

High Dynamic Range Image Compression by Optimizing Tone Mapped Image Quality Index

Kede Ma, *Student Member, IEEE*, Hojatollah Yeganeh, *Member, IEEE*, Kai Zeng, and Zhou Wang, *Fellow, IEEE*

Abstract—Tone mapping operators (TMOs) aim to compress high dynamic range (HDR) images to low dynamic range (LDR) ones so as to visualize HDR images on standard displays. Most existing TMOs were demonstrated on specific examples without being thoroughly evaluated using well-designed and subject-validated image quality assessment models. A recently proposed tone mapped image quality index (TMQI) made one of the first attempts on objective quality assessment of tone mapped images. Here, we propose a substantially different approach to design TMO. Instead of using any predefined systematic computational structure for tone mapping (such as analytic image transformations and/or explicit contrast/edge enhancement), we directly navigate in the space of all images, searching for the image that optimizes an improved TMQI. In particular, we first improve the two building blocks in TMQI—structural fidelity and statistical naturalness components—leading to a TMQI-II metric. We then propose an iterative algorithm that alternatively improves the structural fidelity and statistical naturalness of the resulting image. Numerical and subjective experiments demonstrate that the proposed algorithm consistently produces better quality tone mapped images even when the initial images of the iteration are created by the most competitive TMOs. Meanwhile, these results also validate the superiority of TMQI-II over TMQI.¹

Index Terms—High dynamic range image, image quality assessment, tone mapping operator, perceptual image processing, structural similarity, statistical naturalness.

I. INTRODUCTION

THE luminance of a natural scene often has a high dynamic range (HDR), varying between 10^{-3} to 10^5 cd/m². However, a normal digital display only has a low dynamic range (LDR) about 10^2 cd/m² [1]. Tone mapping operators (TMOs) fill in the gap between HDR imaging and visualizing HDR images on standard displays by compressing the dynamic range of HDR images [2]. TMOs provide a useful surrogate for HDR display technology, which is currently still expensive. Regardless of how fast HDR display technology penetrates the market, there will be a strong need to prepare HDR imagery for display on LDR devices [2]. In addition,

compressing the dynamic range of an HDR image while preserving its structural detail and natural appearance is by itself an interesting and challenging problem for human and computer vision study.

In recent years, many TMOs have been proposed [3]–[9]. Most of them were demonstrated on specific examples without being thoroughly evaluated using well-designed and subject-validated image quality assessment (IQA) models. With multiple TMOs at hand, a natural question is: which TMO produces the best quality tone mapped LDR image? This question could possibly be answered by subjective evaluation [10]–[13], which is expensive, time consuming, and perhaps most importantly, can hardly be used to guide automatic optimization procedures [14].

A promising approach is to develop objective IQA models that can automatically evaluate the performance of TMOs. Traditional objective IQA metrics such as peak signal-to-noise ratio and the structural similarity index (SSIM) [14]–[16] assume that the reference and compared images have the same dynamic range; thus they are not applicable in this scenario. Some attempts have been made for objectively assessing the quality of HDR images. The HDR visible difference predictor [17] tries to predict the visible difference between two HDR images with the same dynamic range. A dynamic range independent quality measure [18] focuses on detecting the loss of visible contrast, the amplification of invisible contrast, and the reversal of visible contrast. It produces three corresponding probability maps but does not integrate them into an overall quality score. Recently, a tone mapped image quality index (TMQI) is proposed [19], which consists of two fundamental building blocks: structural fidelity and statistical naturalness. The fundamental idea behind TMQI is that a high quality tone mapped image should not only preserve the structural details in the HDR image but also present a natural appearance.

In this paper, we propose a novel TMO that utilizes an improved TMQI as the optimization goal. Specifically, we first develop an improved TMQI, namely TMQI-II, that overcomes the limitations underlying the structural fidelity and statistical naturalness components in TMQI. Experiments show that the improved structural fidelity and statistical naturalness terms better correlate with the subjective data. We then propose an iterative optimization algorithm for tone mapping. Substantially different from existing TMOs, we do not pre-define a computational structure that involves analytic image transformations and/or explicit gradient/edge estimation and enhancement operations. Instead, we directly operate in the space

Manuscript received June 24, 2014; accepted May 13, 2015. Date of publication May 21, 2015; date of current version June 9, 2015. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Sergio Goma.

The authors are with the Department of Electrical and Computer Engineering, University of Waterloo, Waterloo, ON, N2L 3G1, Canada (e-mail: k29ma@uwaterloo.ca; hyeganeh@uwaterloo.ca; kzeng@uwaterloo.ca; zhouwang@ieee.org).

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/TIP.2015.2436340

¹Partial preliminary results of this work were presented at *ICASSP 2013* and *ICME 2014*.

of all images. Starting from any given image as the initial point, we move it towards the direction that improves TMQI-II. In each iteration, we alternately improve the structural fidelity and statistical naturalness of the resulting image. Numerical and subjective experiments show that this iterative algorithm consistently produces better quality tone mapped images even when the initial images are created by the most competitive state-of-the-art TMOs. Meanwhile, the superiority of TMQI-II over TMQI is also verified through this process.

II. ITERATIVE TONE MAPPING BY OPTIMIZING TMQI-II

Let \mathbf{X} and \mathbf{Y} be the HDR image and the tone mapped LDR image, respectively. TMQI suggests that a high quality tone mapped image should achieve great structural fidelity with respect to the HDR image and high statistical naturalness simultaneously. The computation of TMQI is given by [19]

$$\text{TMQI}(\mathbf{X}, \mathbf{Y}) = a[S(\mathbf{X}, \mathbf{Y})]^a + (1 - a)[N(\mathbf{Y})]^\beta, \quad (1)$$

where S and N denote the structural fidelity and statistical naturalness measures, respectively. The parameters a and β determine the sensitivities of the two terms, and $0 \leq a \leq 1$ adjusts the relative importance between them. Both S and N are upper bounded by 1 and thus TMQI is also upper bounded by 1.

As one of the first attempts on quality evaluation of images across dynamic ranges, TMQI achieved remarkable success, but as will be shown later, it also has significant limitations. Here, we propose an improved TMQI, namely TMQI-II, that overcomes the limitations to better correlate with subjective evaluations. Details of TMQI-II will be elaborated later along with the discussions regarding the structural fidelity and statistical naturalness components.

Assuming TMQI-II to be the quality criterion of tone mapped images, the problem of optimal tone mapping can be formulated as

$$\mathbf{Y}_{\text{opt}} = \arg \max_{\mathbf{Y}} \text{TMQI-II}(\mathbf{X}, \mathbf{Y}), \quad (2)$$

where \mathbf{Y} has a much lower dynamic range than \mathbf{X} . Solving (2) for \mathbf{Y}_{opt} is a challenging problem due to the complexity of TMQI-II and the high dimensionality (the same as the number of pixels in the image). Therefore, we resort to numerical optimization and propose an iterative approach. Specifically, given any initial image \mathbf{Y}_0 , we move it towards the direction in the space of images that improves TMQI-II. To accomplish that, we first improve the structural fidelity S using a gradient ascent method and then enhance the statistical naturalness N by solving a parameter optimization problem for a point-wise intensity transformation. These two steps constitute one iteration and the iterations continue until convergence.

A. Structural Fidelity Update

The structural fidelity of TMQI is computed using a sliding window across the entire image, which results in a quality map that indicates local structural detail preservation. Let \mathbf{x} and \mathbf{y} be two image patches within the sliding window in the HDR

and tone mapped images, respectively. The local structural fidelity measure is defined as

$$S_{\text{local}}(\mathbf{x}, \mathbf{y}) = \frac{2\tilde{\sigma}_x\tilde{\sigma}_y + C_1}{\tilde{\sigma}_x^2 + \tilde{\sigma}_y^2 + C_1} \cdot \frac{\sigma_{xy} + C_2}{\sigma_x\sigma_y + C_2}, \quad (3)$$

where σ_x , σ_y and σ_{xy} denote the local standard deviations (std) and covariance between the two corresponding patches, respectively. C_1 and C_2 are two small positive constants to avoid instability. The first component is modified from the local contrast comparison term in SSIM [15]. It suggests that the HDR and tone mapped image patches should keep the same contrast visibility; otherwise, the contrast of the tone mapped image patch should be penalized, which corresponds to either artificially creating visible contrast or failing to preserve visible contrast. The second component is the same as the structure comparison term in SSIM [15]. The overall structural fidelity measure of the image is computed by averaging all local structural fidelity measures

$$S(\mathbf{X}, \mathbf{Y}) = \frac{1}{M} \sum_{i=1}^M S_{\text{local}}(\mathbf{x}_i, \mathbf{y}_i), \quad (4)$$

where \mathbf{x}_i and \mathbf{y}_i are the i -th patches in \mathbf{X} and \mathbf{Y} , respectively, and M is the total number of patches.

In TMQI [19], to assess the visibility of local contrast, the local std σ undergoes a nonlinear function motivated by a contrast sensitivity model

$$\tilde{\sigma} = \frac{1}{\sqrt{2\pi}\theta_\sigma} \int_{-\infty}^{\sigma} \exp\left[-\frac{(t - \tau_\sigma)^2}{2\theta_\sigma^2}\right] dt, \quad (5)$$

where τ_σ is a threshold determined by the contrast sensitivity function [20] and $\theta_\sigma = \tau_\sigma/3$ [21].

