



Contents lists available at ScienceDirect

Computer Vision and Image Understanding

journal homepage: www.elsevier.com/locate/cviu

Clustering based content and color adaptive tone mapping

Hui Li^{*,a}, Xixi Jia^{b,a}, Lei Zhang^{*,a}^a Department of Computing, The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong^b School of Mathematics and Statistics, Xidian University, Xi'an, China

ARTICLE INFO

Keywords:

High dynamic range
Tone mapping
Clustering

ABSTRACT

By extracting image luminance channel and separating it into a base layer and a detail layer, the Retinex theory has been widely adopted for tone mapping to visualize high dynamic range (HDR) images on low dynamic range display devices. Many edge-preservation filtering techniques have been proposed to approximate the base layer for Retinex image decomposition; however, the associated tone mapping methods are prone to halo artifacts and false colors because filtering methods are limited in adapting the complex image local structures. We present a statistical clustering based tone mapping method which can more faithfully adapt image local content and colors. We decompose each color patch of the HDR image into three components, *patch mean*, *color variation* and *color structure*, and cluster the patches into a number of clusters. For each cluster, an adaptive subspace can be easily learned by principal component analysis, via which the patches are transformed into a more compact domain for effective tone mapping. Comparing with the popular edge-preservation filtering methods, the proposed clustering based method can better adapt to image local structures and colors by exploiting the image global redundancy. Our experimental results demonstrate that it can produce high-quality image with well-preserved local contrast and vivid color appearance. Furthermore, the proposed method can be extended to multi-scale for more faithful texture preservation, and off-line subspace learning for efficient implementation.

1. Introduction

The dynamic range (i.e., the ratio of maximum to minimum irradiance) of a natural scene is usually very high, approximately 14 orders of magnitude (Duan et al., 2010; Reinhard et al., 2010). However, the generic camera sensors have limited dynamic range, often resulting in under-exposure or over-exposure regions in a captured picture. High dynamic range imaging (HDR) has thus been an important topic in the field of computer vision and computational photography. One widely used strategy to extend the camera dynamic range is to take a sequence of images under different exposures (Ma et al., 2017; 2015b; Wu et al., 2016). With this strategy, there are two categories of approaches to obtain the HDR-like images: multi-exposure image fusion (MEF) (Wu et al., 2016) in image domain, and HDR content reconstruction in radiance domain (Badki et al., 2015; Debevec and Malik, 1997; Mitsunaga and Nayar, 1999).

MEF directly fuses the sequence of images into one image, which is easy to operate but suffers from the problem of ghosting artifacts and the severe dependency on the selection of exposure sequences (Mahmoudabadi et al., 2017; Wu et al., 2016). HDR content reconstruction methods first establish the radiance map by recovering the

camera response function (CRF) and fuse the pixel values in radiance domain (Debevec and Malik, 1997). However, the calculation of CRF is complex and prone to reconstruction errors (Chakrabarti et al., 2014). With the improvement of sensor response sensitivity, high-end cameras can directly generate high-bit raw data without the recovery of CRF.

With the high-bit HDR image available, one important issue is how to display the HDR data. The standard display devices such as LCD, CRT, projectors and printers mostly have a low dynamic range (LDR) and cannot display HDR images directly. To fill in the gap between HDR data and LDR display, techniques have been developed to compress the dynamic range of HDR data for effective display, which are called tone mapping or tone reproduction (Drago et al., 2003; Fattal et al., 2002). A good tone mapping algorithm should faithfully preserve the image detailed features and colors while reducing the irradiance level. In the past two decades, a number of studies have been conducted to develop effective tone mapping algorithms. Generally speaking, the tone mapping methods fall into two primary categories: global tone mapping methods (Drago et al., 2003; Tumblin and Rushmeier, 1993) and local tone mapping methods (Fattal et al., 2002; Reinhard et al., 2002).

Due to the limited computational resources, early studies (Drago et al., 2003; Larson et al., 1997; Tumblin and Rushmeier, 1993; Ward,

^{*} Corresponding author.E-mail addresses: xiaohui102788@126.com (H. Li), xixijia@stu.xidian.edu.cn (X. Jia), cszhang@comp.polyu.edu.hk (L. Zhang).URL: <http://www4.comp.polyu.edu.hk/~cszhang/> (L. Zhang).<http://dx.doi.org/10.1016/j.cviu.2017.11.001>

Received 14 February 2017; Received in revised form 28 June 2017; Accepted 1 November 2017

1077-3142/ © 2017 Elsevier Inc. All rights reserved.

1994) focus on designing simple global tone mapping operators. Tumblin and Rushmeier (1993) proposed a non-linear tone mapping algorithm according to the brightness perception of human visual system. Ward (1994) compressed image contrast instead of absolute luminance using a simple linear compression function. Larson et al. (1997) applied histogram adjustment to tone mapping by preserving the histogram distribution of the original HDR data. The adaptive logarithmic mapping in Drago et al. (2003) compresses the dynamic range with different logarithmic bases. The higher irradiance is compressed via \log_2 , whereas the lower irradiance via \log_{10} , to achieve desirable contrast and detail preserving. Reinhard and Devlin (2005) proposed a simple and practical s curve for global tone mapping in independent channels. The global operators are computationally efficient without halo artifacts. However, the local contrast and visibility of details in the produced LDR images are not satisfactory.

Recent studies focus more on local tone mapping techniques. Fattal et al. (2002) designed a novel local tone mapping operator based on gradient attenuation. They compressed the drastic irradiance changes by reducing the large gradients under a multi-scale framework. Reinhard et al. (2002) classified the dynamic range of display devices into 11 zones according to the different irradiance in HDR data. Li et al. (2005) put forward a multi-resolution image decomposition method using symmetrical analysis-synthesis filter banks for local tone mapping. The gain map of each subband is calculated to alleviate the halo artifacts. Shan et al. (2010) developed a globally local optimization method with a locally linear model, where the guidance map is constructed via local statistical information. Gu et al. (2012) replaced the linear assumption (Shan et al., 2010) with the local non-linear gamma correction. Ma et al. (2015a) designed a tone mapped image quality index (TMQI) and performed dynamic range compression by optimizing this index. Chen et al. (2005) segmented the HDR image into different regions via the earth movers distance (EMD), and applied local tone mapping operation on each component. Ferradans et al. (2011) proposed a two-stage tone mapping method: human visual system based global tone mapping, followed by optimization based local contrast enhancement. Duan et al. (2010) improved the tone mapping performance of Larson et al. (1997) by applying adaptive local histogram adjustment on non-overlapped blocks. In general, local tone mapping methods are spatially adaptive, and can reproduce the local details and contrast well. However, these local operators have higher computational cost and are prone to producing halo artifacts (Li et al., 2005) and ringing effect (Shibata et al., 2016).

In recent years, researchers have been focusing on the design of various edge-preserving filters for tone mapping. The main principle is to decompose an HDR image into a detail layer and a base layer, and impose different operations on the two layers. In particular, the base layer image can be obtained by filtering the HDR data. Tumblin and Turk (1999) made the first attempt to design edge-preserving filters by using anisotropic diffusion to replace Gaussian filtering based on the Retinex theory (Jobson et al., 1997). Durand and Dorsey (2002) developed a fast implementation of bilateral filtering for tone mapping, which can efficiently generate smoothed images while preserving the edges. Based on this framework, many subsequent works (Farbman et al., 2008; Guarnieri et al., 2011; He et al., 2013; Kou et al., 2015; Li and Zheng, 2014; Li et al., 2015; Xu et al., 2011) have been proposed to better remap the HDR data. In Farbman et al. (2008), a weighted least squares based global optimization method was proposed to smooth the HDR data, where a larger weight is given to local details and contours, while a smaller weight is distributed to strong edges. An iterative method was proposed in Guarnieri et al. (2011) to improve the solving of weighted least squares. By minimizing the global gradient of an HDR image, Xu et al. (2011) used the l_0 norm as the regularizer to smooth the HDR image. He et al. (2013) proposed a guided filtering based method for edge preservation. A linear relationship is assumed between the guided image and the image to be filtered to avoid large edge loss. Some works (Kou et al., 2015; Li and Zheng, 2014; Li et al., 2015)

introduce the gradient information as the weight to balance the data term and regularizer term in a local window, which share the similar idea to global weighted least squares.

