Objective Quality Assessment and Optimization for High Dynamic Range Image Tone Mapping

by

Kede Ma

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Master of Applied Science
in
Electrical and Computer Engineering

Waterloo, Ontario, Canada, 2014

Author's Declaration

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

Tone mapping operators aim to compress high dynamic range (HDR) images to low dynamic range ones so as to visualize HDR images on standard displays. Most existing works were demonstrated on specific examples without being thoroughly tested on well-established and subject-validated image quality assessment models. A recent tone mapped image quality index (TMQI) made the first attempt on objective quality assessment of tone mapped images. TMQI consists of two fundamental building blocks: structural fidelity and statistical naturalness. In this thesis, we propose an enhanced tone mapped image quality index (eTMQI) by 1) constructing an improved nonlinear mapping function to better account for the local contrast visibility of HDR images and 2) developing an image dependent statistical naturalness model to quantify the unnaturalness of tone mapped images based on a subjective study. Experiments show that the modified structural fidelity and statistical naturalness terms in eTMQI better correlate with subjective quality evaluations. Furthermore, we propose an iterative optimization algorithm for tone mapping. The advantages of this algorithm are twofold: 1) eTMQI and TMQI can be compared in a more straightforward way; 2) better quality tone mapped images can be automatically generated by using eTMQI as the optimization goal. Numerical and subjective experiments demonstrate that eTMQI is a superior objective quality assessment metric for tone mapped images and consistently outperforms TMQI.

Acknowledgements

I would like to thank all the people who made this thesis possible.

First of all, my supervisor Professor Zhou Wang. It is Dr. Wang who guided me to the field of image quality assessment and numerical optimization. Every time I discussed with him, I learned something new. He not only taught me how to think big but also gave me very detailed suggestions on my research topics. More importantly, he is a really nice and modest person. He seldom criticized me when I progressed slowly or faced a dilemma. What he always does is patient enough to mentor me step by step. It has been my great pleasure to work with Dr. Wang. And I am looking forward to my Ph.D. studies under his guidance.

Second, my fellow Hojatollah Yeganeh. Most of work in this thesis is built up Hojatollah's previous work. He shared selflessly his thoughts and materials with me so that I could be on the same page in the first place. It has been a wonderful experience to cooperate with him.

Third, I am very grateful to Professor Dana Kulić and Professor Liangliang Xie for being the examiners of my thesis and providing constructive comments. Truly thank Professor Dana Kulić for her precious time to attend my seminar.

Moreover, I would like to express my sincere gratitude to all amazing labmates in the image and visual computing group. Particularly, I would like to thank Kai Zeng, Tiesong Zhao, Jiheng Wang, Yuming Fang and Abdul Rehman for their help and support during my Master studies.

Last but not least, I would like to thank my parents, Qingquan Ma and Wei Wu, for giving me privilege to pursue my dream with full supports. I also wish to thank my girlfriend who had not yet appeared during the past two years of my life so that I could concentrate on my research.

Table of Contents

List of Tables

List of Figures

List of Acronyms

CDF cumulative distribution functions. 27

eTMQI enhanced tone mapped image quality index. v, 4, 5, 21, 23, 28–31, 33, 37, 39, 40, 42, 44, 45, 48, 49, 51, 52

HDR high dynamic range. v, 1, 3, 4, 7, 10, 12, 15, 17, 22–25, 27, 29, 31, 36, 39, 40, 42, 44, 51

HVS human visual system. 12, 13, 15, 23

IQA image quality assessment. 3, 4, 7, 12, 20, 29, 31, 49, 51, 52

KRCC Kendall's rank-order correlation coefficient. 29

LDR low dynamic range. 1, 3, 5, 7, 10, 12, 15, 17, 25, 26, 31, 42, 51

MOS mean opinion score. 44

MSE Mean square error. 13, 15

SRCC Spearman's rank-order correlation coefficient. 29

SSIM structural similarity. 13, 15, 17

TMO Tone mapping operators. 1, 3, 5, 7, 10, 12, 24, 31, 37, 39, 42, 48, 49, 52

TMQI tone mapped image quality index. v, 3–5, 7, 15, 17, 18, 20–23, 28–31, 37, 39, 40, 42, 44, 45, 51

Chapter 1

Introduction

1.1 Motivation

The luminance of a natural scene often has a high dynamic range (HDR), varying between 10^{-3} to 10^{5} cd/m^2. However, a normal digital display only has a low dynamic range (LDR) about 10^{2} cd/m^2 [1]. Tone mapping operators (TMO) 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 nice surrogate for HDR display technology, which is currently immature and expensive. Fig. 1.1 gives some examples, where the same scene is shot multiple times with slightly different exposure settings, which may be subsequently merged to an HDR image. The left image shows the best exposure shots that are chosen from corresponding successive shots. Due to the existence of both light and dark areas in the same scene, even the best exposure shots fail to capture the detail and color appearance of the sky in the background and the bricks in the foreground. By contrast, the exposure of both the indoor and the outdoor areas has been greatly improved in tone mapped images. On the other hand, printed media are typically not HDR. Regardless of how fast HDR display technologies penetrate, 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

Figure 1.1: Optimally exposed images (left) versus HDR tone mapped images (right). (Image courtesy of Reinhard's book [2])

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 tested on well-established and subject-validated image quality assessment (IQA) models. Subjective evaluation is a straightforward and useful quality measure [10–13], because the human visual system is the ultimate receiver in most applications. However it is expensive, time consuming, and perhaps most importantly, can hardly guide automatic optimization algorithms [14]. In the context of subjective quality assessment of tone mapped images, subjects usually directly compare several different tone mapped images without referencing the corresponding HDR images. Therefore, they may be unaware of certain structure detail loss in tone mapped images that may convey important information in HDR images. For this reason, subjective evaluation may not be a golden evaluation criterion. Objective quality assessment of tone mapped images is a challenging problem due to the different dynamic ranges between the reference HDR image and the tone mapped LDR image. Traditional objective IQA metrics such as peak signal-to-noise ratio and the structural similarity index [14, 15] 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 [16] tries to predict the visible difference between two HDR images with the same range. A dynamic range independent quality measure [17] focuses on detecting the loss of visible contrast, amplification of invisible contrast and reversal of visible contrast, and produces three corresponding quality maps. However it does not integrate these three components into an overall quality score. Recently, a tone mapped image quality index (TMQI) was proposed in [18], 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 all the structural detail in the HDR image but also look natural. During structural fidelity computation in TMQI [18], one single nonlinear mapping is used to measure the local contrast visibility in both HDR images and tone mapped images. Because the dynamic range of the HDR images is much higher than that of the tone mapped images, the soft thresholds of contrast visibility of HDR images and tone mapped images should be different and

adapted to local luminance levels. More specifically, if the nonlinear mapping in TMQI correctly captures the contrast visibility of tone mapped images, it inevitably treats all local regions in HDR images as contrast visible. In this case, the structural fidelity term will always penalize homogenous regions in tone mapped images, which may cause major problems in the subsequent optimization process. Another drawback of TMQI lies in its image independent statistical naturalness model, which suggests that all statistical natural tone mapped images have the same overall luminance and global contrast. This oversimplifies the reality where different images should have different natural overall luminance and global contrast based on their image contents and lighting conditions. For example, an outdoor scene at noon time may have much higher overall luminance and lower global contrast while an indoor scene containing different objects is exactly the opposite.

1.2 Objectives

The objectives of this thesis are to develop advanced IQA models to accurately predict the perceptual quality of tone mapped images and also to improve their quality based on such IQA models.

1.3 Contributions

The major contributions of this thesis are summarized as follows.

- We propose an enhanced tone mapped image quality index (eTMQI) based on the existing TMQI. We first construct an improved nonlinear mapping function particularly for HDR images to better distinguish the visible and invisible contrast regions. Instead of using standard deviation, we choose standard deviation divided by the mean as an estimate of contrast in HDR images. This estimate adapts to local luminance levels and is thus qualitatively consistent with Weber's law. We then build an image dependent statistical naturalness model and quantify the unnaturalness of

LDR images based on a subjective study. Experiments show that the structural fidelity and statistical naturalness terms in eTMQI better correlate with the subjective data than those in TMQI.

- We propose an iterative optimization algorithm for tone mapping. In each iteration, we move the image towards the direction that optimizes eTMQI in the space of all images. We alternately improve the structural fidelity and statistical naturalness of the image until some stopping criteria are activated. The advantages of this algorithm are twofold: first, eTMQI and TMQI can be compared in a more straightforward way because the test images are sampled from the space of all images that the optimization algorithm navigates; second, better quality tone mapped images can be automatically generated if eTMQI is chosen as the optimization goal. Numerical and subjective experiments show that the iterative algorithm under the guidance of eTMQI faithfully creates better quality tone mapped images compared with those using TMQI as the optimization goal.

1.4 Thesis Outline

Following the introductory chapter, the remainder of this thesis is organized as follows. Chapter 2 reviews several TMOs, the structural similarity index and TMQI that prepare the necessary background knowledge for later chapters. Chapter 3 elaborates on the construction of eTMQI with an emphasis on the differences and improvement over TMQI. Validations from our subjective study indicate that the modified structural fidelity and statistical naturalness terms in eTMQI outperform those in TMQI. Chapter 4 details the iterative optimization algorithm using eTMQI and TMQI as optimization goals followed by extensive numerical and subjective experiments to verify the superiority of eTMQI over TMQI. Chapter 5 concludes the thesis and discusses future work.

Chapter 2

Background

This chapter starts with a literature review of tone mapping operators (TMO). Then it presents the well-known full-reference IQA model, the structural similarity index, followed by the construction of tone mapped image quality index (TMQI) whose fundamental ideas are inspired by the structural similarity index.

2.1 HDR-to-LDR Tone Mapping Operators

High dynamic range (HDR) images[1] represent higher precision of luminance levels in natural scenes than standard low dynamic range (LDR) images [3]. They can be directly acquired with HDR imaging devices [2] or created by fusing differently exposed images of the same real scene, as illustrated in Fig. 2.1. HDR-to-LDR TMOs [5–9] facilitate display of HDR images on viewing devices with lower dynamic range. The goal of TMOs is to compress the dynamic range of HDR images while preserving their structural detail and natural appearance. Tone mapping is a nontrivial problem since simple linear scaling often produces images with severe detail loss as exemplified in Fig. 2.2. Generally, TMOs

[1] The term "dynamic range" for images is defined as the ratio between the lightest and darkest pixels [19].

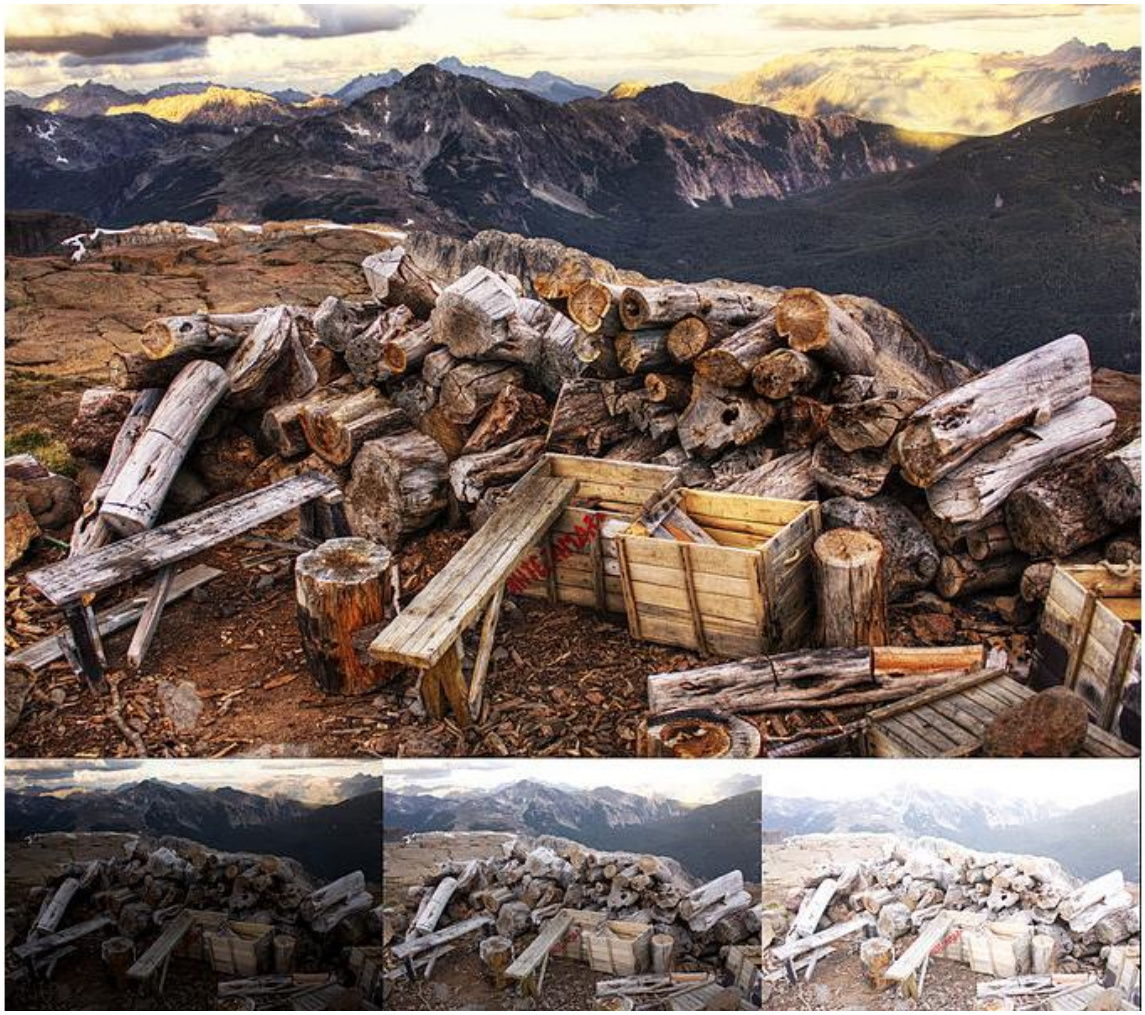


Figure 2.1: HDR image (top) made out of 3 LDR images (bottom) with different exposures. (Image courtesy of Wikipedia)