The above nonlinear function is limited in accurately assessing the contrast visibility of HDR image patches. First, even a small change in local patch of the HDR image (which may result from the HDR camera noise) may contribute to a significant σ value. When Eq. (5) effectively distinguishes the visible and invisible local contrast in the tone mapped image, it tends to label most patches, either visible or invisible in the HDR image, as contrast visible. Fig. 1 illustrates this phenomenon. The homogeneous areas such as the walls and the wood board in the lower middle part of the image are correctly predicted as contrast invisible in the tone mapped image but mistakenly marked as contrast visible in the HDR image due to oversensitivity to noise. This explains the corresponding dark areas of the structural fidelity map in Fig. 1(b) (where brighter indicates higher structural fidelity). Second, different local patches in the HDR image may have substantially different dynamic ranges, which correspond to different thresholds τ_σ . In other words, a single τ_σ is insufficient to account for the local contrast visibility of the HDR image.

The above analysis suggests that a contrast visibility model adapted to local luminance levels is desired for the HDR image. In particular, we follow [22] and choose σ/μ , namely the coefficient of variation, as an estimate of local contrast in the HDR image, where μ is the local mean. This estimate is adapted to local luminance levels, and thus is qualitatively consistent with Weber's law, which has been widely

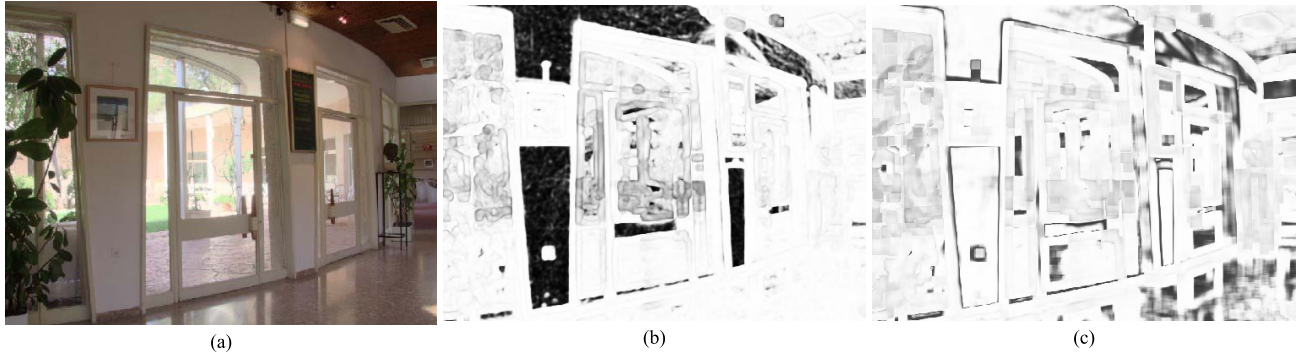


Fig. 1. Tone mapped “Belgium house” image and its structural fidelity maps. (a) Initial image created by Reinhard’s algorithm [3]. (b) and (c) Structural fidelity maps generated by TMQI and TMQI-II respectively, where brighter indicates higher structural fidelity.

used to model luminance masking in the human visual system. An additional benefit of this estimate is that it is invariant to linear contrast stretch, which is a frequently used pre-processing step for HDR images. The reason follows directly from the local luminance adaptation that cancels out the scale factors in the numerator and the denominator. Fig. 1(c) shows an example of the structural fidelity map from the modified structural fidelity term, which captures the contrast visibility of the HDR and the tone mapped images more reasonably. Table I also demonstrates the superiority of the modified structural fidelity term on the subjective database [19] using Spearman’s rank-order correlation coefficient (SRCC) and Kendall’s rank-order correlation coefficient (KRCC) as the evaluation criteria.

Given the modified structural fidelity term, we adopt a gradient ascent algorithm similar to [23] and [24] to improve the structural fidelity of the resulting image \mathbf{Y}_k from the k -th iteration. To do that, we compute the gradient of $S(\mathbf{X}, \mathbf{Y})$ with respect to \mathbf{Y} , denoted by $\nabla_{\mathbf{Y}}S(\mathbf{X}, \mathbf{Y})$ and update the image by

$$\hat{\mathbf{Y}}_k = \mathbf{Y}_k + \lambda \nabla_{\mathbf{Y}}S(\mathbf{X}, \mathbf{Y}) \Big|_{\mathbf{Y}=\mathbf{Y}_k}, \quad (6)$$

where λ is the step size. To compute the gradient $\nabla_{\mathbf{Y}}S(\mathbf{X}, \mathbf{Y})$, we start from the local structural fidelity and rewrite (3) as

$$S_{\text{local}}(\mathbf{x}, \mathbf{y}) = \frac{A_1 A_2}{B_1 B_2}, \quad (7)$$

where

$$A_1 = 2\tilde{\sigma}_x \tilde{\sigma}_y + C_1 \quad (8)$$

$$B_1 = \tilde{\sigma}_x^2 + \tilde{\sigma}_y^2 + C_1 \quad (9)$$

$$A_2 = \sigma_{xy} + C_2 \quad (10)$$

$$B_2 = \sigma_x \sigma_y + C_2. \quad (11)$$

By treating both image patches as column vectors of length N_w , we have the sample statistics given by

$$\mu_y = \frac{1}{N_w} \mathbf{1}^T \mathbf{y} \quad (12)$$

$$\sigma_y^2 = \frac{1}{N_w} (\mathbf{y} - \mu_y)^T (\mathbf{y} - \mu_y) \quad (13)$$

$$\sigma_{xy} = \frac{1}{N_w} (\mathbf{x} - \mu_x)^T (\mathbf{y} - \mu_y), \quad (14)$$

where $\mathbf{1}$ is a N_w -vector with all entries equal to 1. The gradient of the local structural fidelity measure with respect to \mathbf{y} can

then be expressed as

$$\nabla_{\mathbf{y}} S_{\text{local}}(\mathbf{x}, \mathbf{y}) = \frac{(A'_1 A_2 + A_1 A'_2)}{B_1 B_2} - \frac{(B'_1 B_2 + B_1 B'_2) A_1 A_2}{(B_1 B_2)^2}, \quad (15)$$

where

$$A'_1 = \nabla_{\mathbf{y}} A_1, \quad B'_1 = \nabla_{\mathbf{y}} B_1, \quad A'_2 = \nabla_{\mathbf{y}} A_2, \quad B'_2 = \nabla_{\mathbf{y}} B_2. \quad (16)$$

Noting that

$$\nabla_{\mathbf{y}} \sigma_y = \frac{1}{N_w \sigma_y} (\mathbf{y} - \mu_y) \quad (17)$$

$$\nabla_{\mathbf{y}} \sigma_{xy} = \frac{1}{N_w} (\mathbf{x} - \mu_x), \quad (18)$$

we have

$$\begin{aligned} A'_1 &= 2\tilde{\sigma}_x \nabla_{\mathbf{y}} \tilde{\sigma}_y \\ &= \frac{2\tilde{\sigma}_x}{\sqrt{2\pi}\theta_\sigma} \exp\left[-\frac{(\sigma_y - \tau_\sigma)^2}{2\theta_\sigma^2}\right] \cdot \nabla_{\mathbf{y}} \sigma_y \\ &= \sqrt{\frac{2}{\pi}} \frac{\tilde{\sigma}_x}{N_w \theta_\sigma \sigma_y} \exp\left[-\frac{(\sigma_y - \tau_\sigma)^2}{2\theta_\sigma^2}\right] (\mathbf{y} - \mu_y), \end{aligned} \quad (19)$$

$$\begin{aligned} B'_1 &= 2\tilde{\sigma}_y \nabla_{\mathbf{y}} \tilde{\sigma}_y \\ &= \sqrt{\frac{2}{\pi}} \frac{\tilde{\sigma}_y}{N_w \theta_\sigma \sigma_y} \exp\left[-\frac{(\sigma_y - \tau_\sigma)^2}{2\theta_\sigma^2}\right] (\mathbf{y} - \mu_y), \end{aligned} \quad (20)$$

$$A'_2 = \frac{1}{N_w} (\mathbf{x} - \mu_x), \quad (21)$$

$$B'_2 = \sigma_x \nabla_{\mathbf{y}} \sigma_y = \frac{\sigma_x}{N_w \sigma_y} (\mathbf{y} - \mu_y). \quad (22)$$

Plugging (8), (9), (10), (11), (19), (20), (21) and (22) into (15), we obtain the gradient of local structural fidelity. Finally, we compute the gradient of the overall structural fidelity measure with respect to the tone mapped image \mathbf{Y} by summing over all the local gradients

$$\nabla_{\mathbf{Y}} S(\mathbf{X}, \mathbf{Y}) = \frac{1}{M} \sum_{i=1}^M \mathbf{R}_i^T \nabla_{\mathbf{y}} S_{\text{local}}(\mathbf{x}, \mathbf{y}) \Big|_{\mathbf{x}=\mathbf{x}_i, \mathbf{y}=\mathbf{y}_i}. \quad (23)$$

where $\mathbf{x}_i = \mathbf{R}_i(\mathbf{X})$ and $\mathbf{y}_i = \mathbf{R}_i(\mathbf{Y})$ are the i -th image patches, \mathbf{R}_i is the operator that takes the i -th local patch from the image, and \mathbf{R}_i^T places the patch back into the corresponding location in the image.