The luminance edge-preservation filtering based tone mapping algorithms mentioned above can improve the visual quality of tone mapped image; however, the nonlinear filters used by them are not flexible and adaptive enough to fit the various edges and structures in natural images, resulting in halo artifact and false colors. Different from those luminance filtering based methods, in this paper we develop a statistical clustering based tone mapping method to more effectively exploit the image local and global redundancy. We do not separate an image into luminance and chrominance channels to process; instead, we work on image patches, and decompose a color patch into three components: patch mean, color variation and color structure. It is well-known that there exist repetitive patterns/structures in natural images (Dong et al., 2011; Zhang et al., 2010). Based on the color structure component, we group similar patches into clusters, and use statistical signal processing tools such as principal component analysis (PCA) to define a subspace of the patches in a cluster. Consequently, we can project each patch into a more compact domain, where the tone mapping operation can be more effectively performed. Compared with the edge-preservation filtering based methods, our proposed statistical clustering based method is more local content and color adaptive and robust since it exploits the image global redundancy to decompose local structures.

The main contributions of our paper lie in the following aspects. (1) Instead of using the deterministic edge-preserving filters, we leverage statistical clustering methods to better represent the local color structures of HDR images. Each patch will be adaptively processed based on its cluster. (2) We perform tone mapping in the PCA transformed domain other than the intensity domain, where the coefficients have explicit physical meanings and can be more effectively compressed. (3) Different from previous methods which extract luminance channel and perform layer separation on it, we do not extract luminance channel but process image luminance and chrominance information simultaneously.

The rest of our paper is organized as follows. Section 2 presents the proposed method in detail. Section 3 presents extensive experimental results and discussions. Section 4 concludes the paper.

2. Content and color adaptive tone mapping

2.1. The proposed tone mapping framework

Most previous tone mapping methods process luminance and chrominance separately. A typical framework of conventional tone mapping methods is shown in Fig. 1(a). Given an HDR image in RGB format, the luminance channel is first extracted as $L = 0.2126 \cdot R + 0.7152 \cdot G + 0.0722 \cdot B$ for the XYZ color space (Fattal et al., 2002), or $L = 0.299 \cdot R + 0.587 \cdot G + 0.114 \cdot B$ for the YUV color space (Li and Zheng, 2014). In some literature (Gu et al., 2013), the average of R, G, B channels $L = 1/3(R + G + B)$ is employed as the luminance. After dynamic range compression on luminance, the chrominance is processed based on the compressed luminance to reproduce the tone mapped image. The widely used color processing operation is $C_{out} = \left(\frac{C_{in}}{L_{in}}\right)^s \cdot L_{out}$, where C represents the chrominance channel, L_{in} and L_{out} denote the luminance before and after HDR processing, and s adjusts the color saturation of the tone mapped image. The empirical value of s is between 0.5 and 0.9 (Gu et al., 2013).

In our proposed method, we do not separate image into luminance and chrominance channels to process. Instead, we propose a very different approach, whose framework is shown in Fig. 1(b). We partition the input RGB image into overlapped color patches, and decompose each patch into three nearly uncorrelated components. The color patches are clustered into a number of clusters, and statistical analysis is used to compress each HDR patch to an LDR one. The flowchart of the

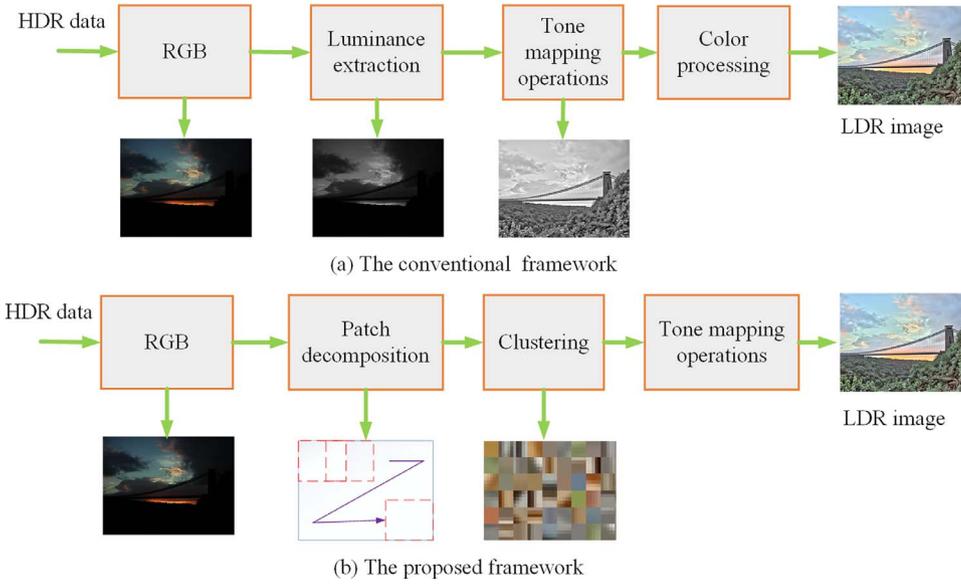


Fig. 1. (a) The traditional tone mapping framework and (b) our proposed framework.

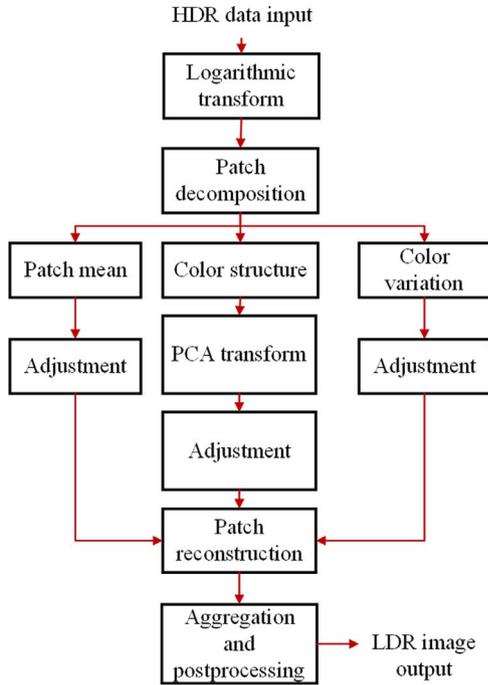


Fig. 2. Flow chart of the proposed tone mapping method.

proposed method is shown in Fig. 2. The main procedures of the proposed method include: logarithmic transform, patch decomposition, clustering and PCA transform, range adjustment, patch reconstruction, aggregation and post-processing. The details of the proposed method are presented in the following.

2.2. Patch decomposition

Like in many existing tone mapping methods (Duan et al., 2010; Gu et al., 2013; Meylan and Susstrunk, 2006), our method needs a simple global tone curve for initialization. Considering the characteristics of human visual system, the logarithmic function is used to this end:

$$\mathbf{L}(i, j, c) = \log(\mathbf{I}(i, j, c) \cdot 10^6 + 1) \quad (1)$$

where \mathbf{I} is the input HDR image, (i, j) refers to the spatial location, and $c \in \{r, g, b\}$ represents the R , G , and B channels. We then apply patch

decomposition to \mathbf{L} . We partition the HDR image \mathbf{L} into many overlapped patches (e.g., of size 7×7) with stride q (e.g., $q = 2$ in our implementation). Denoted by \mathbf{x} an extracted color patch and by \mathbf{x}_c the patch in channel R , G or B . The local mean of each channel \mathbf{x}_c , denoted by m_c , is calculated by averaging all pixels in \mathbf{x}_c . We then subtract the mean from \mathbf{x}_c :

$$\bar{\mathbf{x}}_c = \mathbf{x}_c - \mathbf{1} \cdot m_c \quad (2)$$

where $\mathbf{1}$ is a vector with all elements being 1 and it has the same size as \mathbf{x}_c . One can see that $\bar{\mathbf{x}}_c$ contains the direct current (DC) removed detail structure of \mathbf{x}_c .

