: Addison-Wesley, 1992.

[2] Y.-T. Kim, "Contrast enhancement using brightness preserving bi-histogram equalization," *IEEE Trans. Consum. Electron.*, vol. 43, no. 1, pp. 1–8, Feb. 1997.

[3] Y. Wang, Q. Chen, and B. Zhang, "Image enhancement based on equal area dualistic sub-image histogram equalization method," *IEEE Trans. Consum. Electron.*, vol. 45, no. 1, pp. 68–75, Feb. 1999.

[4] S.-D. Chen and A. R. Ramli, "Contrast enhancement using recursive mean-separate histogram equalization for scalable brightness preservation," *IEEE Trans. Consum. Electron.*, vol. 49, no. 4, pp. 1301–1309, Nov. 2003.

[5] K. S. Sim, C. P. Tso, and Y. Y. Tan, "Recursive sub-image histogram equalization applied to gray scale images," *Pattern Recognit. Lett.*, vol. 28, no. 10, pp. 1209–1221, Jul. 2007.

[6] Q. Wang and R. K. Ward, "Fast image/video contrast enhancement based on weighted thresholded histogram equalization," *IEEE Trans. Consum. Electron.*, vol. 53, no. 2, pp. 757–764, May 2007.

[7] M. Abdullah-Al-Wadud, M. H. Kabir, M. A. A. Dewan, and O. Chae, "A dynamic histogram equalization for image contrast enhancement," *IEEE Trans. Consum. Electron.*, vol. 53, no. 2, pp. 593–600, May 2007.

[8] G.-H. Park, H.-H. Cho, and M.-R. Choi, "A contrast enhancement method using dynamic range separate histogram equalization," *IEEE Trans. Consum. Electron.*, vol. 54, no. 4, pp. 1981–1987, Nov. 2008.

[9] H. Ibrahim and N. S. P. Kong, "Brightness preserving dynamic histogram equalization for image contrast enhancement," *IEEE Trans. Consum. Electron.*, vol. 53, no. 4, pp. 1752–1758, Nov. 2007.

[10] N. S. P. Kong and H. Ibrahim, "Color image enhancement using brightness preserving dynamic histogram equalization," *IEEE Trans. Consum. Electron.*, vol. 54, no. 4, pp. 1962–1968, Nov. 2008.

[11] S.-H. Yun, J. H. Kim, and S. Kim, "Image enhancement using a fusion framework of histogram equalization and Laplacian pyramid," *IEEE Trans. Consum. Electron.*, vol. 56, no. 4, pp. 2763–2771, Nov. 2010.

[12] R. Lai, Y.-T. Yang, B.-J. Wang, and H.-X. Zhou, "A quantitative measure based infrared image enhancement algorithm using plateau histogram," *Opt. Commun.*, vol. 283, no. 21, pp. 4283–4288, Nov. 2010.

[13] 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.

[14] A. Majumder and S. Irani, "Perception-based contrast enhancement of images," *ACM Trans. Appl. Perception*, vol. 4, no. 3, Nov. 2007, Art. ID 17.

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

[16] 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.

[17] K. Gu, G. Zhai, M. Liu, X. Min, X. Yang, and W. Zhang, "Brightness preserving video contrast enhancement using S-shaped transfer function," in *Proc. IEEE Int. Conf. Vis. Commun. Image Process.*, Nov. 2013, pp. 1–6.

[18] X. Hou, J. Harel, and C. Koch, "Image signature: Highlighting sparse salient regions," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 34, no. 1, pp. 194–201, Jan. 2012.

[19] *Kodak Lossless True Color Image Suite*. [Online]. Available: http://r0k.us/graphics/kodak/, accessed Nov. 2014.

[20] L. Zhang, D. 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, Aug. 2011.

[21] 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, Apr. 2012.

[22] 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, Jan. 2013.

[23] K. Gu, G. Zhai, X. Yang, W. Zhang, and M. Liu, "Structural similarity weighting for image quality assessment," in *Proc. IEEE Int. Conf. Multimedia Expo Workshops*, Jul. 2013, pp. 1–6.