B. Statistical Naturalness Update

The statistical naturalness N in TMQI is constructed by modeling the histograms of μ and σ of about 3000 natural images by a Gaussian density function P_m and a Beta density function P_d , respectively [19]. Based on the independence characteristic of image brightness and contrast [22], the two density functions are multiplied to obtain the overall statistical naturalness measure [19]

$$N(\mathbf{Y}) = \frac{1}{K} P_m P_d, \quad (24)$$

where K is a normalization factor.

The above statistical naturalness model has two limitations. First, P_m and P_d are assumed to be completely independent of image content, which is an over simplification. The model suggests that to be highly statistically natural, the tone mapped image of dynamic range $[0, 255]$ should have μ around 116 and σ around 65 which correspond to the peaks in P_m and P_d , respectively [19]. However, each image may have a different μ and σ to look perfectly natural depending on its content. Second, the model is derived from high quality images, with no information about what an unnatural image may look like.

Here we propose an image dependent statistical naturalness model based on a subjective experiment to better quantify the unnaturalness of tone mapped images. First, we estimate the overall luminance and global contrast directly from the HDR image, denoted by μ_e and σ_e , respectively. To do that, we approximate the overall luminance level of the HDR image to its log-mean luminance, which has been successfully used in previous studies [3], [4], [25], [26]. The use of logarithmic function assumes that most structural detail in the HDR image live in a low dynamic range and thus it is reasonable to boost lower luminance levels while compress higher luminance levels. The quantity is computed by

$$L_x = \exp \left(\frac{1}{|\mathbf{X}|} \sum_{i,j} \log(\epsilon + \mathbf{X}(i, j)) \right), \quad (25)$$

where $\mathbf{X}(i, j)$ is the luminance of the HDR image at location (i, j) , $|\mathbf{X}|$ is the cardinality and ϵ is a small positive constant to avoid instability. Next, the luminance is scaled by

$$\mathbf{X}^s(i, j) = \frac{k}{L_x} \mathbf{X}(i, j), \quad (26)$$

where k is a luminance level related quantity typically set between 0.09 and 0.36 for an HDR image with normal luminance level [3]. μ_e and σ_e are then estimated by

$$\mu_e = \frac{L}{|\mathbf{X}|} \left(\sum_{i,j} \frac{\mathbf{X}^s(i, j)}{1 + \mathbf{X}^s(i, j)} \right) \quad (27)$$

and

$$\sigma_e = \frac{1}{|\mathbf{X}| - 1} \sum_{i,j} \left(\frac{\mathbf{X}^s(i, j)}{1 + \mathbf{X}^s(i, j)} L - \mu_e \right)^{\frac{1}{2}}, \quad (28)$$

where L is the dynamic range of the tone mapped image. Here, the high luminance is further compressed by a

factor of \mathbf{X}^s . This inevitably causes detail loss in high luminance areas. Nevertheless, our goal here is to roughly estimate μ_e and σ_e that are relevant to a natural appearance of the tone mapped image. As for the detail loss, it can typically be well captured by the structural fidelity term. This estimation of initial luminance level of the LDR image is closely related to previous works [3], [27], [28].

μ_e and σ_e are only rough estimates of the desired μ and σ values. For each LDR image, there should be certain ranges of μ and σ values surrounding μ_e and σ_e , within which the naturalness of the image is not degraded. To verify this and to provide a quantitative model, we conducted a subjective experiment, in which observers were asked to gradually decrease and then increase μ of the test LDR images until they saw significant degradation in naturalness. A lower bound μ_l and an upper bound μ_r for each LDR image were thus recorded. The same procedure is used to obtain a lower std bound σ_l and an upper std bound σ_r for each LDR image. We selected 60 natural LDR images from the LIVE database [29] with different μ and σ values that cover diverse natural contents. A total of 25 naive observers, including 15 males and 10 females aged between 22 and 30, participated in the experiment. The four bounds for each LDR image are averaged over all 25 observers. Fig. 2 summarizes the experimental results, where we observe that the relationships between μ and the values of μ_l and μ_r are approximately linear. This motivates us to fit two linear models to predict μ_l and μ_r on the basis of μ . The fitted models have slopes $k_1 = 0.60$, $k_2 = 0.70$ and intercepts $b_1 = -0.14$, $b_2 = 83.61$ for μ_l and μ_r , respectively. The R^2 statistics of the two linear models are 0.8008 and 0.8465 respectively, which indicate that the linear models explain most variances in the subjective data. Perhaps the most interesting finding in this experiment is when μ of an image is relatively small, $\mu_r - \mu$ is much larger than $\mu - \mu_l$. By contrast, the situation is reversed when μ of an image is large. In words, the acceptable luminance changes without significantly tampering an image's visual naturalness saturate at both small and large luminance levels. Similarly, σ_l and σ_r of a test LDR image can also be fitted by two linear models using σ as the predictor. The fitted lines for σ_l and σ_r have slopes $k_3 = 0.65$, $k_4 = 0.94$ and intercepts $b_3 = -0.08$, $b_4 = 51.40$, respectively.

Based on the method described above, given an HDR image, we first estimate μ_e and σ_e and then predict μ_l , μ_r , σ_l and σ_r of the tone mapped image. The μ and σ values of a natural looking tone mapped image should at least fall in $[\mu_l, \mu_r]$ and $[\sigma_l, \sigma_r]$, and if possible, close to μ_e and σ_e . We quantify the drop from μ_e and σ_e to their lower and upper bounds using Gaussian cumulative distribution functions (CDF). Specifically, the likelihood of a tone mapped image to be natural given its mean μ is computed by

$$P_m = \begin{cases} \frac{1}{\sqrt{2\pi}\theta_1} \int_{-\infty}^{\mu} \exp \left[-\frac{(t - \tau_1)^2}{2\theta_1^2} \right] dt & \mu \leq \mu_e \\ \frac{1}{\sqrt{2\pi}\theta_2} \int_{-\infty}^{2\mu_r - \mu} \exp \left[-\frac{(t - \tau_2)^2}{2\theta_2^2} \right] dt & \mu > \mu_e \end{cases} \quad (29)$$

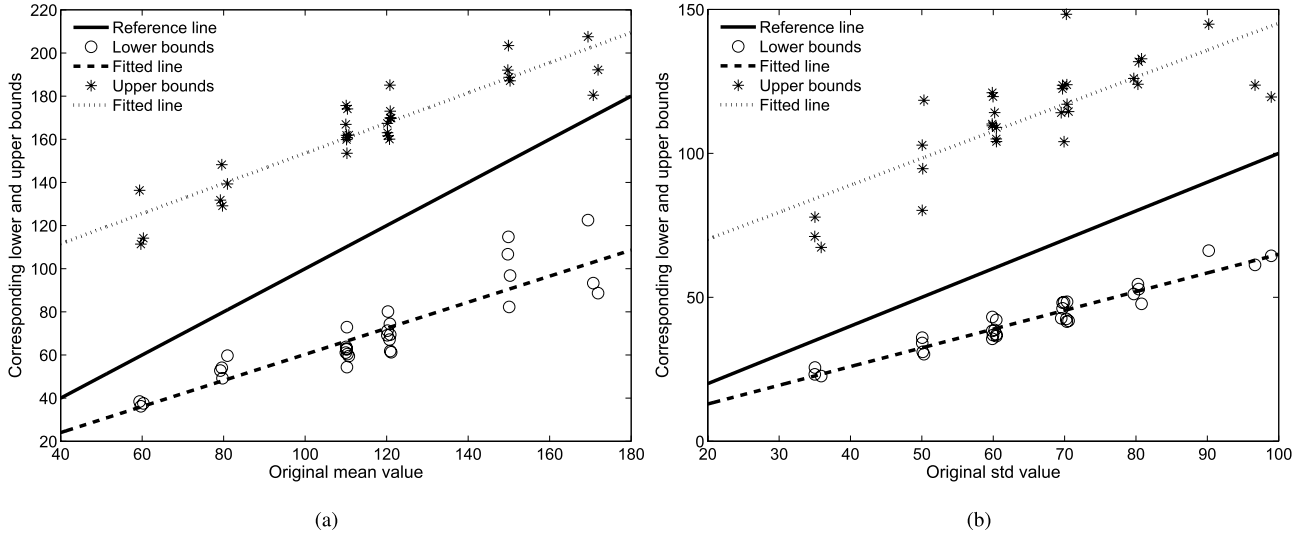


Fig. 2. Subjective experiment results on the naturalness of mean and std values. The asterisks in (a) and (b) represent the upper bounds, given by the average subject, of the mean and std of test images, respectively. The corresponding dotted lines are least square fitted lines. The circles in (a) and (b) represent the lower bounds, given by the average subject, of the mean and std of test images, respectively. The corresponding dashed lines are least square fitted lines. The solid lines in two plots are reference lines that correspond to μ_e and σ_e , respectively.