The mean m_c is a scalar representing the DC amount of patch \mathbf{x} in channel c . The variation of m_c across channels can reflect the color appearance in that patch. For example, if all the three values of m_c are the same, that patch will be a gray level patch. We can calculate the color variation across channels as:

$$\bar{m}_c = m_c - m \quad (3)$$

where $m = (m_r + m_g + m_b)/3$ is the average of the three m_c . Clearly, m is the average of all pixels in the color patch \mathbf{x} .

With the m , \bar{m}_c , and $\bar{\mathbf{x}}_c$ defined above, for each patch we can decompose it into three components:

$$\begin{aligned} \mathbf{x} &= \begin{bmatrix} \bar{\mathbf{x}}_r \\ \bar{\mathbf{x}}_g \\ \bar{\mathbf{x}}_b \end{bmatrix} + \begin{bmatrix} \mathbf{1} \cdot m_r \\ \mathbf{1} \cdot m_g \\ \mathbf{1} \cdot m_b \end{bmatrix} + \begin{bmatrix} \mathbf{1} \\ \mathbf{1} \\ \mathbf{1} \end{bmatrix} \cdot m \\ &= \bar{\mathbf{x}} + \bar{\mathbf{m}} + [\mathbf{1}; \mathbf{1}; \mathbf{1}] \cdot m \end{aligned} \quad (4)$$

We call the 1st component $\bar{\mathbf{x}} = [\bar{\mathbf{x}}_r; \bar{\mathbf{x}}_g; \bar{\mathbf{x}}_b]$ the color structure since it preserved the detailed local structural information in the three channels, the 2nd component $\bar{\mathbf{m}} = [\mathbf{1} \cdot m_r; \mathbf{1} \cdot m_g; \mathbf{1} \cdot m_b]$ the color variation since it reflects the color differences across three channels, and the 3rd component m the patch mean since it is the mean value of all pixels in the three channels.

2.3. Clustering and PCA transform learning

Given an input HDR image, a large number of patches \mathbf{x} will be extracted. For example, we extract $185,754 \times 7 \times 7$ patches with stride 2 for an image of size 1000×750 . It has been widely accepted that there will be many patches sharing a similar structure in an image (Dong et al., 2011; Xu et al., 2015; Zhang et al., 2010). After removing the DC component, some patches with different intensity levels may also have similar structure. Therefore, we can cluster the patches into different

clusters based on the color structure component $\bar{\mathbf{x}}$. The classical clustering methods such as K-means (Zhang et al., 2010) and Gaussian Mixture Model (GMM) (Xu et al., 2015) can be used to this end. We choose K-means because it has much lower computational cost while leading to similar tone mapping results to GMM based on our experiments. We stretch each $\bar{\mathbf{x}}$ to a vector, and apply K-means clustering to the vectorized color structure components $\bar{\mathbf{x}}$ (note that $\bar{\mathbf{x}}$ contains the detailed features from all the R, G and B channels). Suppose that K clusters are obtained. For each cluster, we calculate the covariance matrix of the vectors $\bar{\mathbf{x}}$ within it, denoted by Φ . Since the covariance matrix Φ is positive semidefinite, we can have its eigenvalue decomposition as:

$$\Phi = \mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^{-1} \quad (5)$$

where the orthogonal matrix \mathbf{Q} is composed of the eigenvectors of Φ . The so-called principal component analysis (PCA) transform matrix can be easily obtained as (Zhang et al., 2010):

$$\mathbf{P} = \mathbf{Q}^T \quad (6)$$

Since the patches in one cluster are similar in structure, the eigenvectors associated with the first a few largest eigenvalues will be able to represent the most important common structures in that cluster (i.e., the principal components). With the PCA transform matrix \mathbf{P} , for each patch $\bar{\mathbf{x}}$ within that cluster, we can transform it into the PCA domain as:

$$\bar{\mathbf{y}} = \mathbf{P}\bar{\mathbf{x}} \quad (7)$$

Note that the coefficients in $\bar{\mathbf{y}}$ will be much sparser than those in $\bar{\mathbf{x}}$. The small coefficients correspond to noise interference and trivial structures. The modest coefficients correspond to image fine-scale details. The large coefficients correspond to image principle structures. Usually, only the first a few coefficients in $\bar{\mathbf{y}}$ will be significant, while the remaining being close to zero. Therefore, compressing the dynamic range of $\bar{\mathbf{y}}$ will be much easier and more robust than that of $\bar{\mathbf{x}}$. This is one of the essential reasons that why our method works for tone mapping.

2.4. Dynamic range adjustment and patch reconstruction

To achieve the tone mapping of patch \mathbf{x} , we need to adjust the values of m , $\bar{\mathbf{m}}$, and $\bar{\mathbf{x}}$. For component $\bar{\mathbf{x}}$, we transform it into the PCA domain via Eq. (7) and process $\bar{\mathbf{y}}$. The smallest coefficients in $\bar{\mathbf{y}}$ are usually produced by the trivial structures, fluctuations and even noise in $\bar{\mathbf{x}}$, and therefore we first remove them for a more stable tone mapping. Denote by \max the maximal absolute value of all coefficients in $\bar{\mathbf{y}}$. Since noise mostly corresponds to the smallest PCA coefficients, a simple empirical threshold is good enough to suppress the noise. In order to keep the details of the original data as much as possible while removing noise, a small threshold is empirically selected. We set those coefficients whose absolute value is smaller than $0.1 \max$ to 0.

For the task of tone mapping, the large PCA coefficients (corresponding to image large scale structures) in $\bar{\mathbf{y}}$ should be compressed, while the smaller coefficients (corresponding to image fine scale textures) should be maintained or enhanced slightly. To this end, an s-shaped curve could be employed to adjust the coefficients. The commonly used s-shaped curves include *arctan* and *sigmod* functions. We choose the *arctan* function to adjust coefficients because it exhibits stronger transition ability in both shadows and highlights, and the adjusting function should be symmetrical to 0 to process the negative coefficients in the PCA transform domain. With the *arctan* function, we adjust the coefficients in $\bar{\mathbf{y}}$ as:

$$\bar{\mathbf{y}}_a = (1.6/\pi) \cdot \arctan(a \cdot \bar{\mathbf{y}}) \quad (8)$$

where a is a parameter to control the shape of the curve. Some example curves are plotted in Fig. 3. One can see that the smaller the a is, the stronger compression effect on $\bar{\mathbf{y}}$ will be.

For the color variation component $\bar{\mathbf{m}}$, we also use the *arctan* function but with a different parameter to adjust it:

$$\bar{\mathbf{m}}_b = (1.2/\pi) \cdot \arctan(b \cdot \bar{\mathbf{m}}) \quad (9)$$

where b is the shape parameter. The patch mean component m changes slowly, which can be linearly compressed by multiplying a weight w . After range adjustment on m , $\bar{\mathbf{m}}$ and $\bar{\mathbf{y}}$, the tone mapped patch of \mathbf{x} , denoted by \mathbf{x}_t , can be reconstructed as

$$\mathbf{x}_t = \mathbf{P}^T \bar{\mathbf{y}}_a + \bar{\mathbf{m}}_b + [\mathbf{1}; \mathbf{1}; \mathbf{1}]w \cdot m \quad (10)$$

where w is a scalar ranging from 0 to 1.

2.5. Aggregation and post-processing

The operations described in Sections 2.3 and 2.4 are applied to each extracted patch for the input HDR image, and aggregation of the processed patches is needed to reconstruct the tone mapped LDR image. Each tone mapped patch is put back to its original location, while the overlapped pixels in adjacent patches are averaged. In the post-processing stage, the 1% pixels of lowest and highest values are clamped to enhance the primary contrast. Finally, every patch pixel is linearly stretched to 0 – 1 to fully take advantage of the dynamic range of target display device to show the result.

2.6. Extension to multi-scales

In the proposed patch clustering based tone mapping method, each patch will have a mean component (scalar value). The means of all patches will form a smoothed gray level image of the original image. Fig. 4 shows an example. Fig. 4(a) is the original image (the tone mapped image is shown here for better visibility), and Fig. 4(b) is the mean image after patch decomposition. Note that the resolution of mean image is 1/4 of that of the original image because we use a stride factor of 2 (in both horizontal and vertical directions) to extract the patches (size: $7 \times 7 \times 3$).