Figure 2.2: Linear scaling (shown on the top) fails to preserve the structural detail of a natural scene containing both indoors and outdoors areas. As a comparison, a state-of-the-art TMO [3] (shown on the bottom) does a better job. (Image courtesy of the MathWorks, Inc)

can be classified into global and local ones. Global TMOs are essentially point-wise luminance transformations. They are simple, fast and often produce spatially consistent LDR images. However, they often suffer from fine detail loss in local areas especially in HDR images that contain both light and dark areas. Sigmoid function and histogram-based algorithms [20–22] are two main categories of global operators. Typical sigmoid function-based algorithms involve Gamma mapping and log-normal mapping. Gamma mapping takes the following formula

$$\mathbf{Y} = \mathbf{X}^{\frac{1}{\gamma}}, \tag{2.1}$$

where $\mathbf{X}$ and $\mathbf{Y}$ represent HDR and tone mapped LDR images, respectively. The parameter γ is typically set to 2.2. Assuming the intensity values of LDR images lie in $[0, 1]$, $\mathbf{Y}$ should be further clamped to $[0, 1]$ for display. Log-normal mapping adopts the logarithmic function to boost lower luminance levels and to compress higher luminance levels. The tone mapped image is computed by

$$\mathbf{Y} = \frac{\log_2(\mathbf{X}) - x_{\min}}{x_{\max} - x_{\min}}, \tag{2.2}$$

where $x_{\min}$ and $x_{\max}$ are the minimum and maximum pixel values of $\log_2(\mathbf{X})$, respectively. Gamma mapping often creates dark images with missing detail similar to linear scaling while log-normal mapping often produces blanched images with fuzzy detail and unnatural appearance. A visual example is shown in Fig. 2.3. In 1993, Tumblin and Rushmeier [23] proposed a global TMO based on Stevens' law [24]. Larson et al. developed a histogram adjustment technique to perform tone mapping; they accounted for human contrast sensitivity, veiling luminance, color sensitivity and visual acuity all in the context of a local adaptation model [20]. Fattal et al. proposed a gradient-based TMO [25]. They first attenuated large gradients and then constructed an LDR image by solving a Poisson equation on the modified gradient field. In [5], Shan et al. proposed a globally nonlinear method with overlapping window-based linear functions to reconstruct the image radiance. Other well-known global TMOs are Reinhard's method [3] and Drago's method [4]. They are considered to be among the best TMOs on several independent subjective tests [12, 18].

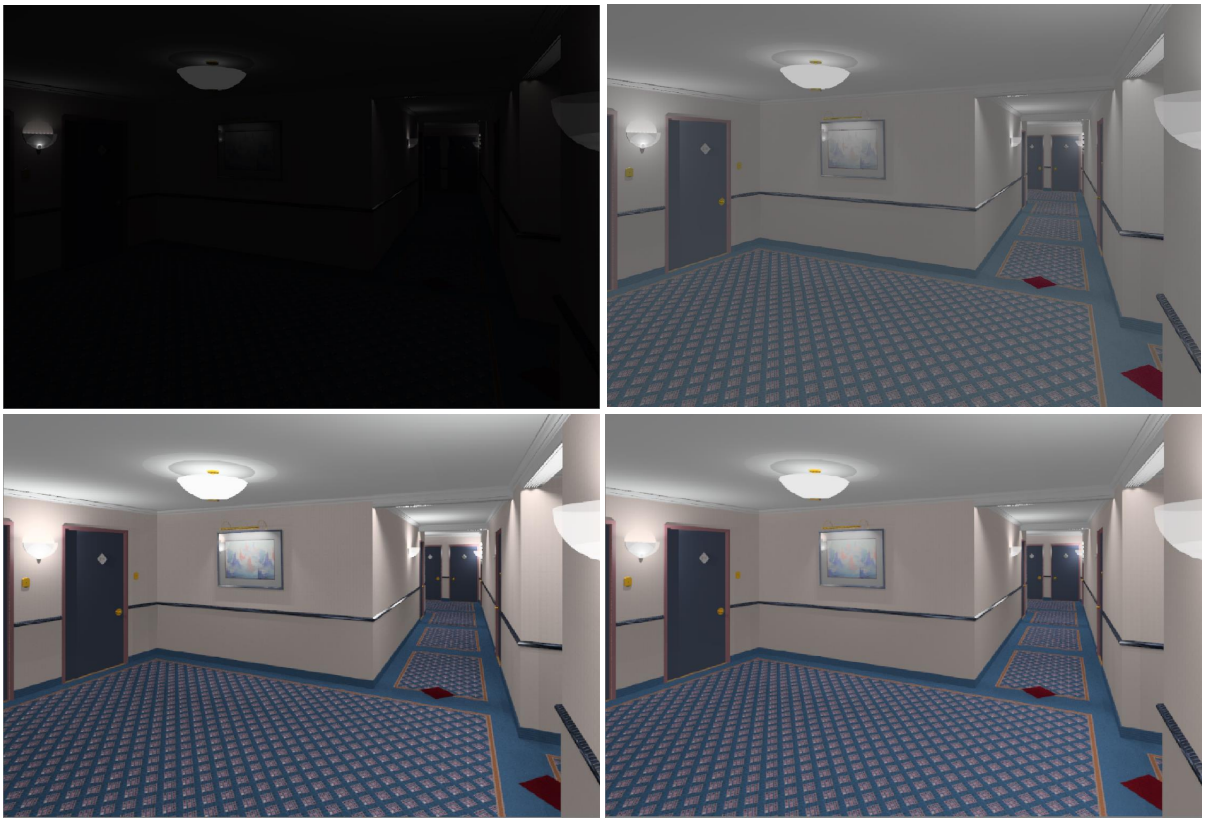


Figure 2.3: Gamma mapping (shown on the top left) fails to preserve the structural detail of a natural scene. Log-normal mapping (shown on the top right) produces extremely fuzzy detail which looks unnatural. Reinhard's method [3] (shown on the bottom left) and Drago's method [4] (shown on the bottom right) faithfully reproduce the visual appearance. (Image courtesy of Gred Ward)

In [3], Reinhard et al. first approximated the key of a scene[2] as the log average luminance and scaled the luminance according to the key. They further compressed high luminance by the function $f(x) = x/(1 + x)$. Finally this resulting tone mapped LDR image was modified locally to further enhance the detail. In [4], Drago et al. proposed an adaptive logarithmic mapping technique with the logarithmic bases varying according to a bias power function. The resulting image further undergoes a Gamma correction mapping to improve contrast in dark areas. A visual example is also shown in Fig. 2.3, where we observe that Reinhard's method [3] and Drago's method [4] faithfully reproduce the visual appearance without significant noticeable artifacts.

Local TMOs are primarily inspired by how the human visual system (HVS) adapts and responds to local scene luminance and contrast. In [26], the edges of an HDR image were first extracted using a bilateral filter [27]. Compression of the dynamic range was performed using linear scaling in the logarithm domain. The final LDR tone mapped image was obtained by adding back the edges. A similar local edge-preserving multiscale decomposition scheme for HDR image tone mapping was proposed in [7]. Chen et al. addressed the tone mapping problem by integrating the local adaptation effect with the consistency in global contrast impression [28]. They decomposed the luminance into a small number of regions that represent the overall impression of an HDR image and constructed a piecewise tone mapping to compress the dynamic range constrained by the estimated global perception. In [29], Li et al. adopted subband architectures for tone mapping with local gain control. Their algorithm can also be used for inverse tone mapping[3].

With so many existing TMOs at hand, a natural question should be asked: which TMO produces the best perceptual tone mapped LDR image? This question could be possibly answered by subjective evaluations. However, it is expensive and extremely time-consuming because whenever new TMOs come up, new subjective experiments need to be done to test their performance. Furthermore, we obtain little clue as to whether there is still room for further improvement. A promising way to solve this problem is to develop ob-

[2]A scene's key is an indicator of how light or dark the overall impression of a scene is [2].

[3]Inverse tone mapping is the inverse process of tone mapping. It upscales LDR images in a perceptually plausible manner to obtain HDR images [2].

jective IQA models that can automatically evaluate the performance of TMOs. Designing objective quality measures for TMOs is the main focus of this thesis.

2.2 The Structural Similarity Index

During their lifetime, digital images may undergo many transformations including acquisition, processing, compression, storage, transmission and reproduction. These transformations may introduce a variety of distortions to images and therefore result in changes in visual quality. The goal of image fidelity measures is to compare two images quantitatively with the assumption that one of the images has pristine quality. Mean square error (MSE) is the most widely used image fidelity measure because it is simple, easy to optimize and has a clear physical meaning [14]. However, it is also widely criticized for its poor correlation with human perception of image fidelity and quality [14, 30]. Fig. 2.4 shows an example [14], where an original "Einstein" image is contaminated by different types of distortion: a contrast stretch, mean luminance shift, contamination by additive white Gaussian noise, impulsive noise distortion, JPEG compression, blur, spatial scaling, spatial shift, and rotation. It can be observed that the MSE values of images (b)-(g) relative to the original image (a) are almost the same; however the visual quality of the images are dramatically different. In addition, small geometrical changes (images (h)-(i)) that barely tamper the perceptual quality of the images lead to very large MSE values. One surrogate of MSE that has been widely recognized is the structural similarity (SSIM) index. Despite different forms of SSIM—whether it is implemented at single scale [15, 31], over multiple scales (MS-SSIM) [32], in the complex wavelet domain [33] (CW-SSIM) or information weighted (IW-SSIM) [34]—it is based on the assumption that the HVS is highly adapted to extract structural information from the viewing field [35] and thus changes in structural information should be penalized for image quality degradation. Suppose $\mathbf{x}$ and $\mathbf{y}$ are two local image patches taken from the same location of two images, the local SSIM index computes three components: the luminance similarity $l(\mathbf{x}, \mathbf{y})$, contrast similarity $c(\mathbf{x}, \mathbf{y})$

Figure 2.4: Comparison of image fidelity measures for "Einstein" image altered with different types of distortions. (a) Reference image. (b) Mean contrast stretch. (c) Luminance shift. (d) Gaussian noise contamination. (e) Impulsive noise contamination. (f) JPEG compression. (g) Blurring. (h) Spatial scaling (zooming out). (i) Spatial shift (to the right). (j) Spatial shift (to the left). (k) Rotation (counter-clockwise). (l) Rotation (clockwise). (Example taken from Wang et al.'s paper [14])

and structure similarity $s(\mathbf{x}, \mathbf{y})$

$$l(\mathbf{x}, \mathbf{y}) = \frac{2\mu_x\mu_y + C_1}{\mu_x^2 + \mu_y^2 + C_1}\,, \tag{2.3}$$

$$c(\mathbf{x}, \mathbf{y}) = \frac{2\sigma_x\sigma_y + C_2}{\sigma_x^2 + \sigma_y^2 + C_2}\,, \tag{2.4}$$

$$s(\mathbf{x}, \mathbf{y}) = \frac{\sigma_{xy} + C_3}{2\sigma_x\sigma_y + C_3}\,. \tag{2.5}$$

μ, σ and σ_{xy} denote the mean, standard deviation (std) and covariance of the image patches, respectively. C_1, C_2 and C_3 are small positive constants to avoid instability. By multiplying these three components and setting $C_3 = C_2/2$ [15], we obtain a simplified version of the SSIM index that is typically used in practice

$$\mathrm{SSIM}(\mathbf{x}, \mathbf{y}) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)}\,. \tag{2.6}$$

The overall SSIM index of the image is computed by averaging all local SSIM indices using a sliding window. Revisiting the example in Fig. 2.4, it can be seen that the SSIM values are much more consistent with the HVS than the MSE values. Image enhancement operations such as luminance shifting and contrast stretching generally preserve image structure and thus lead to very high SSIM values, while noise contamination and JPEG-compression lead to low SSIM values [14]. Fig. 2.5 shows a stronger example, where maximum and minimum SSIM images were automatically generated on the same equal-MSE hypersphere using the method proposed in [36]. This example fails the MSE dramatically.

2.3 The Tone Mapped Image Quality Index

This section provides a brief background description of TMQI proposed in [18]. This is necessary in explaining the enhanced TMQI metric, which is one of the main contributions of this thesis. Detailed descriptions of TMQI can be found in [18].