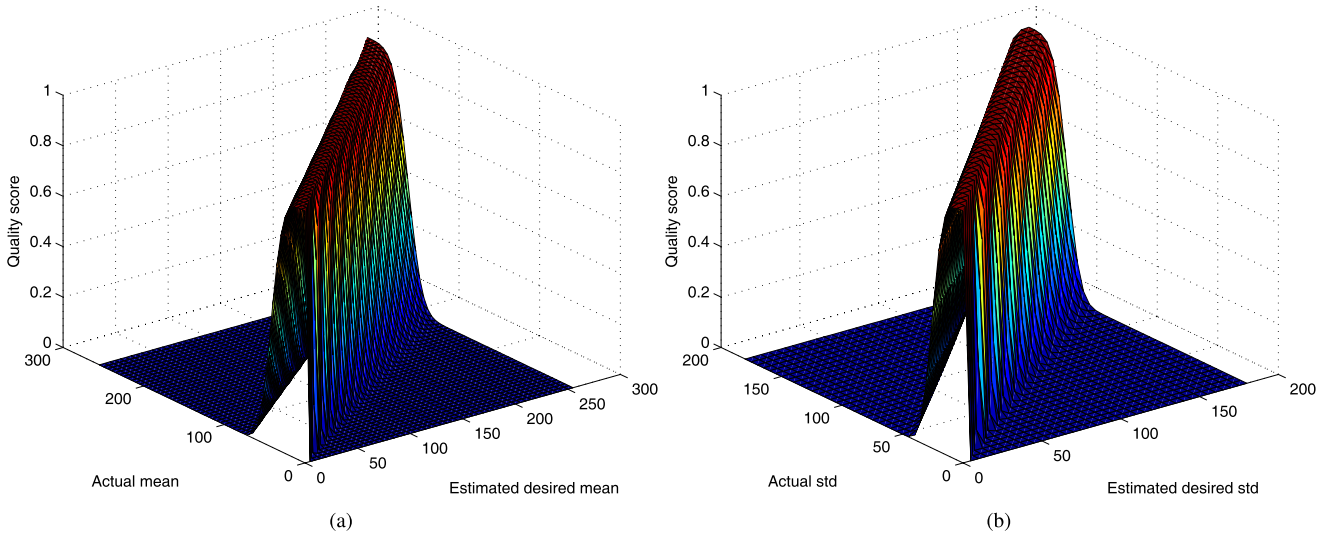


Fig. 3. The surfaces of P_m and P_d in TMQL_II. (a) P_m (b) P_d . It suggests that the μ and σ of the natural looking tone mapped image should be close to μ_e and σ_e , which correspond to the peaks along the diagonal lines. The models give heavy penalty when $|\mu - \mu_e|$ or $|\sigma - \sigma_e|$ is large.

where τ_1 and θ_1 are uniquely determined by two points $(\mu_l, 0.01)$ and $(\mu_e, 1)$ on the Gaussian CDF curve. Correspondingly, τ_2 and θ_2 are uniquely determined by two points $(\mu_r, 0.01)$ and $(\mu_e, 1)$ on the Gaussian CDF curve. Similarly, the likelihood of a tone mapped image to be natural given its std σ is computed by

$$P_d = \begin{cases} \frac{1}{\sqrt{2\pi}\theta_3} \int_{-\infty}^{\sigma} \exp\left[-\frac{(t-\tau_3)^2}{2\theta_3^2}\right] dt & \sigma \leq \sigma_e \\ \frac{1}{\sqrt{2\pi}\theta_4} \int_{-\infty}^{2\sigma_r-\sigma} \exp\left[-\frac{(t-\tau_4)^2}{2\theta_4^2}\right] dt & \sigma > \sigma_e, \end{cases} \quad (30)$$

where τ_3 and θ_3 are uniquely determined by two points $(\sigma_l, 0.01)$ and $(\sigma_e, 1)$ on the Gaussian CDF curve, and τ_4 and θ_4 are uniquely determined by two points $(\sigma_r, 0.01)$ and $(\sigma_e, 1)$ on the Gaussian CDF curve.

The surfaces of P_m and P_d are plotted in Fig. 3. It can be observed that μ and σ of the tone mapped image of high naturalness should be close to μ_e and σ_e , which correspond to the peaks along the diagonal lines. The models give heavy penalty when $|\mu - \mu_e|$ or $|\sigma - \sigma_e|$ is large.

Similar to Eq. (24), assuming the independence of image luminance and contrast [22], we multiply these two quantities and obtain the overall statistical naturalness model

$$N(\mathbf{X}, \mathbf{Y}) = P_m P_d. \quad (31)$$

Since $0 \leq P_m, P_d \leq 1$, N also lies in $[0, 1]$. The superiority of the modified statistical naturalness term over that in TMQI is verified by improved correlation with respect to subjective evaluations as shown in Table I.

To continue with the iterative optimization procedure upon structural fidelity update, we start with the intermediate

TABLE I

PERFORMANCE EVALUATION OF THE STRUCTURAL FIDELITY AND STATISTICAL NATURALNESS TERMS IN TMQI AND TMQI-II ON THE DATABASE [19]

Image set	SRCC				KRCC			
	TMQI		TMQI-II		TMQI		TMQI-II	
	<i>S</i>	<i>N</i>	<i>S</i>	<i>N</i>	<i>S</i>	<i>N</i>	<i>S</i>	<i>N</i>
1	0.6667	0.9048	0.7381	0.9048	0.5000	0.7857	0.6429	0.7857
2	0.8095	0.7619	0.5714	0.5952	0.7143	0.5714	0.3571	0.4286
3	0.2619	0.7381	0.5714	0.7143	0.1429	0.5714	0.3571	0.5714
4	0.8571	0.8571	0.7619	0.7857	0.7143	0.6429	0.5714	0.5714
5	0.1429	0.7381	0.8810	0.5714	0.1429	0.6429	0.7143	0.4286
6	0.7381	0.9524	0.7143	0.9524	0.6429	0.8571	0.5714	0.8571
7	0.8810	0.6429	0.9286	0.7381	0.7857	0.5000	0.8571	0.6429
8	0.3333	0.7143	0.7619	0.7381	0.2857	0.5714	0.5714	0.6429
9	0.8571	0.3571	0.3095	0.7381	0.7143	0.3571	0.3571	0.6429
10	0.6667	0.9048	0.8095	0.8810	0.5000	0.7857	0.7143	0.7143
11	0.6429	0.5476	0.8095	0.7857	0.4286	0.4286	0.6429	0.6429
12	0.7143	0.5714	0.9524	0.5476	0.5714	0.4286	0.8571	0.3571
13	0.9461	0.4311	0.5629	0.7066	0.8365	0.3273	0.4001	0.5455
14	0.9524	0.7381	0.8333	0.7619	0.8571	0.6429	0.6429	0.5714
15	0.9286	0.9048	0.8810	0.7857	0.8571	0.7857	0.7143	0.6429
Average	0.6932	0.7176	0.7391	0.7471	0.5796	0.5933	0.5981	0.6030

image $\hat{\mathbf{Y}}_k$ in Eq. (6) and improve the statistical naturalness to achieve \mathbf{Y}_{k+1} through a three-segment equipartition monotonic piecewise linear function

$$y_{k+1}^i = \begin{cases} (3/L)a\hat{y}_k^i & 0 \leq \hat{y}_k^i \leq L/3 \\ (3/L)(b-a)\hat{y}_k^i + (2a-b) & L/3 < \hat{y}_k^i \leq 2L/3 \\ (3/L)(L-b)\hat{y}_k^i + (3b-2L) & 2L/3 < \hat{y}_k^i \leq L. \end{cases} \quad (32)$$

This is essentially a point-wise intensity transformation with its parameters a and b (where $0 \leq a \leq b \leq L$) chosen so that μ and σ of $\mathbf{Y}_{k+1} = \{y_{k+1}^i \text{ for all } i\}$ better approximate μ_e and σ_e of the desired tone mapped image.

To solve for a and b , we first estimate the mean and std values of \mathbf{Y}_{k+1} by

$$\begin{aligned} \mu_{k+1}^e &= \hat{\mu}_k + \lambda_m(\mu_e - \hat{\mu}_k) \\ \sigma_{k+1}^e &= \hat{\sigma}_k + \lambda_d(\sigma_e - \hat{\sigma}_k), \end{aligned} \quad (33)$$

where $\hat{\mu}_k$ and $\hat{\sigma}_k$ are the mean and std of $\hat{\mathbf{Y}}_k$, respectively. λ_m and λ_d are step sizes that control the updating speed. Finding the parameters a and b is then converted to solving the following constrained optimization problem

$$\begin{aligned} \{a, b\}_{\text{opt}} &= \arg \min_{\{a, b\}} \|\mu_{k+1} - \mu_{k+1}^e\|^2 + \eta \|\sigma_{k+1} - \sigma_{k+1}^e\|^2 \\ &\text{subject to } 0 \leq a \leq b \leq L, \end{aligned} \quad (34)$$

where η adjusts the weights between the mean and std terms. We adopt a standard gradient projection algorithm [30], [31] with a maximum of 30 iterations to solve this problem. Once the optimal values of a and b are obtained, they are plugged into (32) to create the resulting image \mathbf{Y}_{k+1} , which is subsequently fed into the $(k+1)$ -th iteration.

The structural fidelity update and statistical naturalness update alternate until $\|\mathbf{Y}_{k+1} - \mathbf{Y}_k\|_2$ is smaller than a threshold ϵ . Technically, a similar iterative algorithm works as well if TMQI is chosen as the optimization metric. The only difference lies in Eq. (33), where μ_e and σ_e are replaced with two constants corresponding to the peaks of fixed P_m and P_d models in TMQI. We have five free parameters

in the proposed algorithm. For fair comparison between TMQI and TMQI-II, we use the same set of parameters in all experiments, which are $\epsilon = 0.1$, $\lambda = 0.3$, $\lambda_m = \lambda_d = 0.03$ and $\eta = 1$, respectively.