[24] G. Zhai, X. Wu, X. Yang, W. Lin, and W. Zhang, "A psychovisual quality metric in free-energy principle," *IEEE Trans. Image Process.*, vol. 21, no. 1, pp. 41–52, Jan. 2012.

[25] K. Gu, G. Zhai, X. Yang, and W. Zhang, "A new reduced-reference image quality assessment using structural degradation model," in *Proc. IEEE Int. Symp. Circuits Syst.*, May 2013, pp. 1095–1098.

[26] H. Liu and I. Heynderickx, "Visual attention in objective image quality assessment: Based on eye-tracking data," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 21, no. 7, pp. 971–982, Apr. 2011.

[27] K. Gu, G. Zhai, X. Yang, L. Chen, and W. Zhang, "Nonlinear additive model based saliency map weighting strategy for image quality assessment," in *Proc. IEEE 14th Int. Workshop Multimedia Signal Process.*, Sep. 2012, pp. 313–318.

[28] X. Min, G. Zhai, Z. Gao, and K. Gu, "Visual attention data for image quality assessment databases," in *Proc. IEEE Int. Symp. Circuits Syst.*, Jun. 2014, pp. 894–897.

[29] R. Li, B. Zeng, and M. L. Liou, "A new three-step search algorithm for block motion estimation," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 4, no. 4, pp. 438–442, Aug. 1994.

[30] L.-M. Po and W.-C. Ma, "A novel four-step search algorithm for fast block motion estimation," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 6, no. 3, pp. 313–317, Jun. 1996.

[31] M. H. Hayes, J. S. Lim, and A. V. Oppenheim, "Signal reconstruction from phase or magnitude," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 28, no. 6, pp. 672–680, Dec. 1980.

[32] A. V. Oppenheim and J. S. Lim, "The importance of phase in signals," *Proc. IEEE*, vol. 69, no. 5, pp. 529–541, May 1981.

[33] X. Hou and L. Zhang, "Saliency detection: A spectral residual approach," in *Proc. IEEE Int. Conf. Comput. Vis. Pattern Recognit.*, Jun. 2007, pp. 1–8.

[34] K. Gu, G. Zhai, X. Yang, W. Zhang, and M. Liu, "Subjective and objective quality assessment for images with contrast change," in *Proc. IEEE Int. Conf. Image Process.*, Sep. 2013, pp. 383–387.

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

[36] D. H. Johnson and S. Sinanović, "Symmetrizing the Kullback–Leibler distance," *IEEE Trans. Inf. Theory*, vol. 1, no. 1, pp. 1–10, 2001.