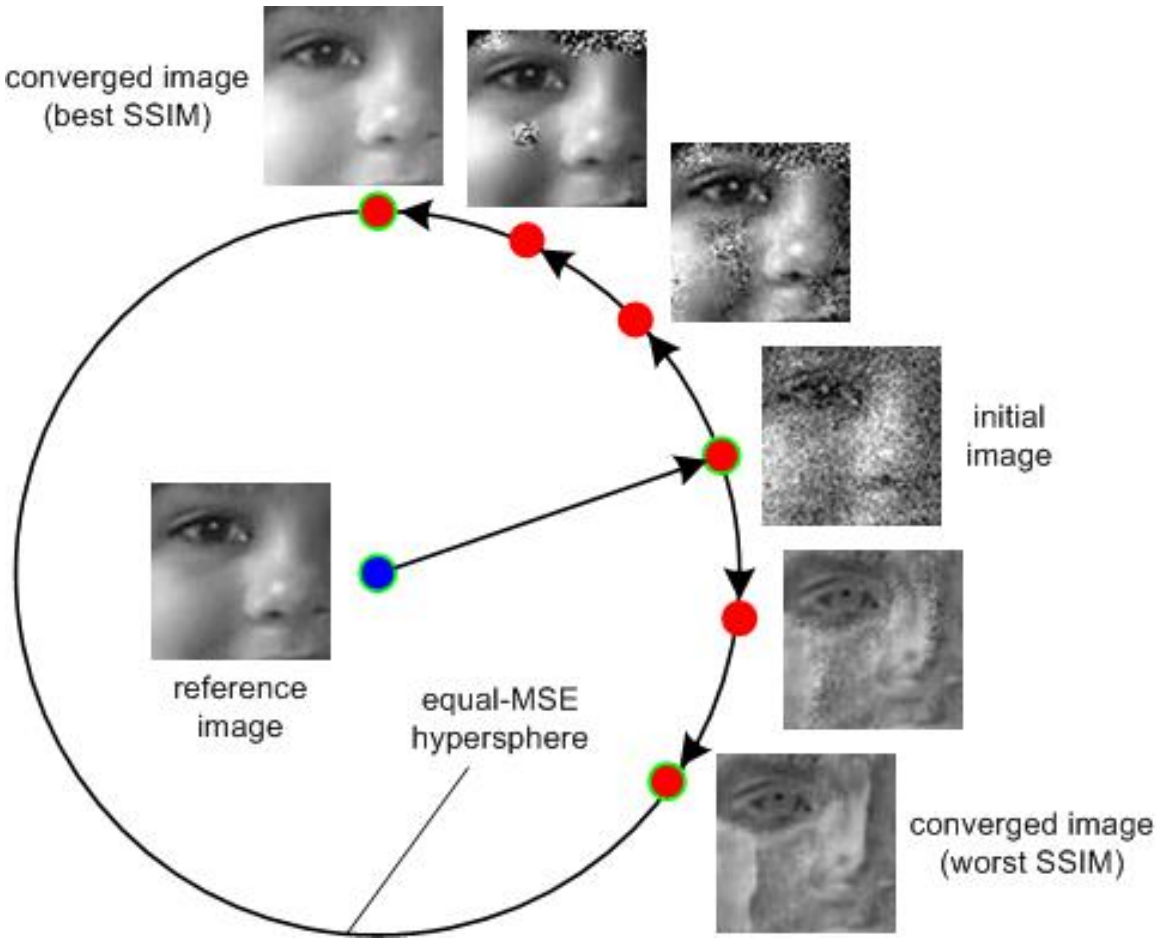


Figure 2.5: Finding the maximum/minimum SSIM images along the equal-MSE hypersphere in image space. (Example taken from Wang et al.'s paper [14])

Let $\mathbf{X}$ and $\mathbf{Y}$ be the HDR image and the tone mapped LDR image, respectively. The fundamental idea behind TMQI is that a good quality tone mapped image should achieve a good compromise between two key factors—structural fidelity and statistical naturalness. The TMQI computation is given by [18]

$$\mathrm{TMQI}(\mathbf{X}, \mathbf{Y}) = a[S(\mathbf{X}, \mathbf{Y})]^{\alpha} + (1-a)[N(\mathbf{Y})]^{\beta}, \tag{2.7}$$

where S and N denote the structural fidelity and statistical naturalness terms, respectively. The parameters α and β determine the sensitivities of the two factors, 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.

The computation of the structural fidelity S is patch-based. Let $\mathbf{x}$ and $\mathbf{y}$ be two image patches extracted from $\mathbf{X}$ and $\mathbf{Y}$, respectively. An SSIM-motivated local structural fidelity measure is defined as

$$S_{\mathrm{local}}(\mathbf{x}, \mathbf{y}) = \frac{2\tilde{\sigma}_x \tilde{\sigma}_y + C_1}{\tilde{\sigma}_x^2 + \tilde{\sigma}_y^2 + C_1} \cdot \frac{\sigma_{xy} + C_2}{\sigma_x \sigma_y + C_2}, \tag{2.8}$$

The first term is a modification of the local contrast comparison component in SSIM [15], and the second term is the same as the structure comparison component in SSIM [15]. The local contrast comparison term is based on two considerations. First, as along as the contrast in the HDR and LDR patches are both significant or both insignificant, the contrast differences should not be penalized. Second, the measure should penalize the cases in which the contrast is significant in one of the patches but not in the other. In TMQI [18], to assess the significance of local contrast, the local standard deviation σ is passed through a contrast sensitivity model-based nonlinear mapping function given by

$$\tilde{\sigma} = \frac{1}{\sqrt{2\pi}\theta_\sigma} \int_{-\infty}^{\sigma} \exp\left[-\frac{(t-\tau_\sigma)^2}{2\theta_\sigma^2}\right] dt, \tag{2.9}$$

where τ_σ is a contrast threshold and $\theta_\sigma = \tau_\sigma/3$ [18]. The local structural fidelity measure S_{local} is applied using a sliding window that runs across the image, resulting in a map that

reflects the variation of structural fidelity across space. Finally, the quality map is averaged to provide a single overall structural fidelity measure of the image

$$S(\mathbf{X}, \mathbf{Y}) = \frac{1}{M} \sum_{i=1}^{M} S_{\text{local}}(\mathbf{x}_i, \mathbf{y}_i) \,, \tag{2.10}$$

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. To account for the sampling density of the image and the distance between the image and the observer, the structural fidelity in TMQI is implemented in a multi-scale fashion [18, 32]. Fig. 2.7 shows such structural fidelity maps at five scales. Here, brighter means better structure preservation.

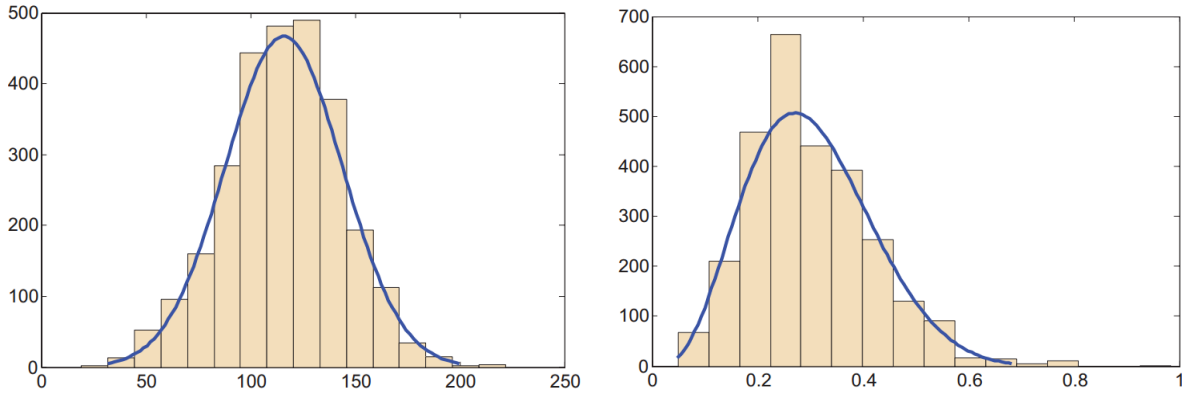


Figure 2.6: Histograms of means fitted by Gaussian PDF (left) and std fitted by Beta PDF (right) of natural images. (Example taken from Yeganeh's paper [18])

The statistical naturalness model N is derived from the statistics of about 3,000 gray-scale images representing many different types of natural scenes [18]. It was found that the histograms of the mean and std can be well fitted by a Gaussian density function P_m and a Beta density function P_d, as shown in Fig. 2.6, respectively [18]. Based on recent vision science studies on the independence of image brightness and contrast [37], the statistical

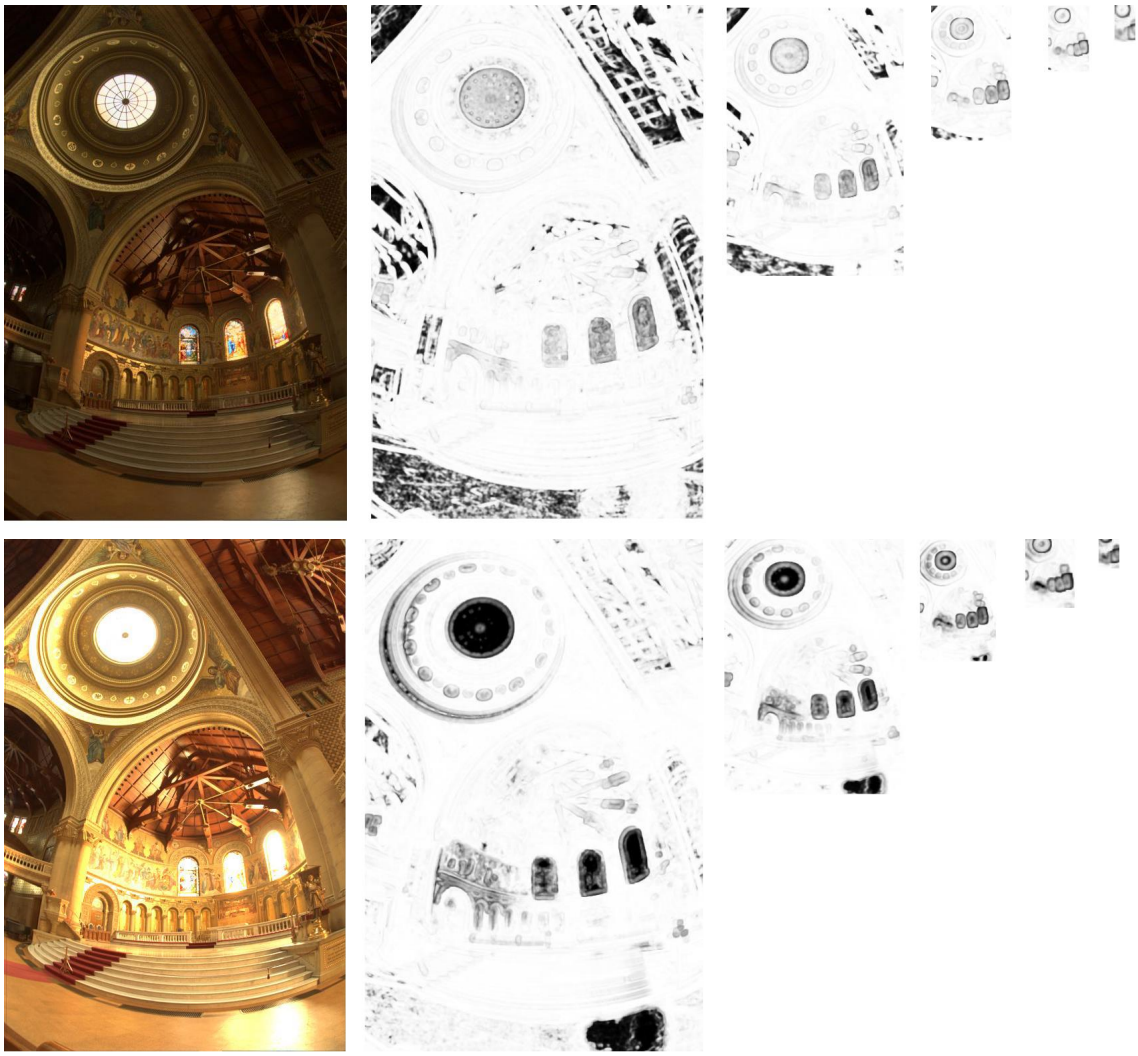


Figure 2.7: Tone mapped LDR images and their structural fidelity maps in five scales. Top: $S = 0.9152$ ($S1 = 0.8940$; $S2 = 0.9341$; $S3 = 0.9428$; $S4 = 0.9143$; $S5 = 0.8277$). Bottom: $S = 0.8614$ ($S1 = 0.9161$; $S2 = 0.9181$; $S3 = 0.8958$; $S4 = 0.8405$; $S5 = 0.7041$). (Example taken from Yeganeh's paper [18])

naturalness is modeled as the product of the two density functions [18]

$$N(\mathbf{Y}) = \frac{1}{K} P_m \, P_d \,, \tag{2.11}$$

where K is a normalization factor given by $K = \max\{P_m P_d\}$. This constrains the N measure to be bounded between 0 and 1.

2.4 Summary

This chapter gives the background information that is highly relavent to our study in later chapters. From the next chapter, we will focus on constructing our new objective IQA model for tone mapped images upon TMQI.

Chapter 3

Enhanced Tone Mapped Image Quality Index

In this chapter, we construct an enhanced tone mapped image quality index (eTMQI) with an emphasis on the differences and improvement over TMQI. Validations on subjective data indicate that the modified structural fidelity and statistical naturalness terms in eTMQI outperform those in TMQI.

3.1 Enhanced Tone Mapped Image Quality Index

Both the structural fidelity and statistical naturalness terms in TMQI [18] have several significant problems. In the following subsections, we will point them out and propose solutions to fix them.