A by-product of the above derivation of the iterative TMO algorithm is a renewed index, TMQI-II, given by

$$\text{TMQI-II}(\mathbf{X}, \mathbf{Y}) = a[S(\mathbf{X}, \mathbf{Y})]^\alpha + (1-a)[N(\mathbf{X}, \mathbf{Y})]^\beta, \quad (35)$$

where both S and N measures have been improved upon those in TMQI. A number of parameters are inherited from TMQI. These include $C1 = 0.01$, $C2 = 10$ and the threshold of the contrast visibility for tone mapped images $\tau_\sigma = 2.6303$. Throughout our study, we set the threshold of contrast visibility for HDR images $\tau_\sigma = 0.06$ and $k = 0.12$. As for the model parameters in TMQI-II, we set $a = 0.5$, $\alpha = 1$ and $\beta = 1$, which emphasize the equal importance between structural fidelity and statistical naturalness terms.

III. EXPERIMENTAL RESULTS

To fully demonstrate the potentials of the proposed iterative algorithm, we select a wide range of HDR images, containing both indoor and outdoor scenes, human and static objects, as well as day and night views. The initial images for this algorithm are also generated by many different TMOs, ranging from simple ones such as Gamma mapping ($\gamma = 2.2$) and log-normal mapping to state-of-the-art ones such as Durand's method [32], Mantiuk's method [33], Drago's method [4] and Reinhard's method [3]. The last one is considered one of the best TMOs based on several independent subjective tests [12], [19].

We first examine the roles of the structural fidelity and statistical naturalness components separately. In Fig. 4, we start with an initial "Desk" image created by Reinhard's TMO [3] and apply the proposed iterative algorithm but using structural fidelity updates only. It can be observed that the structural fidelity map is very effective at detecting the missing structures (e.g., text in the book region, and fine textures on the desk), and the proposed algorithm successfully recovers such details after a sufficient number of iterations. The improvement

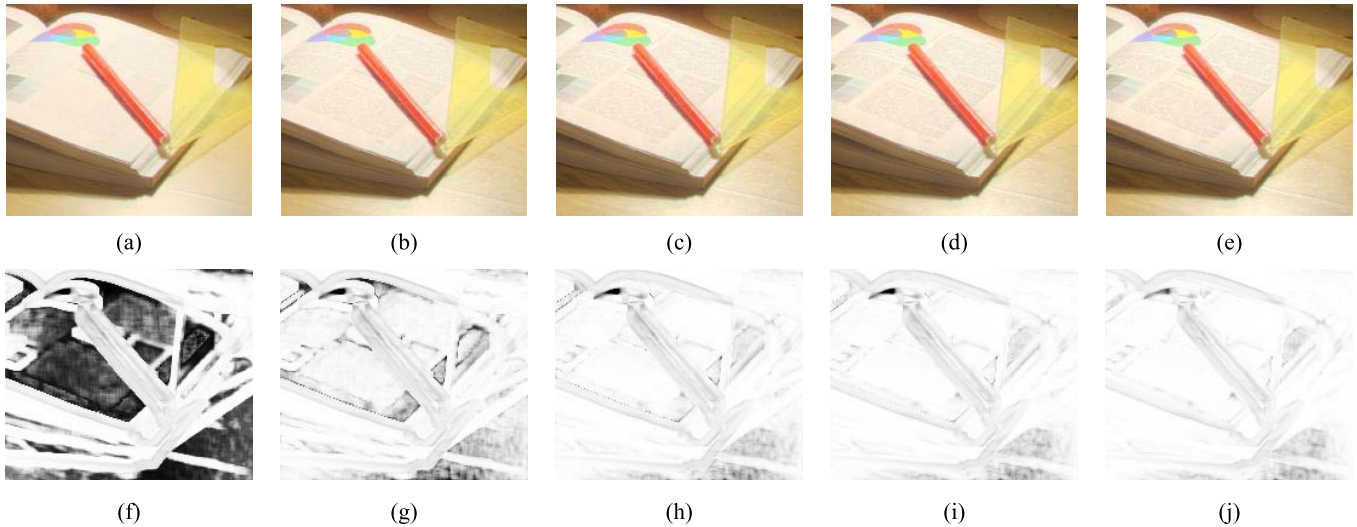


Fig. 4. Tone mapped “Desk” images and their structural fidelity maps. (a) Initial image created by Reinhard’s algorithm [3]. (b)-(e) Images created using iterative structural fidelity update only. (f)-(j) Corresponding structural fidelity maps of (a)-(e), where brighter indicates higher structural fidelity. All images are cropped for better visualization. (a) initial image, $S = 0.689$ (g) 10 iterations, $S = 0.921$ (h) 50 iterations, $S = 0.954$ (i) 100 iterations, $S = 0.961$ (j) 200 iterations, $S = 0.966$.

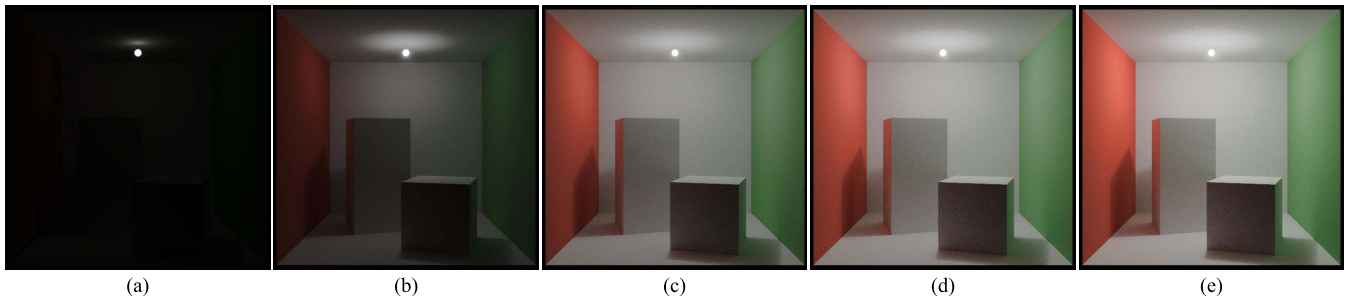


Fig. 5. Tone mapped “Cornell box” images. (a) Initial image created by Gamma mapping ($\gamma = 2.2$). (b)-(e) Images created using iterative statistical naturalness update only. (a) initial image, $N = 0.000$. (b) 10 iterations, $N = 0.0001$. (c) 30 iterations, $N = 0.0329$. (d) 50 iterations, $N = 0.8355$. (e) 100 iterations, $N = 0.9962$.

of structural detail is also well reflected by the structural fidelity map, which eventually evolves to a nearly uniform white image. In Fig. 5, the initial “Cornell box” image is created by the Gamma mapping ($\gamma = 2.2$), and we apply the proposed iterative algorithm but using statistical naturalness updates only. With the iterations, the overall brightness and contrast of the image are significantly improved, leading to a more visually appealing and natural-looking image.

Table II lists the TMQI-II comparison between the initial images and the corresponding converged images after applying the full version of the proposed iterative optimization algorithm. It can be observed that the proposed algorithm consistently converges to images with both high structural fidelity and high statistical naturalness and thus produces higher TMQI-II values even when the initial images are created by the most competitive TMOs.

To validate the superiority of TMQI-II over TMQI as the optimization goal in the proposed iterative algorithm, Fig. 6 shows the comparison of TMQI with TMQI-II optimization results on the “Woman” image initialized by Gamma mapping, which creates dark background with missing structures. Both TMQI and TMQI-II optimized images recover the structures of the background such as the white door, the yellow board

and the photo frame, and present a better overall brightness. However, the TMQI optimized image suffers from heavy noise in homogenous areas such as in the wall and on the floor. The boosted noise artifact are likely due to the structural fidelity term in TMQI, which treats all local areas in the HDR image as contrast visible. In comparison, the TMQI-II optimized image is much cleaner and sharper.

Fig. 7 shows the comparison of TMQI and TMQI-II optimization results on the “Clock building” image. The initial image created by log-normal mapping preserves most structures but looks unrealistic due to its blanced appearance. This problem is largely alleviated in the TMQI-II optimized image, where the overall brightness and contrast of the image are significantly improved, leading to a more visually appealing and natural-looking image. By contrast, the TMQI optimized image suffers from excessive contrast between the lights and the bricks on the wall. This problem is likely rooted in its statistical naturalness term, which drags μ and σ of all tone mapped images towards 116 and 65 regardless of their contents and luminance levels [19]. This is inappropriate for a night scene like “Clock building” which desires lower μ_e and σ_e (In TMQI-II, $\mu_e = 101$ and $\sigma_e = 48$). Moreover, annoying noise appears in the sky region of the TMQI optimized image.