[37] K. Gu, G. Zhai, M. Liu, X. Yang, and W. Zhang, "Details preservation inspired blind quality metric of tone mapping methods," in *Proc. IEEE Int. Symp. Circuits Syst.*, Jun. 2014, pp. 518–521.
[38] K. Gu, G. Zhai, X. Yang, and W. Zhang, "Hybrid no-reference quality metric for singly and multiply distorted images," *IEEE Trans. Broadcast.*, vol. 60, no. 3, pp. 555–567, Sep. 2014.
[39] M. C. Morrone, J. Ross, D. C. Burr, and R. Owens, "Mach bands are phase dependent," *Nature*, vol. 324, pp. 250–253, Nov. 1986.
[40] B. Jähne, H. Haubecker, and P. Geibler, *Handbook of Computer Vision and Applications*. New York, NY, USA: Academic, 1999.
[41] 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, Apr. 2004.
[42] Z. Wang, E. P. Simoncelli, and A. C. Bovik, "Multiscale structural similarity for image quality assessment," in *Proc. IEEE 27th Asilomar Conf. Signals, Syst. Comput.*, Nov. 2003, pp. 1398–1402.
[43] K. Friston, "The free-energy principle: A unified brain theory?" *Nature Rev. Neurosci.*, vol. 11, pp. 127–138, Feb. 2010.
[44] VQEG. (Mar. 2000). *Final Report From the Video Quality Experts Group on the Validation of Objective Models of Video Quality Assessment*. [Online]. Available: http://www.vqeg.org/
[45] A. K. Moorthy and A. C. Bovik, "Visual importance pooling for image quality assessment," *IEEE J. Sel. Topics Signal Process.*, vol. 3, no. 2, pp. 193–201, Apr. 2009.
[46] K. Gu, G. Zhai, X. Yang, and W. Zhang, "No-reference stereoscopic IQA approach: From nonlinear effect to parallax compensation," *J. Elect. Comput. Eng.*, vol. 2012, Sep. 2012, Art. ID 436031.
[47] K. Gu, G. Zhai, X. Yang, and W. Zhang, "A new psychovisual paradigm for image quality assessment: From differentiating distortion types to discriminating quality conditions," *Signal, Image Video Process.*, vol. 7, no. 3, pp. 423–436, May 2013.
[48] T. Judd, F. Durand, and A. Torralba, "A benchmark of computational models of saliency to predict human fixations," Massachusetts Inst. Technol., Cambridge, MA, USA, Tech. Rep., 2012.
[49] D. Martin, C. Fowlkes, D. Tal, and J. Malik, "A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics," in *Proc. 18th IEEE Int. Conf. Comput. Vis.*, 2001, pp. 416–423.
[50] G. A. F. Seber, *Multivariate Observations*. Hoboken, NJ, USA: Wiley, 1984.

Ke Gu received the B.S. degree in electronic engineering from Shanghai Jiao Tong University, Shanghai, China, in 2009, where he is currently working toward the Ph.D. degree.

He was a visiting student with the Department of Electrical and Computer Engineering, University of Waterloo, Waterloo, ON, Canada, in 2014. His research interests include image quality assessment, contrast enhancement, and visual saliency detection.

Dr. Gu is also a reviewer for several IEEE Transactions and Journals, including IEEE TRANSACTIONS ON CYBERNETICS, IEEE SIGNAL PROCESSING LETTERS, the *Journal of Visual Communication and Image Representation*, and *Signal, Image and Video Processing*.

Guangtao Zhai (M'10) received the B.E. and M.E. degrees from Shandong University, Jinan, China, in 2001 and 2004, respectively, and the Ph.D. degree from Shanghai Jiao Tong University, Shanghai, China, in 2009.

He was a Student Intern with the Institute for Infocomm Research, Singapore, from 2006 to 2007, and a visiting student with the School of Computer Engineering, Nanyang Technological University, Singapore, from 2007 to 2008. From 2008 to 2009, he was a visiting student with the Department of Electrical and Computer Engineering, McMaster University, Hamilton, ON, Canada, where he was a Post-Doctoral Fellow from 2010 to 2012. From 2012 to 2013, he was a Humboldt Research Fellow with the Institute of Multimedia Communication and Signal Processing, Friedrich Alexander University of Erlangen-Nuremberg, Erlangen, Germany. He is currently a Research Professor with the Institute of Image Communication and Information Processing, Shanghai Jiao Tong University. His research interests include multimedia signal processing and perceptual signal processing.

Prof. Zhai received the National Excellent Ph.D. Thesis Award from the Ministry of Education of China in 2012.

Xiaokang Yang (M'00–SM'04) received the B.S. degree from Xiamen University, Xiamen, China, in 1994; the M.S. degree from the Chinese Academy of Sciences, Shanghai, China, in 1997; and the Ph.D. degree from Shanghai Jiao Tong University, Shanghai, in 2000.

He was a Research Fellow with the Centre for Signal Processing, Nanyang Technological University, Singapore, from 2000 to 2002, and a Research Scientist with the Institute for Infocomm Research, Singapore, from 2002 to 2004. He is currently a Full Professor and the Deputy Director of the Department of Electronic Engineering with the Institute of Image Communication and Information Processing, Shanghai Jiao Tong University. He actively participates in the International Standards, such as MPEG-4, JVT, and MPEG-21. He has authored over 80 refereed papers, and holds six patents. His research interests include video processing and communication, media analysis and retrieval, perceptual visual processing, and pattern recognition.