3.1.1 Enhanced Structural Fidelity

Recall that to assess the visibility of local contrast of HDR and tone mapped images, the local std σ undergoes a contrast sensitivity motivated nonlinear mapping function:

$$\tilde{\sigma} = \frac{1}{\sqrt{2\pi}\theta_\sigma} \int_{-\infty}^{\sigma} \exp\left[-\frac{(t-\tau_\sigma)^2}{2\theta_\sigma^2}\right] dt\,. \tag{3.1}$$

where τ_σ is a contrast threshold and $\theta_\sigma = \tau_\sigma/3$ [18]. The above nonlinear mapping is limited in accurately assessing the contrast visibility of HDR image patches. First, even a small change in the local patch of the HDR image (usually due to HDR camera noise) may result in a significant σ. While Equation (3.1) 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. 3.1 illustrates this phenomenon. The homogeneous areas such as the walls and the wood board in the lower middle part of the image are supposed to be considered as contrast invisible in both the HDR and tone mapped images. Equation (3.1) correctly predicts those areas as contrast invisible in the tone mapped image but incorrectly considers them as contrast visible in the HDR image. This explains the corresponding dark areas of the structural fidelity map in the middle of Fig. 3.1[1] (brighter indicates higher structural fidelity). Second, different local patches in the HDR image may also have substantially different dynamic ranges, which corresponds to different soft thresholds τ_σ. In summary, a single τ_σ is insufficient to assess the local contrast visibility of the HDR image.

The above analysis suggests that another nonlinear mapping that is adapted to local luminance levels is desired for the HDR image. Therefore, we keep Equation (3.1) for the tone mapped image and construct another nonlinear mapping that takes the same formula with a different local contrast estimate input and a different soft threshold τ_σ. In particular, we choose the std divided by the mean, or σ/μ, as an estimate of local contrast in the HDR image. This estimate is adapted to local luminance levels and thus is qualitatively consistent with Weber's law, which has been widely used to model luminance

[1]Here, only the structural fidelity map of the finest scale in TMQI is shown. The structural fidelity maps of other scales also exhibit the same problem.

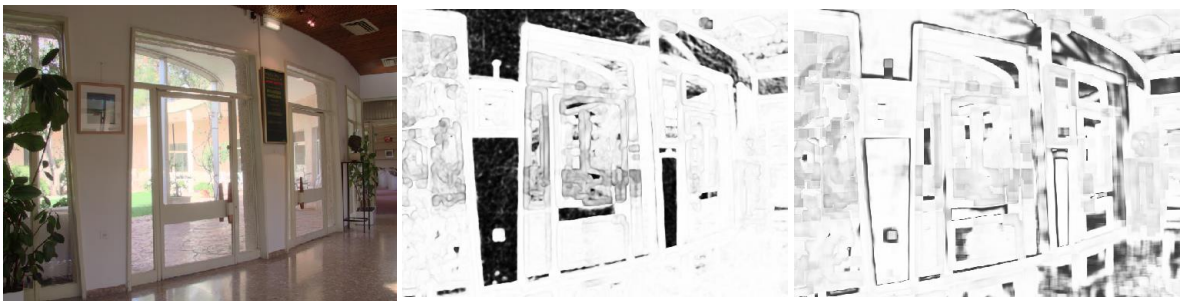


Figure 3.1: Tone mapped "Belgium house" image and its structural fidelity maps. Left: Initial image created by Reinhard's algorithm [3]. Middle and Right: Structural fidelity maps of the left image using TMQI and eTMQI respectively, where brighter indicates higher structural fidelity. We observer that eTMQI gives a more reasonable structural fidelity map which better reflects the contrast visibility of HDR and tone mapped images, respectively.

masking in the HVS. According to Weber's law, the ratio of the just noticeable luminance change ΔI to the background luminance I is approximately a constant for a wide range of luminance values. In the case of complex images, a simple assumption is that μ and σ play the roles of I and ΔI, respectively. The just noticeable ratio corresponds to the soft threshold τ_σ. An additional benefit of this estimate is that it is invariant to linear contrast stretch, which is considered as a standard preprocessing step of the HDR image. The reason follows directly from the local luminance adaptation that cancels out the scale factors in the numerator and the denominator. The right image of Fig. 3.1 shows one example of the structural fidelity map resulting from eTMQI, which better captures the contrast visibility of the HDR and the tone mapped images, respectively.

3.1.2 Enhanced Statistical Naturalness

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 [18]. Due to the independence assumption of image brightness

and contrast [37], the two density functions are multiplied together to obtain the overall statistical naturalness measure [18]

$$N(\mathbf{Y}) = \frac{1}{K} P_m \, P_d \,, \tag{3.2}$$

where K is a normalization factor.

The above statistical naturalness model has two drawbacks. First, the image independent assumption is an over simplification. The model suggests that to be perfectly statistically natural, the tone mapped image should have μ around 116 and σ around 65[2]. However each image may have a different μ and σ to look perfectly natural. In other words, the statistical naturalness model may poorly correlate with the natural appearance of an image. Second, the model is derived from high quality images, with no information about what an unnatural image should look like. Therefore, using this model to penalize the unnaturalness of a tone mapped image is problematic.

Here we propose an image dependent statistical naturalness model based on a subjective experiment to better quantify the unnaturalness of tone mapped images. First of all, we estimate the overall luminance and global contrast of a good quality tone mapped image 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 for many TMOs [3, 4, 38, 39]. The use of logarithmic function assumes that most structural detail in the HDR image lives in a low dynamic range and thus it is reasonable to boost lower luminance levels while compressing higher luminance levels. The quantity is computed by

$$L_x = \exp\left(\frac{1}{|\mathbf{X}|}\sum_{i,j}\log\left(\epsilon + \mathbf{X}(i,j)\right)\right) , \tag{3.3}$$

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. After that, we scale the luminance

[2] $\mu = 116$ and $\sigma = 65$ are values that correspond to the peaks in P_m and P_d, respectively.

by

$$\mathbf{X}^s(i,j) = \frac{k}{L_x}\mathbf{X}(i,j)\,, \tag{3.4}$$

where k is a luminance level related quantity. For an HDR image with normal luminance level, k is typically set from 0.09 to 0.36 [3]. Now we are able to estimate μ_e and σ_e using

$$\mu_e = \frac{L}{|\mathbf{X}|}\left(\sum_{i,j}\frac{\mathbf{X}^s(i,j)}{1+\mathbf{X}^s(i,j)}\right)\,, \tag{3.5}$$

and

$$\sigma_e = \frac{1}{|\mathbf{X}|-1}\sum_{i,j}\left(\frac{\mathbf{X}^s(i,j)}{1+\mathbf{X}^s(i,j)}L-\mu_e\right)^{\frac{1}{2}}\,, \tag{3.6}$$

where L is the dynamic range of the tone mapped image. In the above two equations, we further compress high luminance by a factor of $\mathbf{X}^s$. This may cause detail loss in high luminance regions. 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 our structural fidelity term. This estimation of the initial luminance level of the LDR image is closely related to previous works [3, 28, 40].

After obtaining μ_e and σ_e, we consider them as the most desirable values that a perfect quality tone mapped image should have. On the other hand, we also assume that for each LDR tone mapped image, there should be a certain range of μ and σ, within which the quality degradation of the image by adjusting its μ and σ is negligible. To verify this, we conduct a subjective experiment, in which subjects were asked to first decrease and then increase μ of test LDR images until they saw significant quality degradation. A lower bound μ_l and a upper bound μ_r for each LDR image were recorded. The same procedure is used to obtain a lower std bound σ_l and a upper std bound σ_r for each LDR image[3]. We have selected 60 natural LDR images from the LIVE Database [41] with different μ and σ that cover diverse natural contents. A total of 25 naive observers, including 15 males and

[3]To change μ with σ fixed, we use the formula $\hat{I} = I + \Delta I$, where ΔI is the relative luminance level added to the image globally. To change σ with μ fixed, we use the formula $\hat{I} = s \cdot (I - \mu_I) + \mu_I$, where s is the scale factor and μ_I is the mean of the image.

10 females aged between 22 and 30, participated in the experiment. The four bounds for each LDR image are averaged over all 25 subjects. Fig. 3.2 summarizes the experiment, where we observe that the relationships between μ of test LDR images and their 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 $k1 = 0.6043$, $k2 = 0.6993$ and intercepts

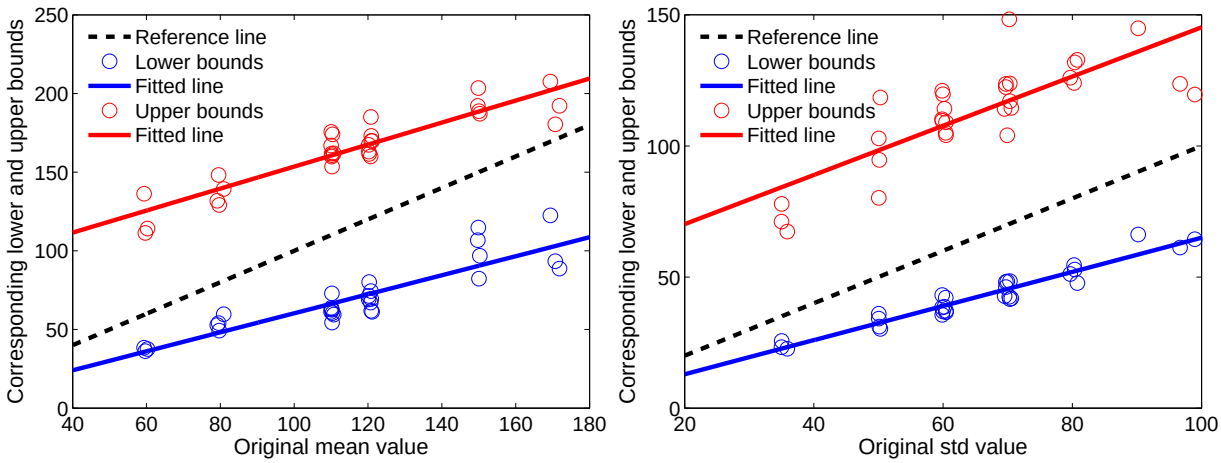


Figure 3.2: Summary of subjective experiment data on acceptable μ and σ values. The blue and red circles in (a) represent the lower and upper bounds of the mean of test LDR images, respectively. The corresponding blue and red solid lines are least square fitted lines. The blue and red circles in (b) represent the lower and upper bounds of the std of test LDR images, respectively. The corresponding blue and red solid lines are lease square fitted lines. The dashed lines in two plots are reference lines that correspond to the original mean and std.

$b1 = -0.1402$, $b2 = 83.6128$ for u_l and u_r, respectively. The R^2 statistics of the two linear models are 0.8008 and 0.8465, which indicate that the linear models explain most variances in the subjective data. Interestingly, when μ of an image is relatively small, $\mu_r - \mu$ is much large than $\mu - \mu_l$. The situation is reversed when μ of an image is large. In other words, the acceptable luminance changes without significantly tampering its visual naturalness saturate at both small and large luminance levels. On the other hand, σ_l and σ_r of a test LDR image can also be well fitted by two linear models using σ as the predictor. The

fitted lines for σ_l and σ_r have slopes $k3 = 0.6504$, $k4 = 0.9386$ and intercepts $b3 = -0.0759$, $b4 = 51.3951$, respectively.

Based on the method described above, given an HDR image, we can estimate μ_e and σ_e of its desired tone mapped image. We can also further predict μ_l, μ_r, σ_l and σ_r of the tone mapped image from μ_e and σ_e with four linear models. Our observation is that μ and σ of a high quality tone mapped image should at least lie in $[\mu_l, \mu_r]$ and $[\sigma_l, \sigma_r]$, and if possible, close to μ_e and σ_e. Therefore, we complete our statistical naturalness model by quantifying the quality drop from μ_e and σ_e to their lower and upper bounds with four Gaussian cumulative distribution functions (CDF). Specifically, the probability of the tone mapped image to be high quality with mean μ is computed by

$$P_m = \begin{cases} \frac{1}{\sqrt{2\pi}\theta_1} \int_{-\infty}^{\mu} \exp\left[-\frac{(t-\tau_1)^2}{2\theta_1^2}\right] dt & \mu \leq \mu_e \\ \frac{1}{\sqrt{2\pi}\theta_2} \int_{-\infty}^{2\mu_r - \mu} \exp\left[-\frac{(t-\tau_2)^2}{2\theta_2^2}\right] dt & \mu > \mu_e \end{cases}, \tag{3.7}$$

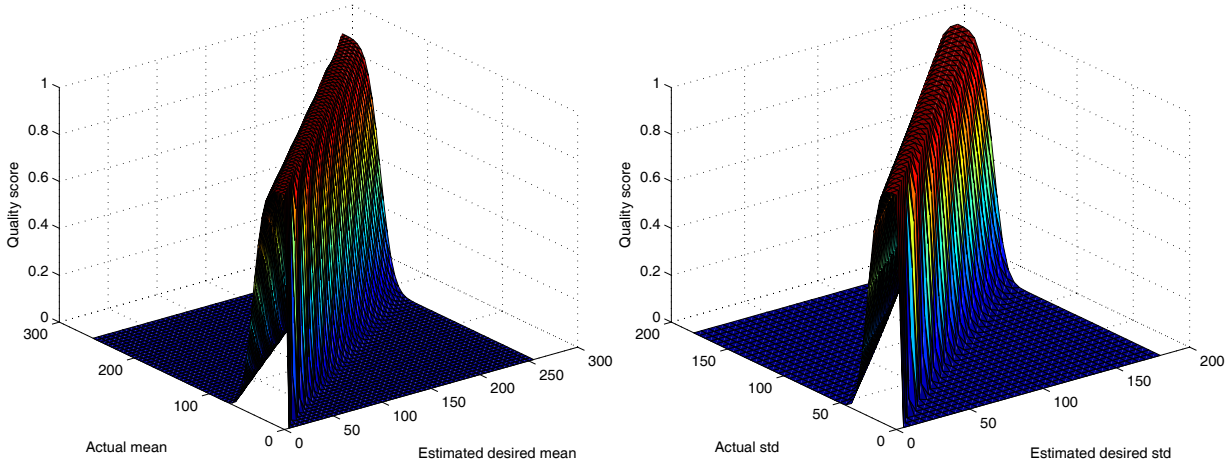
where τ_1 and θ_1 can be uniquely determined by two points $(\mu_l, 0.01)$ and $(\mu_e, 0.99)$ on the Gaussian CDF curve. Correspondingly, τ_2 and θ_2 are uniquely determined by two points $(\mu_r, 0.01)$ and $(\mu_e, 0.99)$ on the Gaussian CDF curve. Similarly, the probability of the tone mapped image to be high quality with std σ is computed by

$$P_d = \begin{cases} \frac{1}{\sqrt{2\pi}\theta_3} \int_{-\infty}^{\sigma} \exp\left[-\frac{(t-\tau_3)^2}{2\theta_3^2}\right] dt & \sigma \leq \sigma_e \\ \frac{1}{\sqrt{2\pi}\theta_4} \int_{-\infty}^{2\sigma_r - \sigma} \exp\left[-\frac{(t-\tau_4)^2}{2\theta_4^2}\right] dt & \sigma > \sigma_e \end{cases}, \tag{3.8}$$

where τ_3 and θ_3 can be uniquely determined by two points $(\sigma_l, 0.01)$ and $(\sigma_e, 0.99)$ on the Gaussian CDF curve. And, τ_4 and θ_4 can be uniquely determined by two points $(\sigma_r, 0.01)$ and $(\sigma_e, 0.99)$ on the Gaussian CDF curve.