TABLE II
TMQI-II COMPARISON BETWEEN INITIAL AND CONVERGED IMAGES

Image		Gamma	Log-normal	Reinhard [3]	Drago [4]	Durand [32]	Mantiuk [33]
BigFogMap	initial image	0.2148	0.3780	0.9333	0.7539	0.8881	0.9116
	converged image	0.9869	0.9906	0.9916	0.9891	0.9750	0.9860
ClockBuilding	initial image	0.3633	0.4582	0.9590	0.8949	0.8507	0.9592
	converged image	0.9874	0.9882	0.9898	0.9932	0.9895	0.9893
CornellBox	initial image	0.0686	0.1256	0.9570	0.6697	0.3332	0.9284
	converged image	0.9932	0.9892	0.9936	0.9904	0.8830	0.9866
DaniCathedral	initial image	0.2555	0.4205	0.9526	0.7803	0.8434	0.4949
	converged image	0.9836	0.9837	0.9773	0.9652	0.9735	0.9842
Desk1	initial image	0.1077	0.3860	0.9410	0.7713	0.7410	0.8554
	converged image	0.9878	0.9866	0.9918	0.9890	0.9883	0.9903
Desk2	initial image	0.1032	0.3968	0.9434	0.8300	0.5452	0.8209
	converged image	0.9898	0.9882	0.9933	0.994	0.9894	0.9933
Display1000	initial image	0.3202	0.4074	0.9356	0.5040	0.9109	0.6025
	converged image	0.9607	0.9635	0.9674	0.9940	0.9693	0.9707
House	initial image	0.3059	0.4392	0.9562	0.8705	0.9474	0.8505
	converged image	0.9899	0.9900	0.9916	0.9887	0.9900	0.9910
Kitchen	initial image	0.2661	0.3702	0.9428	0.4803	0.5532	0.9140
	converged image	0.9930	0.9897	0.9952	0.9950	0.9907	0.9954
Lawn	initial image	0.5955	0.4585	0.9642	0.6769	0.9196	0.9689
	converged image	0.9934	0.9944	0.9939	0.9921	0.9888	0.9912
Memorial	initial image	0.1048	0.4444	0.9727	0.8773	0.4939	0.9109
	converged image	0.9864	0.9881	0.9892	0.9942	0.9875	0.9876
Office	initial image	0.4337	0.4607	0.9549	0.5824	0.4494	0.9623
	converged image	0.9942	0.9950	0.9947	0.9924	0.9703	0.993
VineSunset	initial image	0.4486	0.4030	0.9261	0.4393	0.2898	0.8286
	converged image	0.9961	0.9921	0.9958	0.9855	0.9303	0.9835
Woman	initial image	0.3650	0.4306	0.9318	0.7724	0.9019	0.5830
	converged image	0.9930	0.9921	0.9944	0.9935	0.9931	0.9955
Wreathbu	initial image	0.4190	0.4468	0.9541	0.6861	0.9461	0.8470
	converged image	0.9679	0.9684	0.9702	0.9694	0.9702	0.9706



Fig. 6. Tone mapped “Woman” images. (a) Initial image created by Gamma mapping. (b) TMQI optimized image. (c) TMQI-II optimized image.

Fig. 8 shows the comparison of TMQI and TMQI-II optimization results on the “Bristol bridge” image with initial image created by Reinhard’s method [3]. Although the initial image of Fig. 8(a) has a seemingly reasonable visual appearance, the fine details of the woods and the brick textures of the tower are fuzzy or invisible. In Fig. 8(c), the proposed iterative algorithm using TMQI-II recovers these fine

details and makes them much sharper. Moreover, the overall appearance is softer and thus more pleasant. In Fig. 8(b), we can see that the iterative algorithm using TMQI heavily boosts noise in the sky and cloud regions, which leads to quality degradation when compared with the initial image. This also reveals the problem of TMQI in quality assessment of tone mapped images.

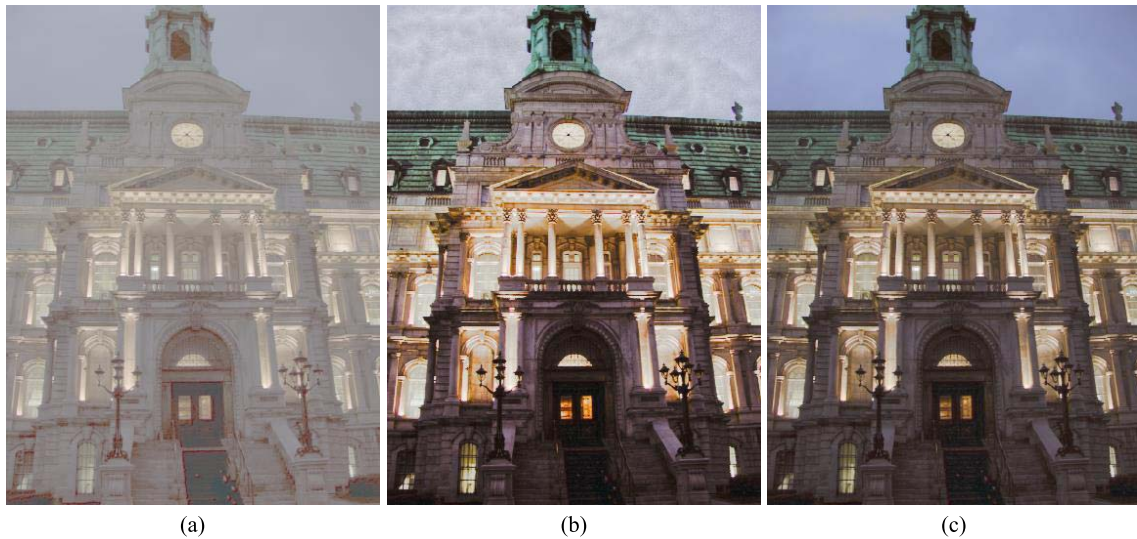


Fig. 7. Tone mapped “Clock building” images. (a) Initial image created by log-normal mapping. (b) TMQI optimized image. (c) TMQI-II optimized image.

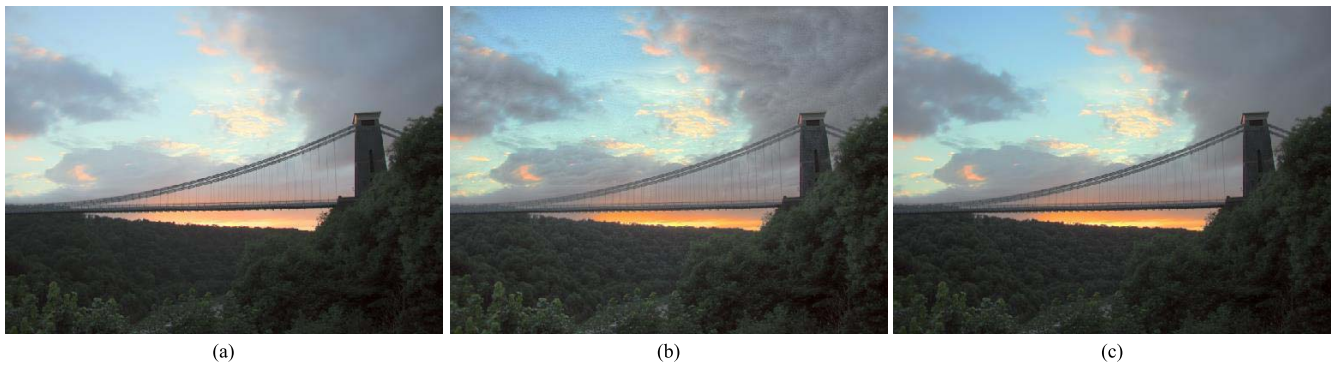


Fig. 8. Tone mapped “Bristol bridge” images. (a) Initial image created by Reinhard’s method [3]. (b) TMQI optimized image. (c) TMQI-II optimized image.



Fig. 9. Tone mapped “Grove” images. (a) Initial image created by Drago’s method [4]. (b) TMQI optimized image. (c) TMQI-II optimized image.

Fig. 9 shows the comparison of TMQI with TMQI-II on the “Grove” image with initial image created by Drago’s method [4]. Again, in the TMQI-II optimized image of Fig. 9(c), fine details such as leaves between the two big trees and the tree barks are well recovered and sharpened. The overall appearance is also more vivid. However, in Fig. 9(b), the iterative algorithm using TMQI over stretches the global contrast, which darkens the tree trunks and whitens the leaves and the sky. This affects the naturalness of the initial image and leads to quality degradation.

To further verify the effectiveness and consistency of the proposed algorithm, we conducted another subjective experiment. In particular, we select 15 HDR images that contain various natural scenes shown in Fig. 10 and adopt Gamma mapping, log-normal mapping and Reinhard’s method [3] to tone-map them to $15 \times 3 = 45$ LDR images. We then use them as initial images of the iterative algorithm and obtain 45 TMQI optimized images and 45 TMQI-II optimized images, respectively. Eventually, we obtain 15 sets of tone mapped images, each of which contains 9 images. 24 naive subjects

TABLE III
MEAN OPINION SCORES OF TONE MAPPED IMAGES

Image set	Mean opinion scores (MOS)								
	Gamma mapping			log-normal mapping			Reinhard's method [3]		
	Initial	TMQI optimized	TMQI-II optimized	Initial	TMQI optimized	TMQI-II optimized	Initial	TMQI optimized	TMQI-II optimized
1	1.00	4.71	7.46	5.58	4.13	7.29	7.29	4.71	7.92
2	1.54	4.92	8.25	2.13	4.33	7.83	8.17	5.25	8.38
3	0.25	4.50	6.71	1.88	3.79	6.25	8.25	4.83	8.33
4	3.33	4.63	6.54	3.33	4.96	7.17	6.96	5.63	7.79
5	0.54	3.88	7.67	4.58	4.25	7.13	7.79	4.46	8.58
6	0.58	3.29	7.38	5.21	3.17	6.58	7.92	3.33	7.50
7	1.29	5.67	7.38	2.38	5.33	7.17	8.00	5.67	8.33
8	1.54	3.79	6.75	3.29	3.83	6.54	8.00	4.08	8.46
9	0.96	4.38	7.25	2.67	3.79	7.29	8.21	4.33	8.50
10	6.50	5.21	6.42	2.54	5.13	7.75	8.04	6.50	8.50
11	0.46	4.83	7.13	4.00	2.79	5.50	7.83	3.63	8.13
12	3.63	5.21	7.17	3.46	3.00	7.08	7.38	3.79	7.63
13	5.33	3.04	6.00	2.67	3.29	6.83	7.79	3.04	8.04
14	2.17	2.75	6.29	3.58	2.83	5.75	8.47	3.04	7.67
15	4.67	5.13	5.83	1.92	3.88	6.17	6.33	6.08	6.79
Average	2.25	4.39	6.95	3.28	3.90	6.82	7.76	4.56	8.04