Prof. Yang is a member of the Visual Signal Processing and Communications Technical Committee of the IEEE Circuits and Systems Society. He was a recipient of the Microsoft Young Professorship Award in 2006, the Best Young Investigator Paper Award at IS&T/SPIE International Conference on Video Communication and Image Processing in 2003, and awards from the Agency for Science, Technology and Research and Tan Kah Kee Foundations. He was the Special Session Chair of Perceptual Visual Processing of the IEEE International Conference on Multimedia and Expo in 2006. He is the local Co-Chair of the International Conference on Communications and Networking in China in 2007 and Technical Program Co-Chair of the IEEE Workshop on Signal Processing Systems in 2007.

Wenjun Zhang (M'02–SM'10–F'11) received the B.S., M.S., and Ph.D. degrees in electronic engineering from Shanghai Jiao Tong University, Shanghai, China, in 1984, 1987, and 1989, respectively.

He was a Post-Doctoral Fellow with Philips Kommunikation Industrie AG, Nuremberg, Germany, from 1990 to 1993, where he was actively involved in developing HD-MAC system. He joined the faculty of Shanghai Jiao Tong University, in 1993, and became a Full Professor with the Department of Electronic Engineering in 1995. He successfully developed the first Chinese HDTV Prototype System as the National HDTV Technical Executive Experts Group Project Leader in 1998. He was one of the main contributors to the Chinese Digital Television Terrestrial Broadcasting Standard issued in 2006 and has led the team in designing the next-generation of broadcast television system in China since 2011. He has authored 90 papers in international journals and conferences and holds over 40 patents. His current research interests include digital video coding and transmission, multimedia semantic processing, and intelligent video surveillance.

Prof. Zhang is the Chief Scientist of the Chinese National Engineering Research Centre of Digital Television, an industry/government consortium in DTV technology research and standardization, and the Chair of the Future of Broadcast Television Initiative Technical Committee.

Chang Wen Chen (F'04) received the B.S. degree from the University of Science and Technology of China, Hefei, China, in 1983, the M.S.E.E. degree from the University of Southern California, Los Angeles, CA, USA, in 1986, and the Ph.D. degree from the University of Illinois at Urbana-Champaign, Champaign, IL, USA, in 1992.

He was with the Faculty of Electrical and Computer Engineering, University of Rochester, Rochester, NY, USA, from 1992 to 1996, and the Faculty of Electrical and Computer Engineering, University of Missouri-Columbia, Columbia, MO, USA, from 1996 to 2003. He was the Allen Henry Endow Chair Professor with the Florida Institute of Technology, Melbourne, FL, USA, from 2003 to 2007. He is currently a Professor of Computer Science and Engineering with the University at Buffalo, The State University of New York, Buffalo, NY, USA.

Prof. Chen is a fellow of the International Society for Optics and Photonics (SPIE). He was a recipient of several research and professional achievement awards, including the Sigma Xi Excellence in Graduate Research Mentoring Award in 2003, the Alexander von Humboldt Research Award in 2009, and the State University of New York at Buffalo Exceptional Scholar Sustained Achievement Award in 2012. He and his students have received seven Best Paper Awards or Best Student Paper Awards over the past two decades. He has been the Editor-in-Chief of the IEEE TRANSCATIONS ON MULTIMEDIA since 2014. He served as the Editor-in-Chief of the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY from 2006 to 2009. He has been an Editor for several other major IEEE TRANSACTIONS and Journals, including the PROCEEDINGS OF THE IEEE, the IEEE JOURNAL OF SELECTED AREAS IN COMMUNICATIONS, and the IEEE JOURNAL ON EMERGING AND SELECTED TOPICS IN CIRCUITS AND SYSTEMS. He has served as the Conference Chair of several major IEEE, ACM, and SPIE conferences related to multimedia video communications and signal processing. His research is supported by NSF, DARPA, Air Force, NASA, Whitaker Foundation, Microsoft, Intel, Kodak, Huawei, and Technicolor.