The surfaces of P_m and P_d are 3-D plotted in Fig. 3.3. It can be observed that μ and σ of the best quality tone mapped image should match the estimated μ_e and σ_e, which correspond to the peaks around the diagonal lines. The models apply a heavy penalty when $|\mu - \mu_e|$ or $|\sigma - \sigma_e|$ is large.

Similar to Equation (2.11), we multiply these two probabilities to evaluate the overall



(a) P_m (b) P_d

Figure 3.3: The Surfaces of P_m and P_d. It suggests that μ and σ of the best quality tone mapped image should match estimated μ_e and σ_e, which corresponds to the peaks around diagonal lines. The models give heavy penalty when $|\mu - \mu_e|$ or $|\sigma - \sigma_e|$ is large.

statistical naturalness model

$$N(\mathbf{X}, \mathbf{Y}) = P_m \, P_d \, . \tag{3.9}$$

Since $0 \leq P_m, \ P_d \leq 1$, N also lies in $[0, 1]$.

We complete this section by renewing Equation (2.7) to

$$\text{eTMQI}(\mathbf{X}, \mathbf{Y}) = a[S(\mathbf{X}, \mathbf{Y})]^{\alpha} + (1 - a)[N(\mathbf{X}, \mathbf{Y})]^{\beta} \, . \tag{3.10}$$

Although the formulae of original TMQI and eTMQI look similar, the foundations behind them are substantially different.

3.2 Validation of eTMQI

In this section, we compare the eTMQI with TMQI on the well known tone mapped image database [18] using the following criteria:

- Spearman's rank-order correlation coefficient (SRCC) is defined as

$$\mathrm{SRCC} = 1 - \frac{6\sum_{N_I}^{i=1} d_i^2}{N_I(N_I^2 - 1)}\,, \tag{3.11}$$

  where d_i is the difference between the ranks of i-th image in subjective and objective test, N_I is the number of considered images.

- Kendall's rank-order correlation coefficient (KRCC) is another rank-order based non-parametric correlation metric, defined as

$$\mathrm{KRCC} = \frac{N_c - N_d}{\frac{1}{2}N_I(N_I - 1)}\,, \tag{3.12}$$

  where N_c and N_d are the number of concordant and discordant pairs in the data set, respectively.

The better tone mapped IQA model needs to have larger SRCC and KRCC values with subjective evaluations. To construct the subjective database in [18], 20 subjects were asked to rank 8 images in one set from the best to the worst for a total of 15 sets of tone mapped images [18]. The subjective rankings for each image is then averaged, resulting in its mean ranking score within the set [18]. The tone mapped images are generated by 8 TMOs: Reinhard et al. [3], Drago et al. [4], Durand and Dorsey [26], Mantiuk et al. [42], Pattanaik et al. [43], "Exposure and Gamma", "Equalize Histogram" and "Local Adaptation", respectively[4]. Although there are some other subjective databases such as [45, 46], they are not publicly available. A number of parameters are inherited from the original TMQI. These include $C1 = 0.01$, $C2 = 10$ and the soft threshold of contrast visibility for tone mapped images $\tau_\sigma = 2.6303$. Throughout our study, we set the soft threshold of contrast visibility for HDR images $\tau_\sigma = 0.06$ and the luminance level related quantity $k = 0.12$. As for the model parameters in eTMQI, we set $a = 0.5$, $\alpha = 1$ and $\beta = 1$, which emphasize the equal importance between structural fidelity and statistical naturalness terms.

[4]The first five TMOs are implemented by the publicly available software Qtpfsgui [44] and the last three are built in Adobe Photoshop [18].

Table 3.1: Performance evaluation of eTMQI and TMQI on the database [18] using SRCC as the criterion

Image set	eTMQI			TMQI		
	S	N	eTMQI	S	N	TMQI
1	0.7381	0.9048	0.9048	0.6667	0.9048	0.9048
2	0.5714	0.5952	0.5000	0.8095	0.7619	0.7857
3	0.5714	0.7143	0.6905	0.2619	0.7381	0.8095
4	0.7619	0.7857	0.6905	0.8571	0.8571	0.8810
5	0.8810	0.5714	0.6667	0.1429	0.7381	0.7381
6	0.7143	0.9524	0.9762	0.7381	0.9524	0.9762
7	0.9286	0.7381	0.8333	0.8810	0.6429	0.6905
8	0.7619	0.7381	0.6667	0.3333	0.7143	0.7143
9	0.3095	0.7381	0.8095	0.8571	0.3571	0.6905
10	0.8095	0.8810	0.9048	0.6667	0.9048	0.9286
11	0.8095	0.7857	0.8333	0.6429	0.5476	0.8810
12	0.9524	0.5476	0.5952	0.7143	0.5714	0.7143
13	0.5629	0.7066	0.7904	0.9461	0.4311	0.6826
14	0.8333	0.7619	0.7619	0.9524	0.7381	0.7381
15	0.8810	0.7857	0.9048	0.9286	0.9048	0.9524
Average	**0.7391**	**0.7471**	**0.7686**	**0.6932**	**0.7176**	**0.8058**

We first compare structural fidelity and statistical naturalness terms in TMQI and eTMQI, individually. Table 3.1 and Table 3.2 summarize the results. We observe that the structural fidelity and statistical naturalness terms in eTMQI consistently outperform those in TMQI. This is due to more accurate modeling of the contrast visibility of HDR images in a locally adaptive fashion and image dependent qualification of the unnaturalness of tone mapped images.

We then investigate the overall performance of eTMQI. The results are also listed in Tables 3.1 and 3.2, from which we observe that eTMQI outperforms its structural fidelity and statistical naturalness components separately. This supports our assumption that the structural fidelity and statistical naturalness are two relatively independent terms that characterize two different aspects of tone mapped images. However, the overall performance of TMQI is better than eTMQI. It is not surprising since TMQI trains the model

parameters on the same database [18] while eTMQI does not involve any training process. We may expect better performance of eTMQI if its parameters are also trained on the database, but it may be unnecessary because of the potential of overfitting. Furthermore, we will compare eTMQI with TMQI in the space of images using an iterative optimization algorithm described in the next chapter.

3.3 Summary

Systematic comparison of different TMOs are important not only for finding the one with the best average performance but also for pointing out directions for further improvement. TMQI made the first move towards objective IQA of tone mapped LDR images using corresponding HDR images as reference. In this chapter, we proposed eTMQI that breaks the limitations underlying the structural fidelity and statistical naturalness terms in TMQI. As expected, the structural fidelity and statistical naturalness terms in eTMQI better correlate with the subjective database [18] than those in TMQI.

Table 3.2: Performance evaluation of eTMQI and TMQI on the database [18] using KRCC as the criterion

Image set	eTMQI			TMQI		
	S	N	eTMQI	S	N	TMQI
1	0.6429	0.7857	0.7857	0.5000	0.7857	0.7857
2	0.3571	0.4286	0.2857	0.7143	0.5714	0.6429
3	0.3571	0.5714	0.5714	0.1429	0.5714	0.6429
4	0.5714	0.5714	0.5000	0.7143	0.6429	0.7143
5	0.7143	0.4286	0.5000	0.1429	0.6429	0.6429
6	0.5714	0.8571	0.9286	0.6429	0.8571	0.9286
7	0.8571	0.6429	0.7143	0.7857	0.5000	0.5714
8	0.5714	0.6429	0.5000	0.2857	0.5714	0.5714
9	0.3571	0.6429	0.7143	0.7143	0.3571	0.5714
10	0.7143	0.7143	0.7857	0.5000	0.7857	0.8571
11	0.6429	0.6429	0.7143	0.4286	0.4286	0.7143
12	0.8571	0.3571	0.4286	0.5714	0.4286	0.5714
13	0.4001	0.5455	0.6183	0.8365	0.3273	0.5455
14	0.6429	0.5714	0.5714	0.8571	0.6429	0.6429
15	0.7143	0.6429	0.7857	0.8571	0.7857	0.8571
Average	**0.5981**	**0.6030**	**0.6269**	**0.5796**	**0.5933**	**0.6840**

Chapter 4

Iterative Tone Mapping for eTMQI Optimization

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

$$\mathbf{Y}_{\text{opt}} = \arg\max_{\mathbf{Y}} \text{eTMQI}(\mathbf{X}, \mathbf{Y}) \,. \tag{4.1}$$

Solving (4.1) for $\mathbf{Y}_{\text{opt}}$ is a challenging problem due to the complexity of eTMQI and the high dimensionality (the same as the number of pixels in the image). Therefore, we resort to numerical optimization and propose an iterative algorithm. Starting from any initial image $\mathbf{Y}_0$, the proposed algorithm searches for the best solution in the space of all images. Specifically, in each iteration, we first adopt a gradient ascent algorithm to improve the structural fidelity S. After that we solve a constrained optimization problem to improve the statistical naturalness N. These two steps are applied alternately until convergence. Details of the algorithm are elaborated as follows.

4.1 Proposed Iterative Tone Mapping Algorithm

In the k-th iteration, given the result image $\mathbf{Y}_k$ from the last iteration, a gradient ascent algorithm is first applied to improve the structural fidelity:

$$\hat{\mathbf{Y}}_k = \mathbf{Y}_k + \lambda \nabla_{\mathbf{Y}} S(\mathbf{X}, \mathbf{Y})|_{\mathbf{Y}=\mathbf{Y}_k}\,, \tag{4.2}$$

where $\nabla_{\mathbf{Y}} S(\mathbf{X}, \mathbf{Y})$ is the gradient of $S(\mathbf{X}, \mathbf{Y})$ with respect to $\mathbf{Y}$ and λ controls the updating speed. To compute the gradient $\nabla_{\mathbf{Y}} S(\mathbf{X}, \mathbf{Y})$, we start from the local structural fidelity and rewrite (2.8) as

$$S_{\text{local}}(\mathbf{x}, \mathbf{y}) = \frac{A_1\, A_2}{B_1\, B_2}\,, \tag{4.3}$$

where

$$\begin{aligned} A_1 &= 2\tilde{\sigma}_x \tilde{\sigma}_y + C_1 && (4.4)\\ B_1 &= \tilde{\sigma}_x^2 + \tilde{\sigma}_y^2 + C_1 && (4.5)\\ A_2 &= \sigma_{xy} + C_2 && (4.6)\\ B_2 &= \sigma_x \sigma_y + C_2\,. && (4.7) \end{aligned}$$

By treating both image patches as as column vectors of length N_w, we can represent

$$\begin{aligned} \mu_y &= \frac{1}{N_w}\mathbf{1}^T\,\mathbf{y} && (4.8)\\ \sigma_y^2 &= \frac{1}{N_w - 1}(\mathbf{y} - \mu_y)^T(\mathbf{y} - \mu_y) && (4.9)\\ \sigma_{xy} &= \frac{1}{N_w - 1}(\mathbf{x} - \mu_x)^T(\mathbf{y} - \mu_y)\,. && (4.10) \end{aligned}$$

The gradient of local structural fidelity measure with respect to $\mathbf{y}$ can then be expressed as

$$\begin{aligned} \nabla_{\mathbf{y}} S_{\text{local}}(\mathbf{x}, \mathbf{y}) &= \frac{(A_1' A_2 + A_1 A_2')}{B_1 B_2} \\ &- \frac{(B_1' B_2 + B_1 B_2') A_1 A_2}{(B_1 B_2)^2}\,, \end{aligned} \tag{4.11}$$

where

$$A_1' = \nabla_{\mathbf{y}} A_1\,,\ B_1' = \nabla_{\mathbf{y}} B_1\,,\ A_2' = \nabla_{\mathbf{y}} A_2\,,\ B_2' = \nabla_{\mathbf{y}} B_2\,. \tag{4.12}$$

Noting that

$$\nabla_{\mathbf{y}} \sigma_y = \frac{1}{N_w \sigma_y}(\mathbf{y} - \mu_y) \tag{4.13}$$

$$\nabla_{\mathbf{y}} \sigma_{xy} = \frac{1}{N_w}(\mathbf{x} - \mu_x)\,, \tag{4.14}$$

we have

$$\begin{aligned} A_1' &= 2\tilde{\sigma}_x \nabla_{\mathbf{y}} \tilde{\sigma}_y \\ &= \frac{2\tilde{\sigma}_x}{\sqrt{2\pi}\theta_\sigma} \exp\left[-\frac{(\sigma_y - \tau_\sigma)^2}{2\theta_\sigma^2}\right] \cdot \nabla_{\mathbf{y}} \sigma_y \\ &= \sqrt{\frac{2}{\pi}}\, \frac{\tilde{\sigma}_x}{N_w \theta_\sigma \sigma_y} \exp\left[-\frac{(\sigma_y - \tau_\sigma)^2}{2\theta_\sigma^2}\right] (\mathbf{y} - \mu_y)\,, \end{aligned} \tag{4.15}$$

$$\begin{aligned} B_1' &= 2\tilde{\sigma}_y \nabla_{\mathbf{y}} \tilde{\sigma}_y \\ &= \sqrt{\frac{2}{\pi}}\, \frac{\tilde{\sigma}_y}{N_w \theta_\sigma \sigma_y} \exp\left[-\frac{(\sigma_y - \tau_\sigma)^2}{2\theta_\sigma^2}\right] (\mathbf{y} - \mu_y)\,, \end{aligned} \tag{4.16}$$

$$A_2' = \frac{1}{N_w}(\mathbf{x} - \mu_x)\,, \tag{4.17}$$

$$B_2' = \sigma_x \nabla_{\mathbf{y}} \sigma_y = \frac{\sigma_x}{N_w \sigma_y}(\mathbf{y} - \mu_y)\,. \tag{4.18}$$

Plugging (4.4), (4.5), (4.6), (4.7), (4.15), (4.16), (4.17) and (4.18) into (4.11), we obtain the gradient of local structural fidelity. Finally, we can 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

$$\nabla_{\mathbf{Y}} S(\mathbf{X}, \mathbf{Y}) = \frac{1}{M} \sum_{i=1}^{M} \mathbf{R}_i^T \nabla_{\mathbf{y}} S_{\text{local}}(\mathbf{x}, \mathbf{y})|_{\mathbf{x}=\mathbf{x}_i, \mathbf{y}=\mathbf{y}_i} \,, \tag{4.19}$$

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.