One can see that there is still certain amount of textures in the mean image. If we compress the mean image by a weight w as shown in Eq. (10), some detailed texture information can be lost in the final tone mapped image. To solve this problem, we could extend the proposed method to multi-scales. More specifically, we extract patches from the mean image, and decompose each patch into two components: patch mean and patch structure. The patch mean is the average of all pixels in a patch, while the patch structure component is obtained by subtracting the mean from the patch. Note that we do not have a color variation component here since the mean image is gray scale. The clustering and PCA transform can then be applied to the patch structure components. By embedding such operations into the framework in Fig. 2, we could have a two-scale implementation of the proposed method, which is illustrated in Fig. 5.

Our method can be easily extended to more scales by further decomposing the mean image generated on the 2nd scale. Nonetheless, our experiments show that a 2-scale decomposition is enough for most of our test images. In Fig. 6(a) and (b), we show the single-scale and two-scale tone mapping results by our method. One can see that some detailed structures of the cloud region are lost in the single-scale result image, but they can be preserved in the two-scale result image. In addition, since the mean image is gray scale and has a lower resolution, the two-scale decomposition scheme has similar implementation time to the single-scale scheme.

2.7. Offline PCA transform learning

The color structure clustering step is the most time-consuming part in our proposed method. With the K-means clustering algorithm, it will take about 147 s to process an image of size $1000 \times 750 \times 3$ (patch size: $7 \times 7 \times 3$) under the MATLAB R2014a programming environment on a PC equipped with an i7-4790K CPU, 4G HZ and 32GB memory.

To reduce the computational cost, we can pre-calculate the clusters

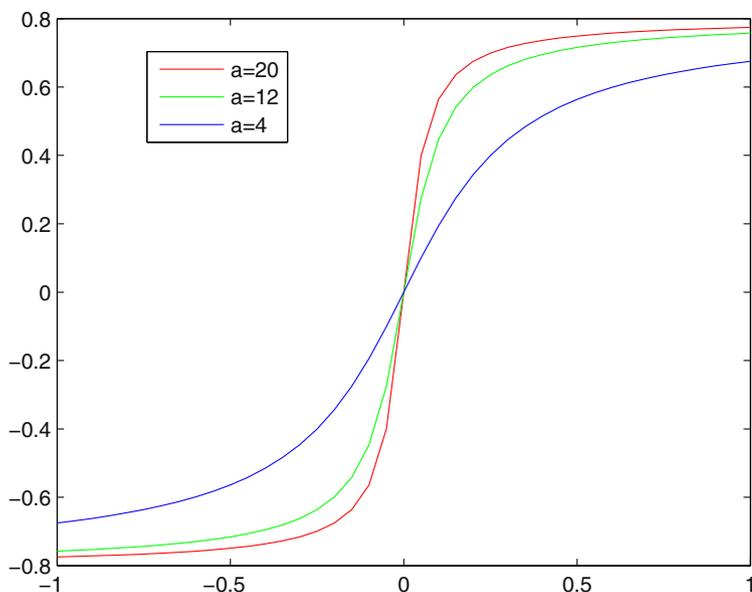


Fig. 3. The arctan function in Eq. (8) with different parameters.

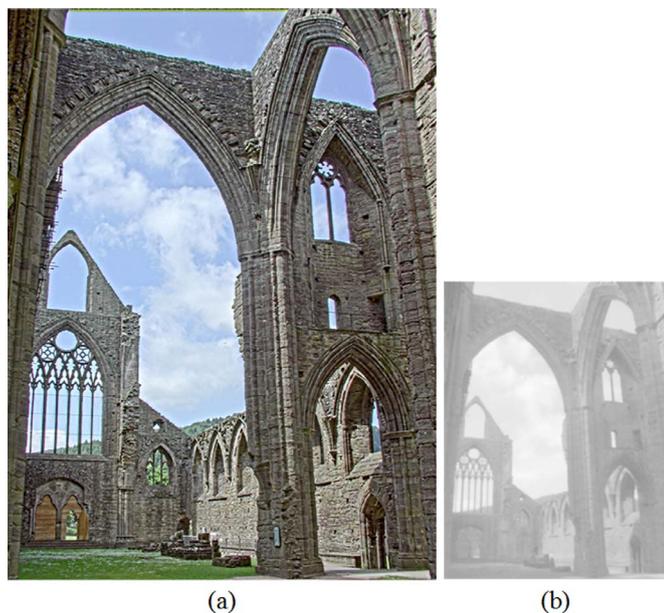


Fig. 4. (a) The original HDR image (the tone mapped image is shown here for better visibility). (b) The mean image formed by the patch means.

and their PCA transform matrices using an external dataset, as illustrated in Fig. 7. We use the Kodak database¹ as the training dataset. About 300,000 patches (patch size: $7 \times 7 \times 3$) are extracted and their color structure components are computed for clustering. For each cluster, we have a cluster mean and its PCA transform matrix. In the test stage, for each patch of the input HDR image, we determine its corresponding cluster based on the minimum Euclidean distance between its color structure component and the centroids of clusters. Then the PCA transform matrix of that cluster is used to process that patch. Without the online clustering, the running time of our method is significantly improved. On average, it costs about 7 seconds to process an image of size $1000 \times 750 \times 3$, about 21 times faster than the online version of our method. In Fig. 6(c) and (d), we show the single-scale and two-scale tone mapping results by our offline method. We can see that the offline

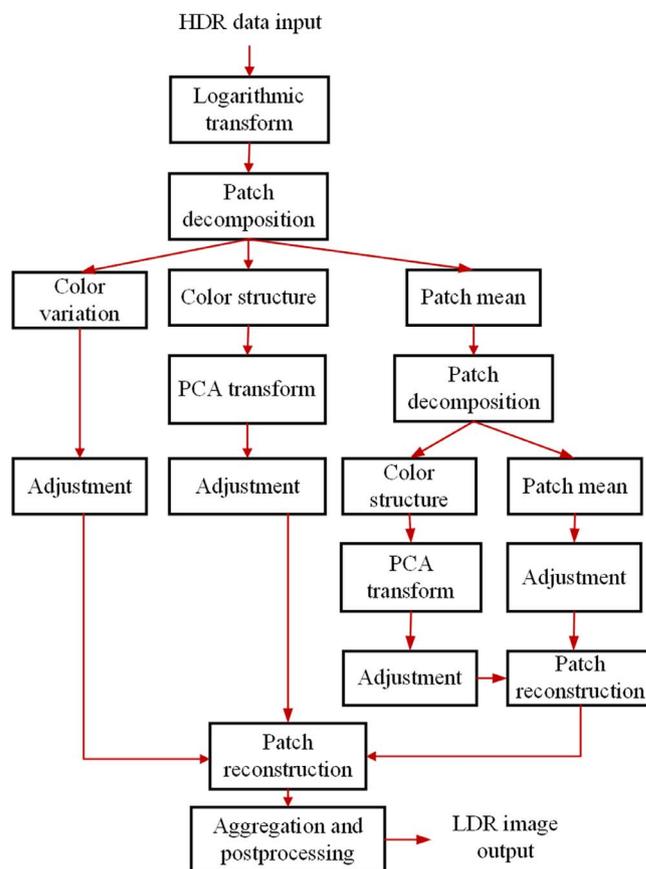


Fig. 5. The two-scale implementation flow chart of the proposed method.

method achieves similar tone mapping results to the online method in terms of objective assessment (See Tables 2 and 3.)

3. Experimental results and discussions

3.1. Implementation details

Our method is a patch based approach, and we need to fix the patch size first. Based on our experimental experience, setting the patch size

¹ <http://r0k.us/graphics/kodak/>.