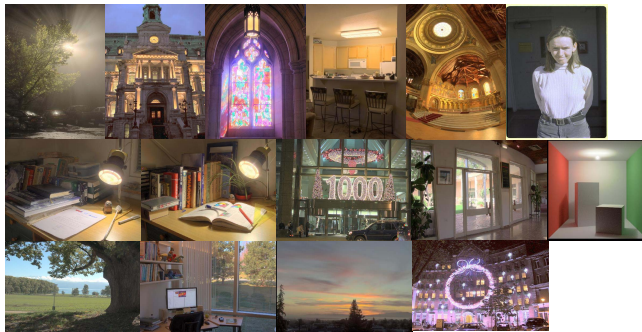


Fig. 10. HDR test images used in the subjective experiment. The images shown here are TMQI-II optimized with initial images created by Reinhard's method [3].

(9 males and 15 females aged between 22 and 30) were asked to give an integer score between 0 and 10 for the perceptual quality of each tone mapped image, where 0 denotes the worst quality and 10 the best. The final quality score for each individual image is computed as the average of subjective scores, named mean opinion score (MOS), from all subjects. The results are listed in Table III, from which we have several interesting observations. First, using TMQI-II as the optimization goal, the proposed algorithm leads to consistently perceptual gain for all three different types of initial images. By contrast, the perceptual gain obtained by optimizing TMQI is much less, when initial images are created by Gamma and log-normal mapping. Indeed, the quality of TMQI optimized images decreases when the initial images are created by Reinhard's method [3]. Second, the best quality image on average is TMQI-II optimized with the initial image created by Reinhard's method [3]. Note that because the image space is extremely complicated and the proposed algorithm can only guarantee to find a local optimum; thus better initial images often lead to better local optima, which correspondingly have better perceptual quality.

We use a hypothesis testing approach (based on t-statistics [34]) to evaluate the statistical significance of the

TABLE IV

STATISTICAL SIGNIFICANCE MATRIX BASED ON THE HYPOTHESIS TESTING. "1" MEANS THAT THE ROW CATEGORY IS STATISTICALLY BETTER THAN THE COLUMN CATEGORY. "0" MEANS THAT THE COLUMN CATEGORY IS STATISTICALLY BETTER THAN THE ROW CATEGORY. "-" MEANS THAT THE ROW AND COLUMN CATEGORIES ARE STATISTICALLY INDISTINGUISHABLE. G: GAMMA MAPPING INITIALIZED; L: LOG-NORMAL MAPPING INITIALIZED; R: REINHARD'S METHOD [3] INITIALIZED; (I): TMQI OPTIMIZED AND (II): TMQI-II OPTIMIZED

	G	G(I)	G(II)	L	L(I)	L(II)	R	R(I)	R(II)
G	-	0	0	0	0	0	0	0	0
G(I)	1	-	0	1	1	0	0	-	0
G(II)	1	1	-	1	1	-	0	1	0
L	1	0	0	-	0	0	0	0	0
L(I)	1	0	0	1	-	0	0	0	0
L(II)	1	1	-	1	1	-	0	1	0
R	1	1	1	1	1	1	-	1	0
R(I)	1	-	0	1	1	0	0	-	0
R(II)	1	1	1	1	1	1	1	1	-

subjective experimental results. Specifically, we treat the MOS values of each column in Table III as one category. The null hypothesis is that the MOS values in one category is statistically indistinguishable (with 95% confidence) from those in another category. The test is carried out for all possible combinations of pairs of categories, and the results are summarised in Table IV, from which we can see that TMQI-II optimized images have statistically significantly better MOS values in all cases.

In summary, we believe that the proposed iterative optimization procedure provides a strong test that not only verifies the superiority of TMQI-II over TMQI in predicting the perceptual quality of tone mapped images but also shows that the robustness and usefulness of TMQI-II to guide the optimization process with a variety of initial images.

Because of the complexity of the initial TMOs, TMQI-II, and the dimension of the search space, analytical convergence

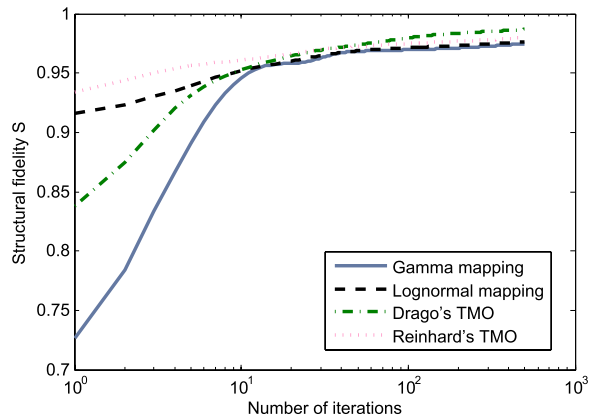


Fig. 11. Structural fidelity as a function of iteration with initial “woods” images created by different TMOs.

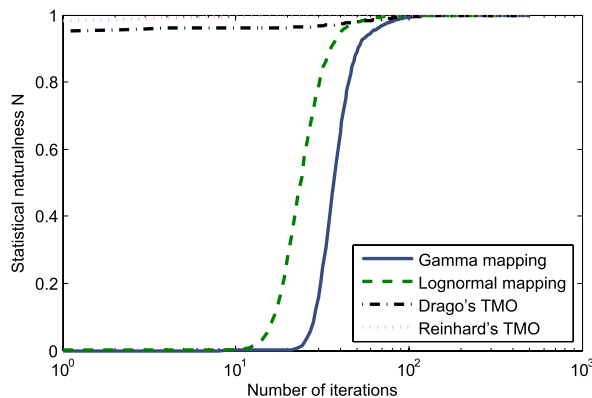


Fig. 12. Statistical naturalness as a function of iteration with initial “woods” images created by different TMOs.

assessment of the proposed algorithm is difficult. Therefore, we observe the convergence performance empirically. Figs. 11 and 12 show the structural fidelity and statistical naturalness measures as functions of iteration using different initial images as the starting points. There are several useful observations. First, both measures increase monotonically with iterations. Second, the proposed algorithm converges in all cases regardless of using simple or sophisticated TMO results as initial images. Third, different initial images may result in different converged images. From these observations, we conclude that the proposed iterative algorithm is well behaved, but the high-dimensional search space is complex and contains many local optima, and the proposed algorithm may be trapped in one of them.

The computation complexity of the proposed algorithm increases linearly with the number of pixels in the image. Our unoptimized MATLAB implementation takes around 4 seconds per iteration for a 341×512 image on a computer with an Intel Quad-Core 2.67 GHz processor.

IV. CONCLUSION AND FUTURE WORKS

We propose a novel approach to design TMOs by navigating in the space of images to find the optimal image in terms of an improved TMQI or TMQI-II. TMQI-II overcomes the

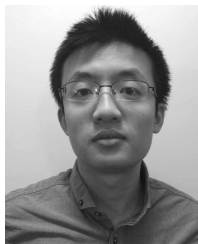
limitations underlying the structural fidelity and statistical naturalness components in TMQI and thus has better correlation with subjective quality evaluation. Optimizing TMQI-II is based on an iterative approach that alternates between improving the structural fidelity preservation and enhancing the statistical naturalness of the image. Numerical and subjective experimental results show that both steps contribute significantly to the improvement of the overall quality of the tone mapped image. The proposed algorithm further verifies the superiority of TMQI-II over TMQI. Finally, our experiments show that the proposed method is well behaved, and effectively enhances image quality from a wide variety of initial images, including those created from state-of-the-art TMOs.

The current work opens the door to a new class of tone mapping approaches. Many topics are worth further investigations. First, as is the case for any algorithm operating in complex high-dimensional spaces, the current approach only finds local optima. Deeper understanding of the search space is desirable to better solve the optimization problem. Second, the current implementation is computationally costly and requires a large number of iterations to converge. Fast search algorithms are necessary to accelerate the iterations. Third, objective quality assessment of tone mapped images still has much space to improve. In the future, better objective IQA models may be incorporated into the proposed framework to create better quality tone mapped images.