Upon finishing the structural fidelity update, we obtain an intermediate image $\hat{\mathbf{Y}}_k$. Next, we improve the statistical naturalness to obtain $\mathbf{Y}_{k+1}$ through a three-segment equipartition monotonic piecewise linear function

$$y_{k+1}^i = \begin{cases} (3/L)a\hat{y}_k^i & 0 \le \hat{y}_k^i \le L/3 \\ (3/L)(b-a)\hat{y}_k^i + (2a-b) & L/3 < \hat{y}_k^i \le 2L/3 \\ (3/L)(L-b)\hat{y}_k^i + (3b-2L) & 2L/3 < \hat{y}_k^i \le L \end{cases} . \tag{4.20}$$

This is essentially a point-wise intensity transformation with its parameters a and b ($0 \le a \le b \le L$) chosen so that μ and σ of $\mathbf{Y}_{k+1} = \{y_{k+1}^i \text{ for all } i\}$ better matches μ_e and σ_e of the desired tone mapped image. Recall that μ_e and σ_e are estimated from the corresponding HDR image using Equation (3.5) and Equation (3.6), respectively.

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

$$\begin{aligned} \mu_{k+1}^e &= \hat{\mu}_k + \lambda_m(\mu_e - \hat{\mu}_k) \\ \sigma_{k+1}^e &= \hat{\sigma}_k + \lambda_d(\sigma_e - \hat{\sigma}_k) \,, \end{aligned} \tag{4.21}$$

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. We then compute the parameters a and b by solving the

following constrained optimization problem

$$\{a, b\}_{\text{opt}} = \arg\min_{\{a,b\}} ||\mu_{k+1} - \mu^e_{k+1}||^2 + \eta||\sigma_{k+1} - \sigma^e_{k+1}||^2$$
$$\text{subject to } 0 \leq a \leq b \leq L\,, \tag{4.22}$$

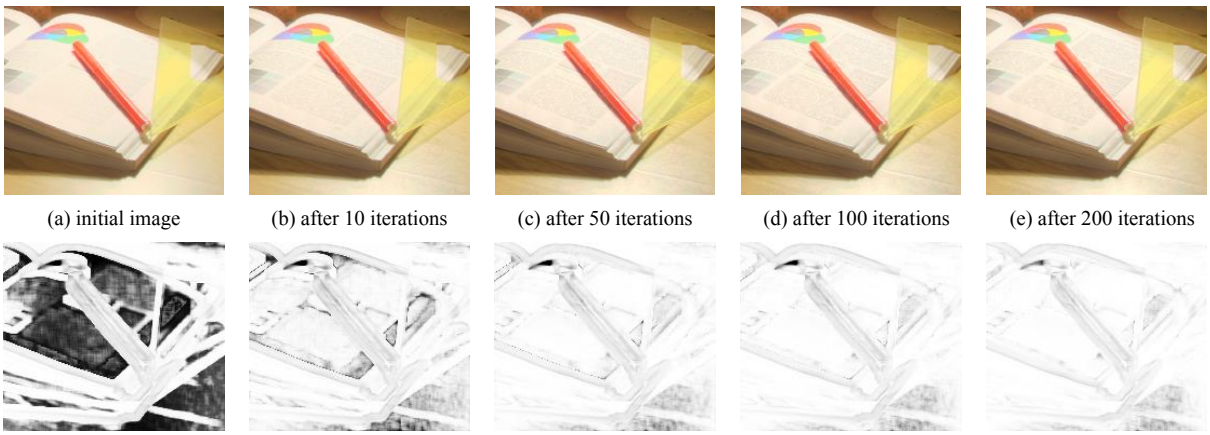
where η adjusts the weights between the mean and std terms. We adopt a standard interior-point algorithm [47] with a maximum 10 iterations to solve this problem. Once the optimal values of a and b are obtained, they are plugged into (4.20) to create the resulting image $\mathbf{Y}_{k+1}$, which is subsequently fed into the $(k+1)$-th iteration.

These two steps alternate until $||\mathbf{Y}_{k+1} - \mathbf{Y}_k||_2$ is smaller than a threshold ϵ. Technically, the above iterative algorithm works as well if TMQI is chosen as the optimization metric [9]. The only difference lies in Equation (4.21), where μ_e and σ_e are replaced with two constants corresponding to the peaks in its P_m and P_d models[1]. We have five free parameters in the proposed algorithm. For fair comparison between eTMQI and TMQI, we use the same set of parameter values in all experiments, which are $\epsilon = 0.1$, $\lambda = 0.3$, $\lambda_m = \lambda_d = 0.03$ and $\eta = 1$.

4.2 Experimental Results

We first examine the roles of the structural fidelity and statistical naturalness components in eTMQI separately. In Fig. 4.1, 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 structures after a sufficient number of iterations. The improvement of structure details is also well reflected by the structural fidelity maps, which eventually evolve to a nearly uniform white image. By contrast, in Fig. 4.2, the initial "building" image is created by a Gamma correction mapping ($\gamma = 2.2$),

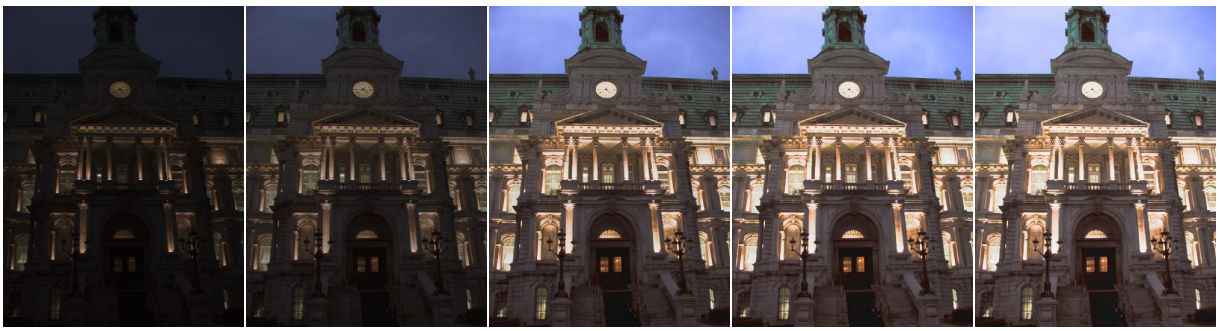
[1] In order to optimize TMQI, we have to work only on its finest scale.



(a) initial image (b) after 10 iterations (c) after 50 iterations (d) after 100 iterations (e) after 200 iterations

(f) 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

Figure 4.1: 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 visulization.



(a) initial image, N = 0.000 (b) 10 iterations, N = 0.001 (c) 50 iterations, N = 0.428 (d) 100 iterations, N = 0.868 (e) 200 iterations, N = 0.971

Figure 4.2: Tone mapped "building" images. (a) Initial image created by Gamma mapping ($\gamma = 2.2$). (b)-(e) Images created using iterative statistical naturalness update only.

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.

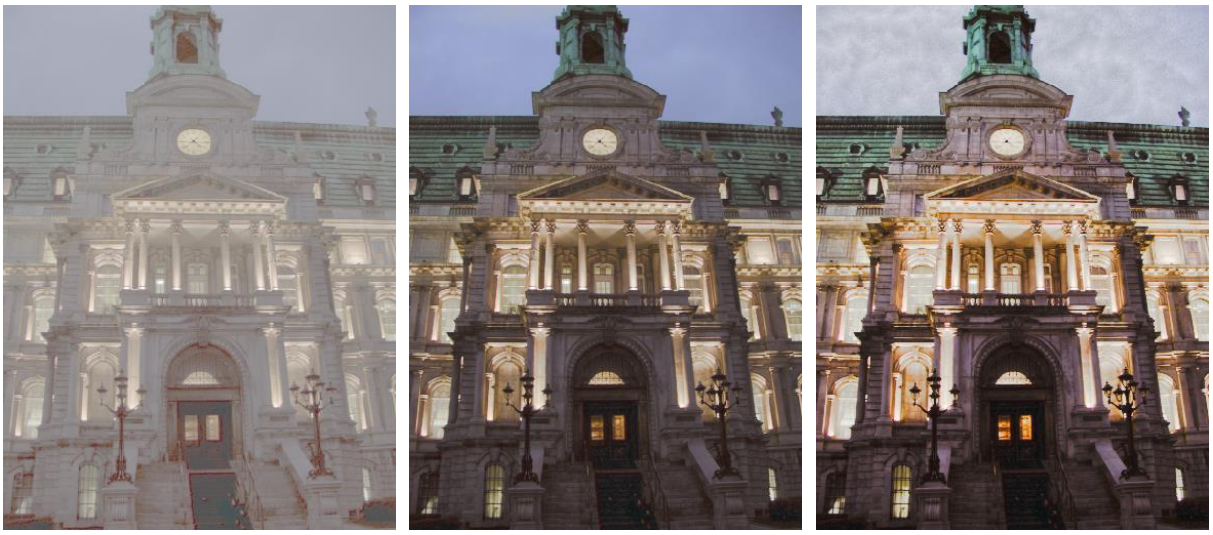
To fully compare the performance of eTMQI and TMQI using this iterative optimization algorithm, we select a wide range of HDR images, containing both indoor and outdoor scenes, human and static objects, as well as day and night scenes. The initial images for this algorithm are also generated by many different TMOs, ranging from simple ones such as Gamma mapping ($\gamma = 2.2$) and log-normal mapping to sophisticated ones such as Reinhard's method [3] and Drago's method [4], both of which are considered as one of the best TMOs on several independent subjective tests [12, 18].

(a) (b) (c)

Figure 4.3: Tone mapped "Woman" images. (a) Initial image created by Gamma mapping. (b) eTMQI-optimized image with $S = 0.9860$, $N = 0.9998$ and eTMQI $= 0.9929$. (c) TMQI-optimized image with $S = 0.9674$, $N = 0.9995$ and TMQI $= 0.9919$.

Fig. 4.3 shows the comparison of eTMQI with TMQI on the "Woman" image initialized by Gamma mapping, which creates dark background with missing structures. Both

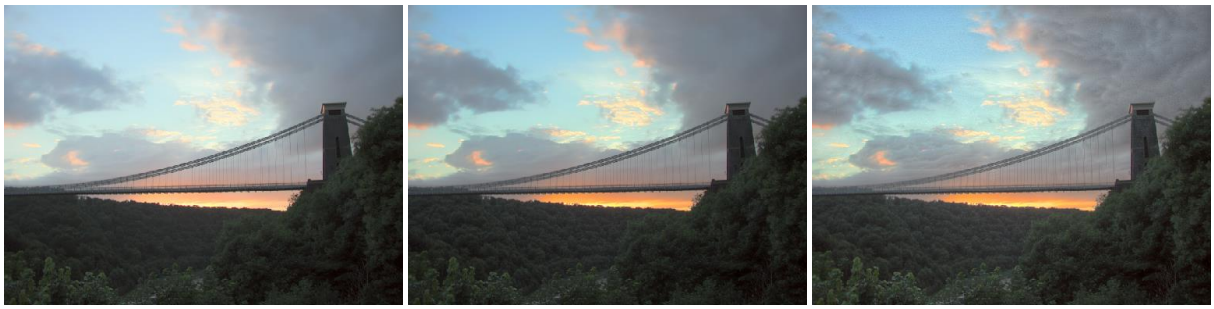
eTMQI- and TMQI-optimized images recover the structures of the background such as the white door, the yellow board and the photo frame, and present a better overall brightness. However, the TMQI-optimized image suffers from heavy noise in homogenous regions such as in the wall and on the floor. The boosted noise artifact boils down to its structural fidelity term, which treats all local regions in the HDR image as contrast visible. In comparison, the eTMQI-optimized image is much cleaner and sharper.