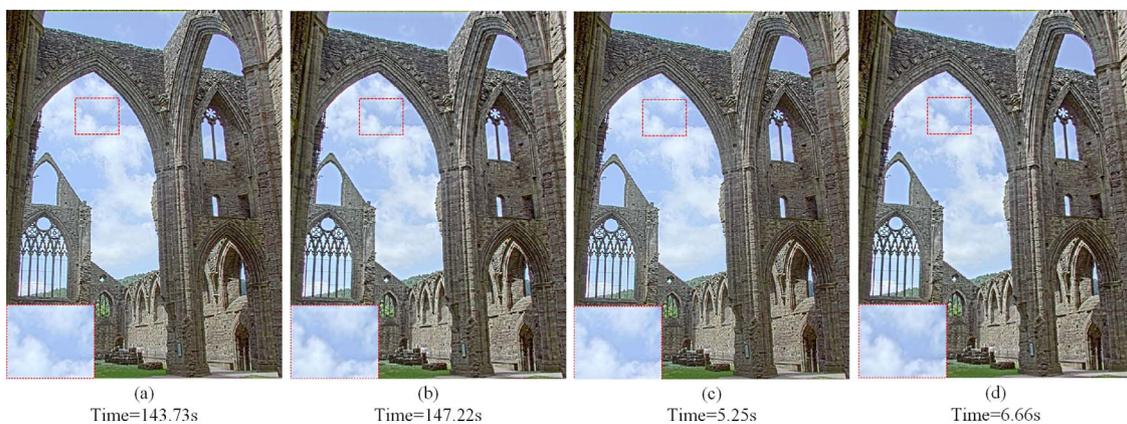


Fig. 6. (a) and (b) are the tone mapped images by single-scale and two-scale decompositions, respectively, and (c) and (d) are the single-scale and two-scale results by off-line pre-learning of the PCA transforms.

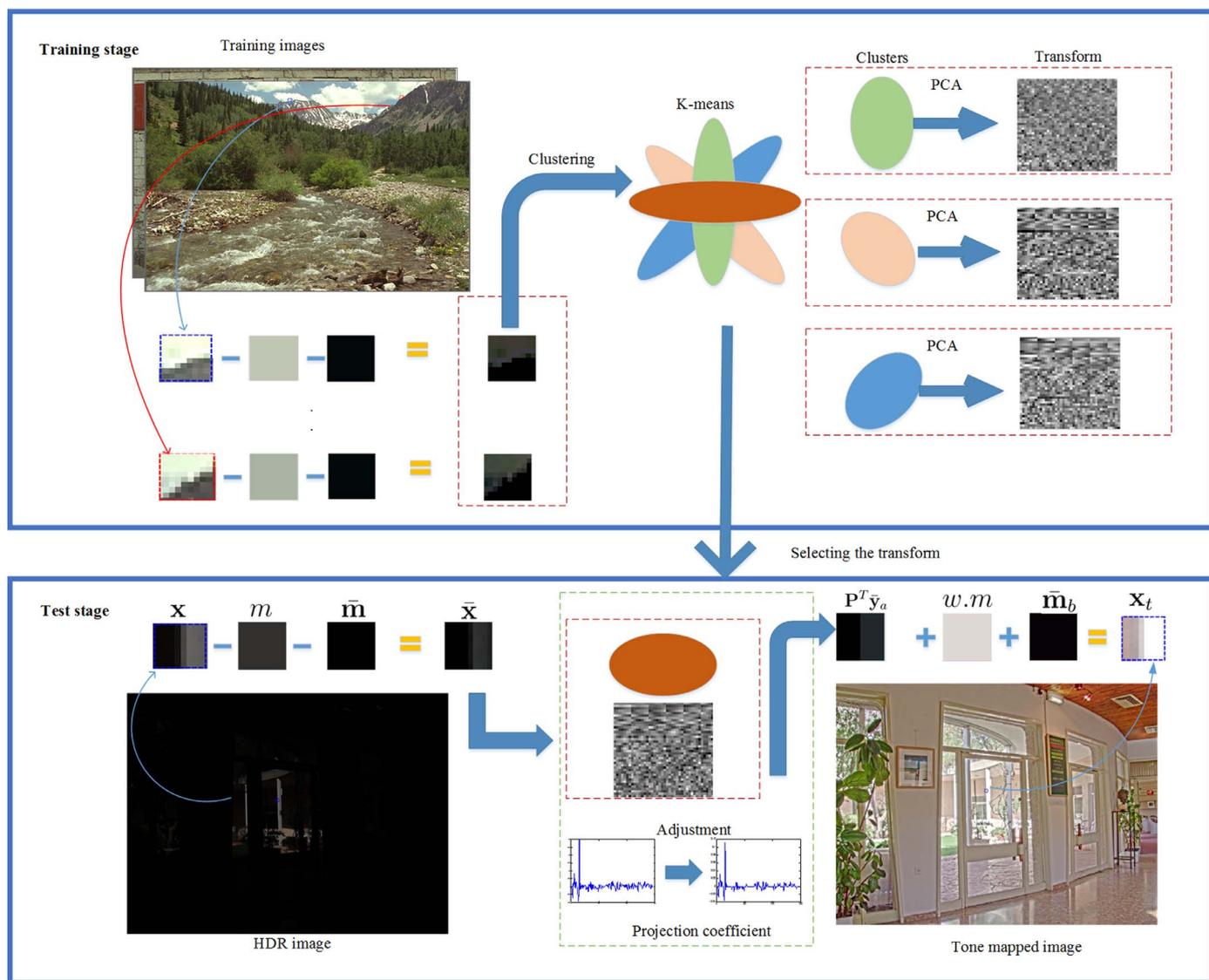


Fig. 7. Top box: the offline patch clustering and PCA transform learning by using an external dataset. Bottom box: the online cluster selection and tone mapping.

from $5 \times 5 \times 3$ to $8 \times 8 \times 3$ will lead to similar results, and we set the patch size to $7 \times 7 \times 3$ in all our experiments. We extract the patches from an image with stride 2 in both horizontal and vertical directions. For clustering, we use the K-means algorithm (Dong et al., 2011; Zhang et al., 2010) with initial cluster number 100 for scale 1 and 50 for scale

2. Note that some small clusters will be merged in the clustering process so that the final number of clusters will be less than 100 and 50 on the two scales. For our offline clustering method, the final numbers of clusters are 83 (scale 1) and 13 (scale 2), respectively.

The parameter a in Eq. (8) controls the adjustment of local



Fig. 8. The impact of parameter a on the reconstruction of image local structure.



Fig. 9. The impact of parameter b on the reconstruction of local color appearance.

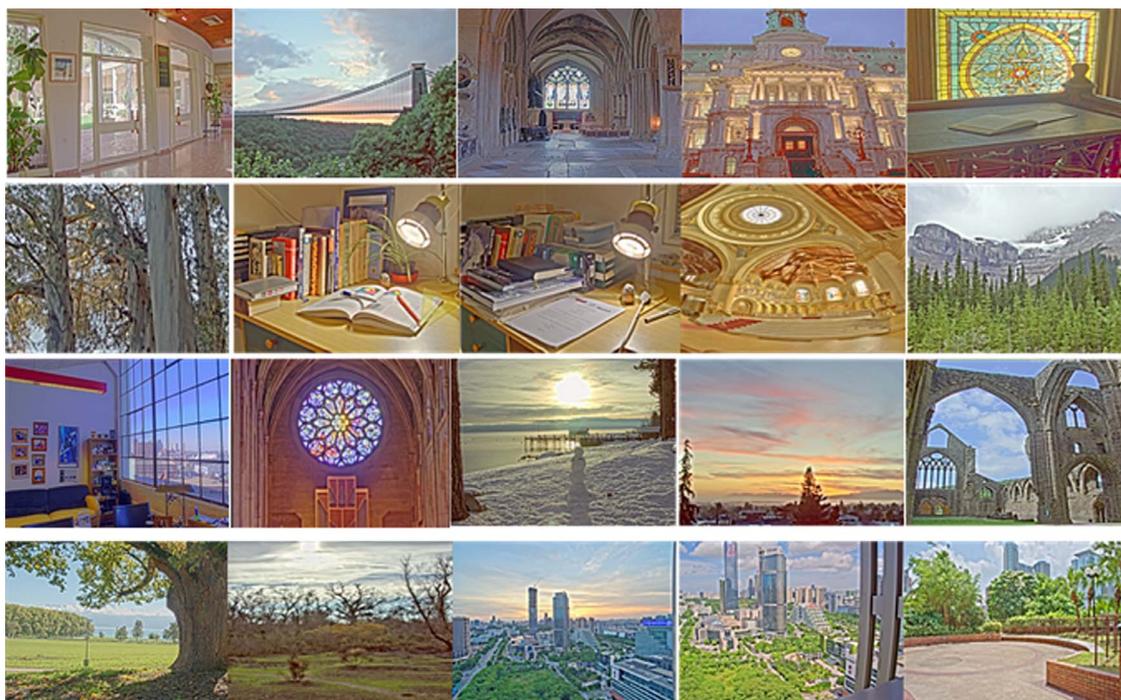


Fig. 10. Source image scenes used in our experiment. The HDR data are represented by the tone mapped results for better visualization.