REFERENCES

- [1] B. A. Wandell, *Foundations of Vision*. Sunderland, MA, USA: Sinauer Associates, 1995.
- [2] E. Reinhard, W. Heidrich, P. Debevec, S. Pattanaik, G. Ward, and K. Myszkowski, *High Dynamic Range Imaging: Acquisition, Display, and Image-Based Lighting*. San Mateo, CA, USA: Morgan Kaufmann, 2010.
- [3] E. Reinhard, M. Stark, P. Shirley, and J. Ferwerda, “Photographic tone reproduction for digital images,” *ACM Trans. Graph.*, vol. 21, no. 3, pp. 267–276, 2002.
- [4] F. Drago, K. Myszkowski, T. Annen, and N. Chiba, “Adaptive logarithmic mapping for displaying high contrast scenes,” *Comput. Graph. Forum*, vol. 22, no. 3, pp. 419–426, 2003.
- [5] Q. Shan, J. Jia, and M. S. Brown, “Globally optimized linear windowed tone mapping,” *IEEE Trans. Vis. Comput. Graphics*, vol. 16, no. 4, pp. 663–675, Jul./Aug. 2010.
- [6] S. Ferradans, M. Bertalmio, E. Provenzi, and V. Caselles, “An analysis of visual adaptation and contrast perception for tone mapping,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 33, no. 10, pp. 2002–2012, Oct. 2011.
- [7] B. Gu, W. Li, M. Zhu, and M. Wang, “Local edge-preserving multiscale decomposition for high dynamic range image tone mapping,” *IEEE Trans. Image Process.*, vol. 22, no. 1, pp. 70–79, Jan. 2013.
- [8] H. Yeganeh and Z. Wang, “High dynamic range image tone mapping by maximizing a structural fidelity measure,” in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process.*, May 2013, pp. 1879–1883.
- [9] K. Ma, H. Yeganeh, K. Zeng, and Z. Wang, “High dynamic range image tone mapping by optimizing tone mapped image quality index,” in *Proc. IEEE Int. Conf. Multimedia Expo*, Jul. 2014, pp. 1–6.
- [10] F. Drago, W. L. Martens, K. Myszkowski, and H.-P. Seidel, “Perceptual evaluation of tone mapping operators,” in *Proc. ACM SIGGRAPH Sketches Appl.*, 2003, p. 1.
- [11] M. Čadík and P. Slavík, “The naturalness of reproduced high dynamic range images,” in *Proc. 9th Int. Conf. Inf. Vis.*, 2005, pp. 920–925.
- [12] M. Čadík, M. Wimmer, L. Neumann, and A. Artusi, “Evaluation of HDR tone mapping methods using essential perceptual attributes,” *Comput. Graph.*, vol. 32, no. 3, pp. 330–349, 2008.
- [13] C. Cavaro-Ménard, L. Zhang, and P. Le Callet, “Diagnostic quality assessment of medical images: Challenges and trends,” in *Proc. IEEE 2nd Eur. Workshop Vis. Inf. Process.*, Jul. 2010, pp. 277–284.

- [14] Z. Wang and A. C. Bovik, "Mean squared error: Love it or leave it? A new look at signal fidelity measures," *IEEE Signal Process. Mag.*, vol. 26, no. 1, pp. 98–117, Jan. 2009.
- [15] 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.
- [16] Z. Wang and A. C. Bovik, "A universal image quality index," *IEEE Signal Process. Lett.*, vol. 9, no. 3, pp. 81–84, Mar. 2002.
- [17] R. Mantiuk, S. J. Daly, K. Myszkowski, and H.-P. Seidel, "Predicting visible differences in high dynamic range images: Model and its calibration," *Proc. SPIE*, vol. 5666, pp. 204–214, Mar. 2005.
- [18] T. O. Aydin, R. Mantiuk, K. Myszkowski, and H.-P. Seidel, "Dynamic range independent image quality assessment," *ACM Trans. Graph.*, vol. 27, no. 3, 2008, Art. ID 69.
- [19] H. Yeganeh and Z. Wang, "Objective quality assessment of tone-mapped images," *IEEE Trans. Image Process.*, vol. 22, no. 2, pp. 657–667, Feb. 2013.
- [20] W. J. Crozier, "On the variability of critical illumination for flicker fusion and intensity discrimination," *J. General Physiol.*, vol. 19, no. 3, pp. 503–522, 1936.
- [21] P. G. J. Barten, *Contrast Sensitivity of the Human Eye and Its Effects on Image Quality*. Bellingham, WA, USA: SPIE, 1999.
- [22] V. Mante, R. A. Frazor, V. Bonin, W. S. Geisler, and M. Carandini, "Independence of luminance and contrast in natural scenes and in the early visual system," *Nature Neurosci.*, vol. 8, no. 12, pp. 1690–1697, 2005.
- [23] Z. Wang and E. P. Simoncelli, "Stimulus synthesis for efficient evaluation and refinement of perceptual image quality metrics," *Proc. SPIE*, vol. 5292, pp. 99–108, Jun. 2004.
- [24] Z. Wang and E. P. Simoncelli, "Maximum differentiation (MAD) competition: A methodology for comparing computational models of perceptual quantities," *J. Vis.*, vol. 8, no. 12, p. 8, 2008.
- [25] G. Ward, "A contrast-based scalefactor for luminance display," in *Graphics Gems IV*, P. Heckbert, Ed. San Diego, CA, USA: Academic, 1994, pp. 415–421.
- [26] J. Holm, "Photographic tone and color reproduction goals," in *Proc. CIE Expert Symp., Color Standards Image Technol.*, 1996, pp. 51–56.
- [27] N. Salvaggio, L. Stroebel, and R. Zakia, *Basic Photographic Materials and Processes*. New York, NY, USA: Taylor & Francis, 2008.
- [28] Y. Li, L. Sharan, and E. H. Adelson, "Compressing and companding high dynamic range images with subband architectures," *ACM Trans. Graph.*, vol. 24, no. 3, pp. 836–844, 2005.
- [29] H. R. Sheikh, Z. Wang, L. Cormack, and A. C. Bovik. *Live Image Quality Assessment Database*. [Online]. Available: <http://live.ece.utexas.edu/research/quality>, accessed Nov. 2006.
- [30] S. P. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2004.
- [31] J. Nocedal and S. J. Wright, *Numerical Optimization*, 2nd ed. New York, NY, USA: Springer-Verlag, 2006.
- [32] F. Durand and J. Dorsey, "Fast bilateral filtering for the display of high-dynamic-range images," *ACM Trans. Graph.*, vol. 21, no. 3, pp. 257–266, 2002.
- [33] R. Mantiuk, S. Daly, and L. Kerofsky, "Display adaptive tone mapping," *ACM Trans. Graph.*, vol. 27, no. 3, 2008, Art. ID 68.
- [34] D. C. Montgomery and G. C. Runger, *Applied Statistics and Probability for Engineers*, 6th ed. New York, NY, USA: Wiley, 2013.



Kede Ma (S'13) received the B.E. degree from the University of Science and Technology of China, Hefei, China, in 2012, and the M.A.Sc. degree from the University of Waterloo, Waterloo, ON, Canada, where he is currently pursuing the Ph.D. degree in electrical and computer engineering. His research interest lies in perceptual image processing.



Hojatollah Yeganeh (S'10–M'14) received the B.S. degree in electrical engineering from the Isfahan University of Technology, Isfahan, Iran, in 2006, the M.Sc. degree in electrical engineering with a specialization in speech processing and speech recognition from the Amirkabir University of Technology, Tehran, Iran, in 2009, and the Ph.D. degree in electrical engineering from the University of Waterloo, ON, Canada, in 2014. He is currently a Post-Doctoral Fellow with the Department of Electrical and Computer Engineering, University of Waterloo. His main research interests include image and video processing, quality assessment of high dynamic range images and videos, quality of experience of Internet videos, multimedia communications, and biomedical signal processing.



Kai Zeng received the B.E. and M.A.Sc. degrees in electrical engineering from Xidian University, Xi'an, China, in 2009, and the Ph.D. degree in electrical and computer engineering from the University of Waterloo, Waterloo, ON, Canada, in 2013. He is currently a Post-Doctoral Fellow with the Department of Electrical and Computer Engineering, University of Waterloo. His research interests include computational video and image pattern analysis, multimedia communications, and image and video processing (coding, denoising, analysis, and representation), with an emphasis on image and video quality assessment and corresponding applications. He was a recipient of the IEEE Signal Processing Society Student Travel Grant at the 2010 and 2012 IEEE International Conference on Image Processing, and the prestigious 2013 Chinese Government Award for Outstanding Students Abroad.



Zhou Wang (S'99–M'02–SM'12–F'14) received the Ph.D. degree in electrical and computer engineering from The University of Texas at Austin, in 2001. He is currently a Professor with the Department of Electrical and Computer Engineering, University of Waterloo, Canada. His research interests include image processing, coding, and quality assessment; computational vision and pattern analysis; multimedia communications; and biomedical signal processing. He has over 100 publications in his research fields with over 24 000 citations (a Google Scholar).

He was a member of the IEEE Multimedia Signal Processing Technical Committee (2013–2015). He served as an Associate Editor of the IEEE TRANSACTIONS ON IMAGE PROCESSING (2009–2014), *Pattern Recognition* (2006–present), and the IEEE SIGNAL PROCESSING LETTERS (2006–2010), and a Guest Editor of the IEEE JOURNAL OF SELECTED TOPICS IN SIGNAL PROCESSING (2013–2014 and 2007–2009), the *EURASIP Journal of Image and Video Processing* (2009–2010), and *Signal, Image and Video Processing* (2011–2013). He was a recipient of the 2014 NSERC E.W.R. Steacie Memorial Fellowship Award, the 2013 IEEE Signal Processing Best Magazine Paper Award, the 2009 IEEE Signal Processing Society Best Paper Award, the 2009 Ontario Early Researcher Award, and the ICIP 2008 IBM Best Student Paper Award (as a senior author).