(a) (b) (c)

Figure 4.4: Tone mapped "Clock building" images. (a) Initial image created by log-normal mapping. (b) eTMQI-optimized image with $S = 0.9763$, $N = 0.9998$ and eTMQI $= 0.9881$. (c) TMQI-optimized image with $S = 0.9380$, $N = 1.0000$ and TMQI $= 0.9845$.

Fig. 4.4 shows the comparison of eTMQI with TMQI on the "Clock building" image. The initial image created by log-normal mapping preservers most structures but looks unrealistic due to its blanched appearance. This problem is largely alleviated in the eTMQI-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 stems from its statistical naturalness term, which drags μ and σ of all tone mapped images towards 116 and 65 regardless of their contents and luminance



(a) (b) (c)

Figure 4.5: Tone mapped "Bristol bridge" images. (a) Initial image created by Reinhard's method [3]. (b) eTMQI-optimized image with $S = 0.9938$, $N = 0.9998$ and eTMQI $= 0.9968$. (c) TMQI-optimized image with $S = 0.9408$, $N = 0.9999$ and TMQI $= 0.9852$.

(a) (b) (c)

Figure 4.6: Tone mapped "Grove" images. (a) Initial image by Drago's method [4]. (b) eTMQI-optimized image with $S = 0.9782$, $N = 0.9998$ and eTMQI $= 0.9890$. (c) TMQI-optimized image with $S = 0.9614$, $N = 0.9998$ and TMQI $= 0.9904$.

levels, respectively [18]. This is in appropriate for a night scene like "Clock building" which should have the desired values to be $\mu_e < 116$ and $\sigma_e < 65$ (In eTMQI, $\mu_e = 101$ and $\sigma_e = 48$). Moreover, annoying noise appears in the sky region of the TMQI-optimized image.

Fig. 4.5 shows the comparison of eTMQI with TMQI on the "Bristol bridge" image with initial image created by Reinhard's method [3]. Although the initial image of Fig. 4.5(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. 4.5(b), the proposed iterative algorithm using eTMQI recovers these fine details and makes them much sharper. Moreover, the overall appearance is softer and thus more pleasant. In Fig. 4.5(c), we can see that the iterative algorithm using TMQI heavily boosts noise in the sky and cloud regions, which may lead to quality degradation when compared with the initial image. This reveals the problem of TMQI in quality assessment of tone mapped images.

Fig. 4.6 shows the comparison of eTMQI with TMQI on the "Grove" image with initial image created by Drago's method [4]. Again, in the eTMQI-optimized image of Fig. 4.6(b), fine details such as leafs between the two big trees and the tree barks are faithfully recovered and sharpened. The overall appearance is also more vivid. However, in Fig. 4.6(c), the iterative algorithm using TMQI over stretches the global contrast, which darkens the tree trunks and whitens the leafs and the sky. This inevitably damages the naturalness of the initial image and leads to quality degradation.

Table 4.1 lists eTMQI values of the initial and converged images. It can be seen that the proposed algorithm consistently converges to images with both high structural fidelity and high statistical naturalness, and thus produces high eTMQI values even when the initial images are created by the most competitive state-of-the-art TMOs.

We also conduct another extensive subjective experiment to compare eTMQI with TMQI using the iterative optimization algorithm. In particular, we select 15 HDR images which contain various natural scenes shown in Fig. 4.7 and adopt gamma mapping, log-normal algorithm and Reinhard's method [3] to tone map them to 15×3 LDR images. We then consider them as initial images of the iterative optimization algorithm and obtain

Table 4.1: eTMQI comparison between initial and converged images

Image		Gamma	Reinhard [3]	Drago [4]	Log-normal
Bridge	initial image	0.8093	0.9232	0.8848	0.7439
	converged image	0.9928	0.9929	0.9938	0.9944
Lamp	initial image	0.5006	0.9387	0.8717	0.7371
	converged image	0.9906	0.9925	0.9910	0.9894
Memorial	initial image	0.4482	0.9138	0.8685	0.7815
	converged image	0.9895	0.9894	0.9868	0.9867
Woman	initial image	0.6764	0.8891	0.8918	0.8026
	converged image	0.9941	0.9947	0.9943	0.9937

Figure 4.7: 15 HDR test images to compare eTMQI with TMQI using the iterative optimization algorithm. The tone mapped images shown here are eTMQI-optimized with initial images created by Reinhard's method [3].

15×3 eTMQI-optimized images and 15×3 TMQI-optimized images, respectively. In summary, we have 15 sets of tone mapped images, each of which contains 9 images.

A desktop PC with Intel(R) Core(TM) i7-2600 dual 3.40GHz CPU was used in the subjective user study. The test environment in the lab was setup as a normal indoor office workspace with ordinary illumination level. All images are displayed on an LCD monitor at a resolution of 2560×1600 pixels with Truecolor (32bit) at 60Hz. The monitor was calibrated in accordance with the recommendations of ITU-T BT.500 [48]. A customized MATLAB figure window was used to render the images on the screen. During the test, all 9 tone mapped images in the same set were shown to the subject at the same time in random spatial order on one computer screen at actual pixel resolution. The study adopted a multi-stimulus quality scoring strategy without showing the reference HDR image[2]. A total of 24 naïve observers, including 9 males and 15 females aged between 22 and 30, mostly graduate students at the University of Waterloo, participated in the subjective experiment. The subjects had the freedom to move their positions to get closer or further from the screen for better observation. All subject ratings were recorded with pen and paper during the study. To minimize the influence of fatigue effect, the length of a session was limited to 30 minutes. For each image set, the subject was asked to give an individual integer score for the perceptual quality of each tone mapped image. The score ranges from 0 to 10, where 0 denotes the worst quality and 10 denotes the best. Compared with pair-comparison and ranking strategies, the advantages of the adopted method are (1) more efficient, as multiple scores are collected at one time; (2) instead of ranking order, absolute quality scores are collected to facilitate overall performance evaluation across image sets; (3) cross-content quality comparison can be conducted to develop more generalized quality models. The final quality score for each individual image is computed as the average of subjective scores, named mean opinion score (MOS), from all valid subjects. The results are listed in Tables 4.2, 4.3 and 4.4, from which we have several interesting observations. First, using eTMQI as the optimization metric, the iterative optimization algorithm leads to consistent perceptual gain for all three different types of initial images. By contrast, the perceptual gain obtained by optimizing TMQI is much less compared with that obtained by

[2]In the absence of HDR displays, we actually are unable to show HDR images directly.

optimizing eTMQI, when initial images are created by Gamma and log-normal mapping. Even worse, the quality of TMQI-optimized images decreased dramatically compared with initial images created by Reinhard's method [3]. Second, the best quality image on average is eTMQI-optimized with the initial image created by Reinhard's method [3]. It is not surprising because the proposed optimization algorithm can only guarantee to find a local optimum due to the complexity of the search space; thus better initial images often lead to better local optima, which correspondingly have better perceptual quality.

Table 4.2: Mean opinion scores of test tone mapped images with initial images created by Gamma mapping

Images	Mean opinion scores (MOS)		
	Initial	TMQI-optimized	eTMQI-optimized
1	1.00	4.71	7.46
2	1.54	4.92	8.25
3	0.25	4.50	6.71
4	3.33	4.63	6.54
5	0.54	3.88	7.67
6	0.58	3.29	7.38
7	1.29	5.67	7.38
8	1.54	3.79	6.75
9	0.96	4.38	7.25
10	6.50	5.21	6.42
11	0.46	4.83	7.13
12	3.63	5.21	7.17
13	5.33	3.04	6.00
14	2.17	2.75	6.29
15	4.67	5.13	5.83
Average	**2.25**	**4.39**	**6.95**

To ascertain the differences in average MOSs are statistically significant, we test a hypothesis using t-statistics. Since the number of samples in each column of Tables 4.2, 4.3 and 4.4 is $15 \times 24 = 360$, which exceeds 40 [49], the Gaussian assumption of the subjective score is approximately verified by the central limit theorem. The test statistic is the

Table 4.3: Mean opinion scores of test tone mapped images with initial images created by log-normal mapping

Images	Mean opinion scores (MOS)		
	Initial	TMQI-optimized	eTMQI-optimized
1	5.58	4.13	7.29
2	2.13	4.33	7.83
3	1.88	3.79	6.25
4	3.33	4.96	7.17
5	4.58	4.25	7.13
6	5.21	3.17	6.58
7	2.38	5.33	7.17
8	3.29	3.83	6.54
9	2.67	3.79	7.29
10	2.54	5.13	7.75
11	4.00	2.79	5.50
12	3.46	3.00	7.08
13	2.67	3.29	6.83
14	3.58	2.83	5.75
15	1.92	3.88	6.17
Average	**3.28**	**3.90**	**6.82**

Table 4.4: Mean opinion scores of test tone mapped images with initial images created by Reinhard's method [3]

Images	Mean opinion scores (MOS)		
	Initial	TMQI-optimized	eTMQI-optimized
1	7.29	4.71	7.92
2	8.17	5.25	8.38
3	8.25	4.83	8.33
4	6.96	5.63	7.79
5	7.79	4.46	8.58
6	7.92	3.33	7.50
7	8.00	5.67	8.33
8	8.00	4.08	8.46
9	8.21	4.33	8.50
10	8.04	6.50	8.50
11	7.83	3.63	8.13
12	7.38	3.79	7.63
13	7.79	3.04	8.04
14	8.47	3.04	7.67
15	6.33	6.08	6.79
Average	**7.76**	**4.56**	**8.04**

difference of means divided by the pool estimator of the standard deviation [49]. The null hypothesis is that the average MOS of one image category (there are 9 categories which correspond to 9 columns in Tables 4.2, 4.3 and 4.4) is statistically indistinguishable (with 95% confidence) from the average MOS of another image category. The results are listed in Table 4.5, from which we conclude that the eTMQI-optimized images have statistically higher average MOS, which lead to better perceptual quality.

Table 4.5: Statistical significance matrix based on average MOS. 1 means that the average MOS for the row is statistically higher than that of the column. 0 means that it is statistically lower, and "-" means that it is statistically indistinguishable. GI: Gamma mapping initialized; LI: log-normal mapping initialized; RI: Reinhard's method [3] initialized; TO: TMQI-optimized and EO: eTMQI-optimized

	GI	GITO	GIEO	LI	LITO	LIEO	RI	RITO	RIEO
GI	-	0	0	0	0	0	0	0	0
GITO	1	-	0	1	1	0	0	-	0
GIEO	1	1	-	1	1	-	0	1	0
LI	1	0	0	-	0	0	0	0	0
LITO	1	0	0	1	-	0	0	0	0
LIEO	1	1	-	1	1	-	0	1	0
RI	1	1	1	1	1	1	-	1	0
RITO	1	-	0	1	1	0	0	-	0
RIEO	1	1	1	1	1	1	1	1	-

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

4.3 Convergence Analysis

Because of the complexity of initial TMOs, eTMQI, and the dimension and complexity of the search space, analytical performance assessment of the proposed algorithm is not possible. Therefore, we observe the convergence performance empirically. Figs. 4.8 and 4.9

show the structural fidelity and statistical naturalness measures as functions of iteration using different initial images as starting points. There are several useful observations. First, both measures increase monotonically with the number of iterations. Second, the proposed algorithm converges in all cases whether 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 the local optima.

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 an Intel Quad-Core 2.67 GHz computer.

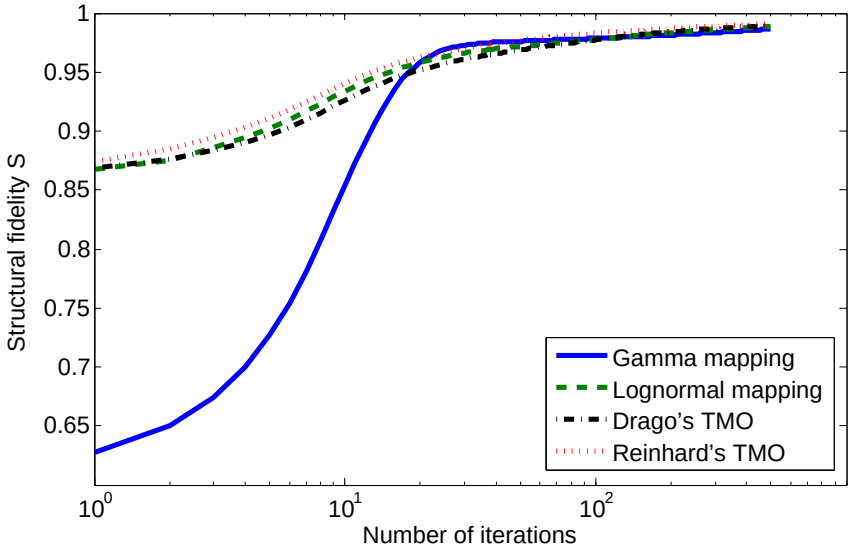


Figure 4.8: Structural fidelity as a function of iteration with initial "woods" images created by different TMOs.