Table 1

Average execution time in seconds on 5 scenes of size $713 \times 535 \times 3$.

Alg	Drago	Fattal	Kuang	Frabman	Shan	Shibata	Our
Env	MATLAB	MATLAB	MATLAB	MATLAB	MATLAB	MATLAB	MATLAB
Time (s)	0.13	1.15	1.23	2.89	10.28	15.01	3.86

structures. For simplicity, we set a the same for both the two scales. Fig. 8 shows the tone mapping results by letting a be 2, 6, 10, 20, respectively. We can see that a bigger a will make the local contrast stronger, but a too big a will make local structures and colors unnatural. We choose $a = 6$ in our experiment to achieve a good balance between

contrast enhancement and color/structure preservation. The parameter b in Eq. (9) controls the adjustment of local color appearance. Fig. 9 shows the tone mapping results by letting b be 2, 4, 8, 16, respectively. We can see that a too big b will lead to over-saturation, while a too small b will lead to under-saturation. We choose $b = 4$ in our

Table 2
The TMQI scores of the tone mapping images.

Source	Log	Exp	Mantiuk et al. (2008)	Drago et al. (2003)	Fattal et al. (2002)	Kuang et al. (2007)	Farbman et al. (2008)	Shan et al. (2010)	Shibata et al. (2016)	Ours1	Ours2	Ours3
1	0.550	0.866	0.889	0.907	0.701	0.823	0.918	0.843	0.866	0.934	0.939	0.940
2	0.698	0.748	0.867	0.842	0.844	0.928	0.727	0.805	0.908	0.871	0.872	0.876
3	0.576	0.886	0.936	0.896	0.818	0.855	0.974	0.973	0.850	0.933	0.939	0.938
4	0.630	0.810	0.912	0.951	0.790	0.911	0.844	0.835	0.774	0.857	0.856	0.869
5	0.709	0.922	0.699	0.960	0.808	0.817	0.844	0.905	0.781	0.944	0.934	0.932
6	0.807	0.813	0.852	0.958	0.885	0.880	0.821	0.933	0.771	0.809	0.811	0.816
7	0.759	0.910	0.916	0.921	0.758	0.843	0.844	0.936	0.822	0.959	0.954	0.952
8	0.740	0.895	0.874	0.870	0.752	0.824	0.807	0.919	0.813	0.945	0.949	0.946
9	0.639	0.856	0.896	0.948	0.769	0.832	0.807	0.921	0.755	0.909	0.898	0.906
10	0.729	0.887	0.890	0.870	0.880	0.962	0.719	0.960	0.930	0.890	0.884	0.885
11	0.706	0.882	0.923	0.952	0.854	0.952	0.882	0.804	0.838	0.943	0.933	0.916
12	0.536	0.938	0.888	0.938	0.748	0.796	0.776	0.869	0.781	0.892	0.887	0.885
13	0.534	0.854	0.915	0.944	0.786	0.790	0.928	0.950	0.839	0.952	0.954	0.953
14	0.828	0.782	0.878	0.788	0.803	0.867	0.521	0.789	0.844	0.812	0.817	0.817
15	0.773	0.913	0.908	0.901	0.903	0.986	0.741	0.886	0.890	0.909	0.899	0.898
16	0.777	0.868	0.949	0.933	0.953	0.946	0.774	0.905	0.823	0.930	0.828	0.835
17	0.667	0.879	0.883	0.906	0.912	0.979	0.718	0.957	0.908	0.953	0.947	0.945
18	0.773	0.817	0.921	0.873	0.804	0.957	0.662	0.897	0.922	0.919	0.919	0.915
19	0.800	0.862	0.960	0.881	0.936	0.939	0.728	0.909	0.851	0.846	0.845	0.845
20	0.790	0.880	0.976	0.916	0.941	0.952	0.751	0.920	0.859	0.848	0.847	0.847
Average	0.701	0.863	0.897	0.908	0.832	0.892	0.790	0.896	0.841	0.903	0.896	0.896

Table 3
The FSITM scores of the tone mapping images.

Source	Log	Exp	Mantiuk et al. (2008)	Drago et al. (2003)	Fattal et al. (2002)	Kuang et al. (2007)	Farbman et al. (2008)	Shan et al. (2010)	Shibata et al. (2016)	Ours1	Ours2	Ours3
1	0.863	0.783	0.852	0.829	0.821	0.857	0.848	0.774	0.779	0.802	0.804	0.803
2	0.756	0.779	0.849	0.854	0.736	0.868	0.771	0.809	0.844	0.852	0.853	0.855
3	0.927	0.857	0.903	0.896	0.861	0.897	0.894	0.868	0.823	0.830	0.839	0.838
4	0.696	0.846	0.901	0.915	0.732	0.917	0.901	0.842	0.869	0.901	0.901	0.902
5	0.795	0.726	0.719	0.790	0.748	0.823	0.801	0.715	0.759	0.792	0.791	0.792
6	0.922	0.879	0.932	0.951	0.783	0.948	0.916	0.856	0.905	0.930	0.931	0.932
7	0.811	0.808	0.866	0.869	0.723	0.878	0.871	0.831	0.826	0.872	0.872	0.873
8	0.711	0.802	0.855	0.860	0.717	0.872	0.861	0.832	0.826	0.864	0.864	0.864
9	0.803	0.827	0.913	0.923	0.757	0.932	0.884	0.877	0.884	0.920	0.920	0.921
10	0.863	0.868	0.904	0.924	0.760	0.924	0.797	0.895	0.896	0.906	0.907	0.908
11	0.735	0.757	0.838	0.846	0.724	0.831	0.732	0.732	0.807	0.842	0.846	0.852
12	0.606	0.803	0.861	0.875	0.750	0.888	0.857	0.838	0.853	0.878	0.880	0.881
13	0.797	0.748	0.819	0.800	0.811	0.837	0.821	0.774	0.780	0.785	0.788	0.788
14	0.818	0.778	0.834	0.838	0.743	0.861	0.602	0.792	0.845	0.840	0.840	0.846
15	0.906	0.817	0.884	0.862	0.799	0.872	0.777	0.753	0.837	0.833	0.835	0.835
16	0.803	0.802	0.914	0.926	0.762	0.930	0.882	0.835	0.827	0.921	0.921	0.927
17	0.796	0.835	0.907	0.914	0.749	0.926	0.804	0.890	0.898	0.908	0.909	0.909
18	0.825	0.795	0.843	0.852	0.747	0.874	0.785	0.837	0.827	0.855	0.858	0.861
19	0.864	0.856	0.874	0.902	0.755	0.889	0.785	0.837	0.866	0.882	0.886	0.889
20	0.870	0.880	0.915	0.913	0.764	0.903	0.810	0.835	0.887	0.906	0.906	0.909
Average	0.808	0.809	0.869	0.877	0.762	0.886	0.823	0.820	0.845	0.866	0.868	0.870

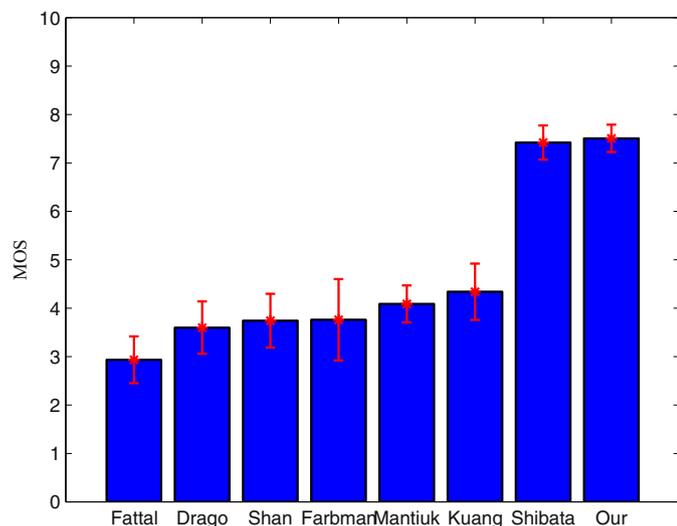


Fig. 11. Mean and std of subjective rankings of the 8 competing tone mapping algorithms.