4.4 Summary

We propose a substantially different approach to design TMO, where instead of using any pre-defined systematic computational structure (such as image transformation or con-

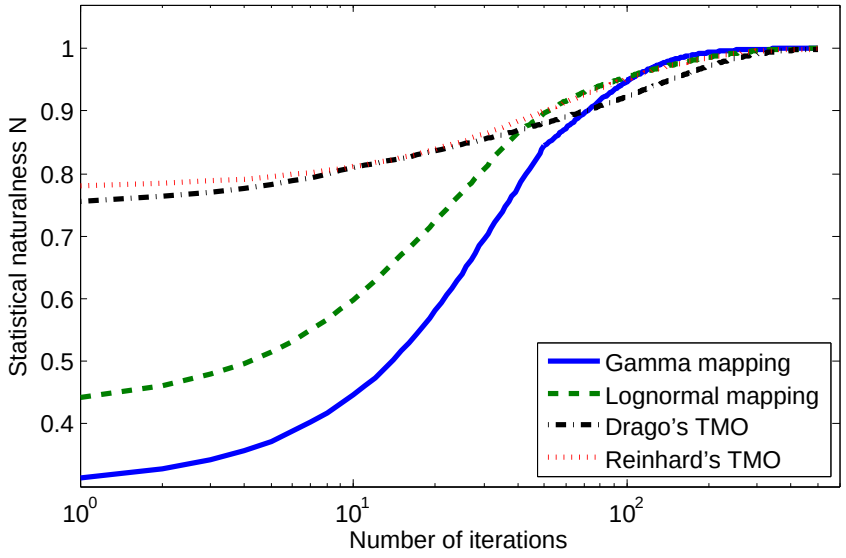


Figure 4.9: Statistical naturalness as a function of iteration with initial "woods" images created by different TMOs.

trast/edge enhancement) for tone mapping, we navigate in the space of all images, searching for the image that optimizes eTMQI. The navigation involves an iterative process that alternately improves the structural fidelity and statistical naturalness of the resulting image. The current work opens the door to a new class of TMO approaches. In the future, if other and better IQA models for tone mapped images are available, their performance may also be tested or improved using our optimization framework.

Chapter 5

Conclusions and Future Work

This thesis presents an objective IQA model, namely eTMQI, to predict the perceptual quality of tone mapped images. On the basis of the original TMQI, we first construct an improved contrast sensitivity model-based nonlinear mapping to better capture the local contrast visibility of HDR images. The modified structural fidelity term results in a more meaningful and useful structural fidelity map and guides the proposed optimization algorithm to recover fine detail with more robustness to noise artifact. We then build an image dependent statistical naturalness model and quantify the unnaturalness of tone mapped LDR images based on a subjective experiment. Validations on an independent subjective database indicate that eTMQI outperforms TMQI on structural fidelity and statistical naturalness terms separately. Furthermore, we propose an iterative optimization algorithm not only to compare eTMQI and TMQI in the space of all images but also to produce better quality tone mapped images when eTMQI is chosen as the optimization goal. Numerical and subjective experiments suggest that eTMQI is a robust objective quality metric for quality prediction of tone mapped images and consistently outperforms TMQI.

The quality assessment and enhancement of tone mapped images are still at an early state. Many topics are worth further exploring. First, there is still room for improving the accuracy of objective IQA models in predicting the perceived quality of tone mapped

images. In other words, better objective IQA models are desired. In the case of improving eTMQI, the current statistical naturalness model is still crude and global; finding local patch-based models has great potentials in creating more natural-looking tone mapped images by generating a local statistical naturalness quality map. What's more, eTMQI works with luminance only; to better predict the perceptual quality of tone mapped images, color information needs to be taken into consideration for designing better quality models. Second, when performing tone mapping, existing TMOs also introduce a variety of distortions such as color saturation, blanched appearances and artificial edges; simple and efficient algorithms are desired to classify and quantify these distortions, which may in return contribute to tone mapped IQA. Third, the current eTMQI-based optimization algorithm can only find local optima. Deeper understanding of the search space is necessary to find better initial guesses and to avoid being easily trapped to local optima.

References

[1] B. A. Wandell, *Foundations of vision.* Sinauer Associates, 1995. 1

[2] E. Reinhard, W. Heidrich, P. Debevec, S. Pattanaik, G. Ward, and K. Myszkowski, *High Dynamic Range Imaging: Acquisition, Display, and Image-based Lighting.* Morgan Kaufmann, 2010. xiii, 1, 2, 7, 12

[3] E. Reinhard, M. Stark, P. Shirley, and J. Ferwerda, "Photographic tone reproduction for digital images," *ACM Transactions on Graphics*, vol. 21, no. 3, pp. 267–276, 2002. xi, xiii, xiv, xv, xvi, 3, 7, 9, 10, 11, 12, 23, 24, 25, 29, 37, 38, 39, 41, 42, 43, 45, 47, 48

[4] F. Drago, K. Myszkowski, T. Annen, and N. Chiba, "Adaptive logarithmic mapping for displaying high contrast scenes," *Computer Graphics Forum*, vol. 22, pp. 419–426, 2003. xiii, xvi, 3, 10, 11, 12, 24, 29, 39, 41, 42, 43

[5] Q. Shan, J. Jia, and M. S. Brown, "Globally optimized linear windowed tone mapping," *IEEE Trans. on Visualization and Computer Graphics*, vol. 16, no. 4, pp. 663–675, 2010. 3, 7, 10

[6] S. Ferradans, M. Bertalmio, E. Provenzi, and V. Caselles, "An analysis of visual adaptation and contrast perception for tone mapping," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 33, no. 10, pp. 2002–2012, 2011. 3, 7

[7] B. Gu, W. Li, M. Zhu, and M. Wang, "Local edge-preserving multiscale decomposition for high dynamic range image tone mapping," *IEEE Trans. on Image Processing*, vol. 22, no. 1, pp. 70–79, 2013. 3, 7, 12

[8] H. Yeganeh and Z. Wang, "High dynamic range image tone mapping by maximizing a structural fidelity measure," in *IEEE International Conference on Acoustics, Speech and Signal Processing*, pp. 1879–1883, May 2013. 3, 7

[9] K. Ma, H. Yeganeh, K. Zeng, and Z. Wang, "High dynamic range image tone mapping by optimizing tone mapped image quality index," in *IEEE International Conference on Multimedia and Expo*, 2014. 3, 7, 37

[10] F. Drago, W. L. Martens, K. Myszkowski, and H.-P. Seidel, "Perceptual evaluation of tone mapping operators," in *ACM SIGGRAPH 2003 Sketches & Applications*, pp. 1–1, 2003. 3

[11] M. Čadík and P. Slavik, "The naturalness of reproduced high dynamic range images," in *International Conference on Information Visualisation*, pp. 920–925, 2005. 3

[12] M. Čadík, M. Wimmer, L. Neumann, and A. Artusi, "Evaluation of hdr tone mapping methods using essential perceptual attributes," *Computers & Graphics*, vol. 32, no. 3, pp. 330–349, 2008. 3, 10, 39

[13] C. Cavaro-Ménard, L. Zhang, and P. Le Callet, "Diagnostic quality assessment of medical images: Challenges and trends," in *IEEE 2nd European Workshop on Visual Information Processing*, pp. 277–284, 2010. 3

[14] Z. Wang and A. C. Bovik, "Mean squared error: love it or leave it? a new look at signal fidelity measures," *IEEE Signal Processing Magazine*, vol. 26, no. 1, pp. 98–117, 2009. xiv, 3, 13, 14, 15, 16

[15] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *IEEE Trans. on Image Processing*, vol. 13, no. 4, pp. 600–612, 2004. 3, 13, 15, 17

[16] R. Mantiuk, S. J. Daly, K. Myszkowski, and H.-P. Seidel, "Predicting visible differences in high dynamic range images: model and its calibration," in *Proc. SPIE*, vol. 5666, pp. 204–214, 2005. 3

[17] T. O. Aydin, R. Mantiuk, K. Myszkowski, and H.-P. Seidel, "Dynamic range independent image quality assessment," *ACM Transactions on Graphics*, vol. 27, no. 3, p. 69, 2008. 3

[18] H. Yeganeh and Z. Wang, "Objective quality assessment of tone-mapped images," *IEEE Trans. on Image Processing*, vol. 22, no. 2, pp. 657–667, 2013. xi, xiv, 3, 10, 15, 17, 18, 19, 20, 21, 22, 23, 24, 28, 29, 30, 31, 32, 39, 42

[19] J. M. DiCarlo and B. A. Wandell, "Rendering high dynamic range images," in *Electronic Imaging*, pp. 392–401, International Society for Optics and Photonics, 2000. 7

[20] G. W. Larson, H. Rushmeier, and C. Piatko, "A visibility matching tone reproduction operator for high dynamic range scenes," *IEEE Trans. on Visualization and Computer Graphics*, vol. 3, no. 4, pp. 291–306, 1997. 10

[21] J. Duan, M. Bressan, C. Dance, and G. Qiu, "Tone-mapping high dynamic range images by novel histogram adjustment," *Pattern Recognition*, vol. 43, no. 5, pp. 1847–1862, 2010. 10

[22] T. Pouli and E. Reinhard, "Progressive histogram reshaping for creative color transfer and tone reproduction," in *Proceedings of the 8th International Symposium on Non-Photorealistic Animation and Rendering*, pp. 81–90, ACM, 2010. 10

[23] J. Tumblin and H. Rushmeier, "Tone reproduction for realistic images," *IEEE Computer Graphics and Applications*, vol. 13, no. 6, pp. 42–48, 1993. 10

[24] J. Stevens and S. S. Stevens, "Brightness function: Effects of adaptation," *JOSA*, vol. 53, no. 3, pp. 375–385, 1963. 10

[25] R. Fattal, D. Lischinski, and M. Werman, "Gradient domain high dynamic range compression," *ACM Transactions on Graphics*, vol. 21, no. 3, pp. 249–256, 2002. 10

[26] F. Durand and J. Dorsey, "Fast bilateral filtering for the display of high-dynamic-range images," *ACM Transactions on Graphics*, vol. 21, no. 3, pp. 257–266, 2002. 12, 29

[27] C. Tomasi and R. Manduchi, "Bilateral filtering for gray and color images," in *6th International Conference on Computer Vision*, pp. 839–846, 1998. 12

[28] H.-T. Chen, T.-L. Liu, and C.-S. Fuh, "Tone reproduction: A perspective from luminance-driven perceptual grouping," *International journal of computer vision*, vol. 65, no. 1-2, pp. 73–96, 2005. 12, 25

[29] Y. Li, L. Sharan, and E. H. Adelson, "Compressing and companding high dynamic range images with subband architectures," *ACM Transactions on Graphics*, vol. 24, no. 3, pp. 836–844, 2005. 12

[30] B. Girod, "What's wrong with mean-squared error?," in *Digital images and human vision*, pp. 207–220, MIT press, 1993. 13

[31] Z. Wang and A. C. Bovik, "A universal image quality index," *IEEE Signal Processing Letters*, vol. 9, no. 3, pp. 81–84, 2002. 13

[32] Z. Wang, E. P. Simoncelli, and A. C. Bovik, "Multiscale structural similarity for image quality assessment," in *IEEE the Thirty-Seventh Asilomar Conference on Signals, Systems and Computers*, vol. 2, pp. 1398–1402, 2003. 13, 18

[33] M. P. Sampat, Z. Wang, S. Gupta, A. C. Bovik, and M. K. Markey, "Complex wavelet structural similarity: A new image similarity index," *IEEE Trans. on Image Processing*, vol. 18, no. 11, pp. 2385–2401, 2009. 13

[34] Z. Wang and Q. Li, "Information content weighting for perceptual image quality assessment," *IEEE Trans. on Image Processing*, vol. 20, no. 5, pp. 1185–1198, 2011. 13

[35] Z. Wang and A. C. Bovik, "Modern image quality assessment," *Synthesis Lectures on Image, Video, and Multimedia Processing*, vol. 2, no. 1, pp. 1–156, 2006. 13

[36] Z. Wang and E. P. Simoncelli, "Maximum differentiation (mad) competition: A methodology for comparing computational models of perceptual quantities," *Journal of Vision*, vol. 8, no. 12, p. 8, 2008. 15

[37] 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 neuroscience*, vol. 8, no. 12, pp. 1690–1697, 2005. 18, 24

[38] G. Ward, "A contrast-based scalefactor for luminance display," *Graphics gems IV*, pp. 415–421, 1994. 24

[39] J. Holm, "Photographic tone and colour reproduction goals," in *Proceedings, CIE Expert Symposium 96: Colour Standards for Image Technology*, pp. 51–56, 1996. 24

[40] N. Salvaggio, L. D. Stroebel, and R. D. Zakia, *Basic photographic materials and processes*. Taylor & Francis, 2008. 25

[41] H. R. Sheikh, Z. Wang, L. Cormack, and A. C. Bovik, "Live image quality assessment database." `http://live.ece.utexas.edu/research/quality`. 25

[42] R. Mantiuk, K. Myszkowski, and H.-P. Seidel, "A perceptual framework for contrast processing of high dynamic range images," *ACM Transactions on Applied Perception*, vol. 3, no. 3, pp. 286–308, 2006. 29

[43] S. N. Pattanaik, J. Tumblin, H. Yee, and D. P. Greenberg, "Time-dependent visual adaptation for fast realistic image display," in *Proceedings of the 27th annual conference on Computer graphics and interactive techniques*, pp. 47–54, ACM Press/Addison-Wesley Publishing Co., 2000. 29

[44] "Open source community. (2007)." `http://qtpfsgui.sourceforge.net/index.php`. 29

[45] P. Ledda, A. Chalmers, T. Troscianko, and H. Seetzen, "Evaluation of tone mapping operators using a high dynamic range display," *ACM Transactions on Graphics*, vol. 24, no. 3, pp. 640–648, 2005. 29

[46] A. Oğuz Akyüz, M. Levent Eksert, and M. Selin Aydin, "An evaluation of image reproduction algorithms for high contrast scenes on large and small screen display devices," *Computers & Graphics*, vol. 37, no. 7, pp. 885–895, 2013. 29

[47] S. P. Boyd and L. Vandenberghe, *Convex optimization*. Cambridge university press, 2004. 37

[48] I. Recommendation, "500-11, methodology for the subjective assessment of the quality of television pictures," *International Telecommunication Union, Geneva, Switzerland*, 2002. 44

[49] D. C. Montgomery and G. C. Runger, *Applied statistics and probability for engineers*. John Wiley & Sons, 2010. 45, 48