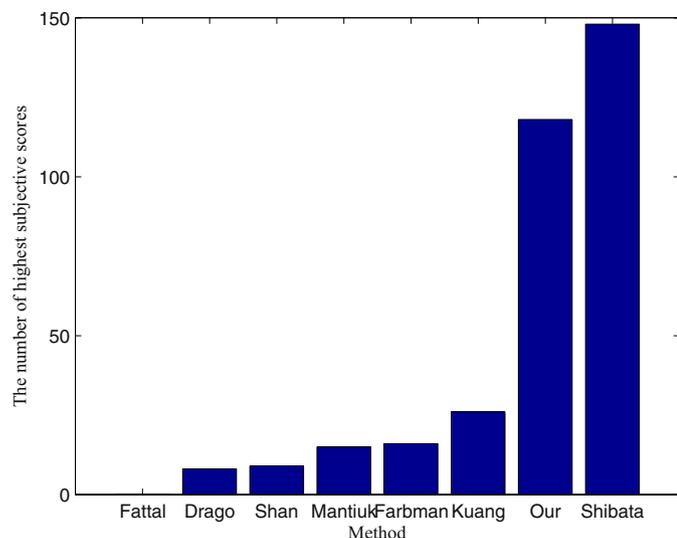


Fig. 12. The number of highest subjective scores obtained by different methods.

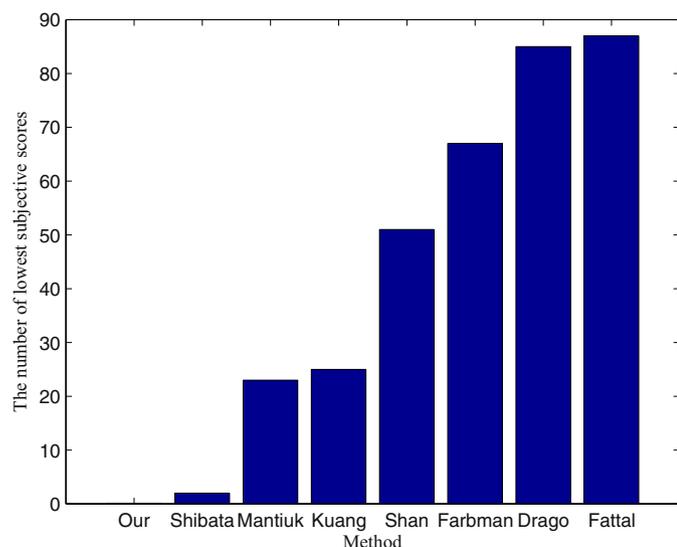


Fig. 13. The number of lowest subjective scores obtained by different methods.

experiments.

Finally, the parameter $w \in [0, 1]$ in Eq. (10) is used to adjust the luminance of the tone mapped image. Clearly, the image luminance will be lower with a smaller w . We set $w = 0.8$ based on experimental experience.

3.2. Test data and comparison algorithms

To verify the effectiveness of the proposed method, we collect 17 sets of widely used HDR image data from links^{2,3,4} and capture 3 sets of HDR image data by two digital cameras (Sony a7 and DJI Phantom3). The scenes of the 20 sets of HDR images are shown in Fig. 10. These 20 images cover both outdoor and indoor scenes, as well as different objects such as trees, sky, sun, cloud, books, and windows.

We compare our algorithm with 7 representative tone mapping algorithms, including “Mantiuk” (Mantiuk et al., 2008), “Drago” (Drago et al., 2003), “Fattal” (Fattal et al., 2002), “Kuang” (Kuang et al., 2007), “Farbman” (Farbman et al., 2008), “Shan” (Shan et al., 2010), and “Shibata” (Shibata et al., 2016). The source codes of these comparison methods are publicly available in the “HDR-Toolbox” (Banterle et al., 2011) or provided in the authors’ homepages^{5,6,7}. We use the default parameters of those codes which were optimized by the authors. The running time of competing algorithms are summarized in Table 1, from which we can see that our two-scale offline method is slower than “Drago” (Drago et al., 2003), “Fattal” (Fattal et al., 2002), “Kuang” (Kuang et al., 2007), and “Farbman” (Farbman et al., 2008), but faster than “Shan” (Shan et al., 2010), and “Shibata” (Shibata et al., 2016). Since “Mantiuk” (Mantiuk et al., 2008) et al.’s method is implemented by HDR Luminance⁸, we do not report it in running time comparison.

3.3. Objective evaluation

Since there is not a groundtruth LDR image for the HDR data, classical objective measures such as PSNR cannot be used to evaluate the quality of tone mapped images and the performance of a tone mapping algorithm. Recently, researchers have proposed some objective measures (Aydin et al., 2008; Gu et al., 2016; Kundu et al., 2017; Nafchi et al., 2015; Song et al., 2016; Yeganeh and Wang, 2013) to evaluate the tone mapping results. The objective metrics TMQI (Yeganeh and Wang, 2013) and FSITM (Nafchi et al., 2015) are employed in our manuscript and they are based on structural similarity (SSIM) (Wang et al., 2004) and feature similarity (FSIM) (Zhang et al., 2011). TMQI combines SSIM-motivated structural fidelity with statistical naturalness to assess the tone mapped images. FSITM measures local phase similarity of the original HDR and the tone mapped LDR image. Apart from the 7 representative methods (Drago et al., 2003; Farbman et al., 2008; Fattal et al., 2002; Kuang et al., 2007; Mantiuk et al., 2008; Shan et al., 2010; Shibata et al., 2016), we also list the results of the baseline Log and Exp operators in the “HDR-toolbox” (Banterle et al., 2011). The TMQI and FSITM results are shown in Tables 2 and 3, respectively, where Ours1, Ours2 and Ours3 represent the single-scale, two-scale and the off-line two-scale implementations of our method. For each image, the best result is highlighted in bold face.

3.4. Subjective comparison

Let’s then present some visual comparisons of the competing

² <http://www.ok.ctrl.titech.ac.jp/res/IC/ProxPoisson/ProxPoisson.html>.

³ <http://cadik.posvete.cz/tmo/>.

⁴ <https://people.csail.mit.edu/sparis/publi/2011/siggraph/>.

⁵ http://www.cse.cuhk.edu.hk/leo/jia/programs/optimize_tone_mapping_code.zip.

⁶ <http://www.cs.huji.ac.il/~danix/epd/>.

⁷ <http://www.ok.ctrl.titech.ac.jp/res/IC/ProxPoisson/ProxPoisson.html>.

⁸ <http://qtptfsgui.sourceforge.net/>.



Fig. 14. The tone mapping results on image 7 (refer to Fig. 10) by competing tone mapping operators. From (a) to (h): results by “Mantiuk” (Mantiuk et al., 2008), “Drago” (Drago et al., 2003), “Fattal” (Fattal et al., 2002), “Kuang” (Kuang et al., 2007), “Farbman” (Farbman et al., 2008), “Shan” (Shan et al., 2010), “Shibata” (Shibata et al., 2016), and ours. From (i) to (p): the close-ups of (a)–(h). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article).

methods. For our method, we present the results by the offline two-scale implementation. Figs. 14–17 show the tone mapped images of scenes 7, 9, 17, and 18 (see Fig. 10), respectively.

The results by “Mantiuk” (Mantiuk et al., 2008) present the loss of details especially in dark regions. For example, in the close-up images in Figs. 14 and 17, the books and trees cannot be seen. The adaptive global method “Drago” (Drago et al., 2003) presents better results, but it suffers from the loss of local contrast. One can see from Fig. 16 that the contrast of tree branches and cloud background is low. Fattal et al.s method (Fattal et al., 2002) has the problem of detail and contrast loss such as the wall in Fig. 14 and green tree in Fig. 17. Kuang et al.s method (Kuang et al., 2007) shows much distortion of color appearance, although it preserves well local details and contrasts. For instance, it produces a purple color of sky in Fig. 16, which is not natural. The tone mapped images by multi-scale decomposition based method “Farbman” (Farbman et al., 2008) suffer from information loss in some regions, such as the sky in Figs. 16 and 17. Shan et al.s method (Shan et al., 2010) over-smooths much the image local textures. There are neither clear contours of the cloud in Fig. 16 nor fine structures of tree leaves in Fig. 17. Shibata et al.s method (Shibata et al., 2016) shows good local contrast but meanwhile generates much visual artifacts. The surfaces of the wall and desk in Fig. 14 and the roofs in Fig. 15 are over-exaggerated.

Compared with the above methods, our method demonstrates competitive visual quality with good local structure preservation and color reproduction. For instance, in Fig. 14 the local details and contrast labeled in the red box can be seen clearly with decent overall visual

effect. Furthermore, the colors of trees, cloud and grass look natural and saturated. This is mainly because our method clusters image patches based on their local colors and structures and it processes each patch adaptively based on the color and structure statistical information in that cluster.

3.5. Subjective study

A formal subjective study is conducted to further evaluate the proposed tone mapper and compared methods. The subjective testing was operated in an indoor environment with stable illumination. We adopted the strategy in Ma et al. (2015b) in our subjective testing. The tone mapped images of 20 scenes by 8 representative algorithms are shown on a PA328 Display, 32 inch (7680*4320), controlled by a Mac Pro-with Intel Core i5 2.9 GHz CPU. A total number of 17 volunteer subjects, including 8 females and 9 males, were asked to give an integer score ranging from 1 to 10 to each image shown on the display, where 1 means the worst visual quality and 10 means the best visual quality. The mean and std of mean opinion score (MOS) values are shown in Fig. 11. It can be seen that our method and Shibata et al.s method have much better performance than other competing methods. The MOS of our method is 7.50 with std 0.56, while that of Shibata et al.s method is 7.42 with std 0.71. In the subjective experiments, our method obtains 118 highest subjective scores and 0 lowest subjective score among 340 highest and lowest scores. The distributions of numbers of highest and lowest scores by different methods are shown in Figs. 12 and 13. Overall, our method demonstrates highly competitive and stable tone



Fig. 15. The tone mapping results on image 9 (refer to Fig. 10) by competing tone mapping operators. From (a) to (h): results by “Mantiuk” (Mantiuk et al., 2008), “Drago” (Drago et al., 2003), “Fattal” (Fattal et al., 2002), “Kuang” (Kuang et al., 2007), “Farbman” (Farbman et al., 2008), “Shan” (Shan et al., 2010), “Shibata” (Shibata et al., 2016), and ours.

mapping performance.

It should be pointed out that the subjective testing results are not well consistent with the objective metrics used in this paper. Existing objective metrics for tone mapping operators are primarily focused on structural similarity (Yeganeh and Wang, 2013), feature similarity (Nafchi et al., 2015), visibility (Aydin et al., 2008; Kundu et al., 2017; Song et al., 2016), contrast (Kundu et al., 2017; Song et al., 2016), naturalness (Nafchi et al., 2015; Yeganeh and Wang, 2013), and chrominance (Song et al., 2016). These quality measures are derived from general image quality assessment methods and they may not be suitable for the tone mapping problem. It is still a challenging issue to design a faithful perceptual quality measure to assess tone mapping operators. In addition, we found that the naturalness index should not be over-emphasized for evaluating tone mapping methods via our subjective experiments, and that the color information plays an important role in assessing tone mapped images.

4. Conclusion

In this paper, we presented a clustering based content and color adaptive tone mapping method. Different from previous methods which are mostly filtering based, our method works on image patches, and it decomposes each patch into three components: patch mean, color variation and color structure. Based on the color structure component, we clustered image patches into clusters, and calculated the PCA transform matrix for each cluster. The patches were then transformed into its PCA domain, and the s-shaped arctan function was used to adjust their PCA coefficients. We further extended our method to two scales and proposed an offline clustering implementation to improve its fine-texture preservation and efficiency. Experiments on 20 sets of HDR data demonstrated the superior performance of our method to representative tone mapping methods.

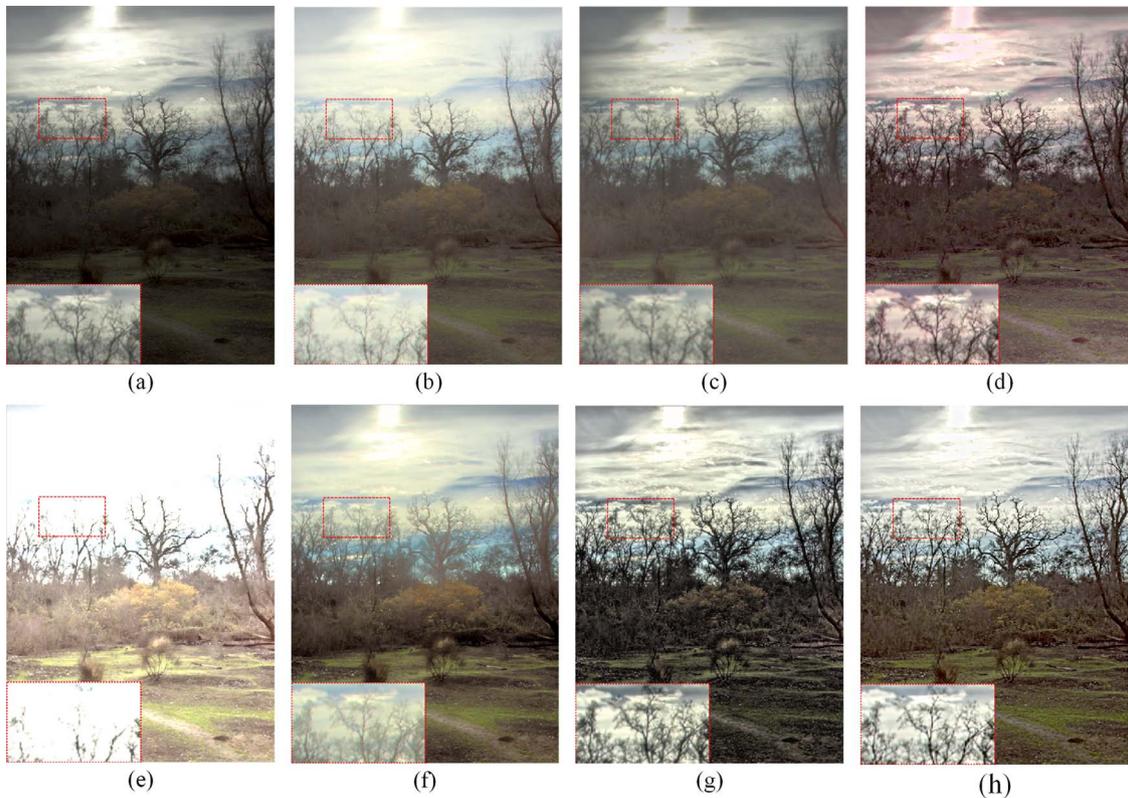


Fig. 16. The tone mapping results on image 17 (refer to Fig. 10) by competing tone mapping operators. From (a) to (h): results by “Mantiuk” (Mantiuk et al., 2008), “Drago” (Drago et al., 2003), “Fattal” (Fattal et al., 2002), “Kuang” (Kuang et al., 2007), “Farbman” (Farbman et al., 2008), “Shan” (Shan et al., 2010), “Shibata” (Shibata et al., 2016), and ours.

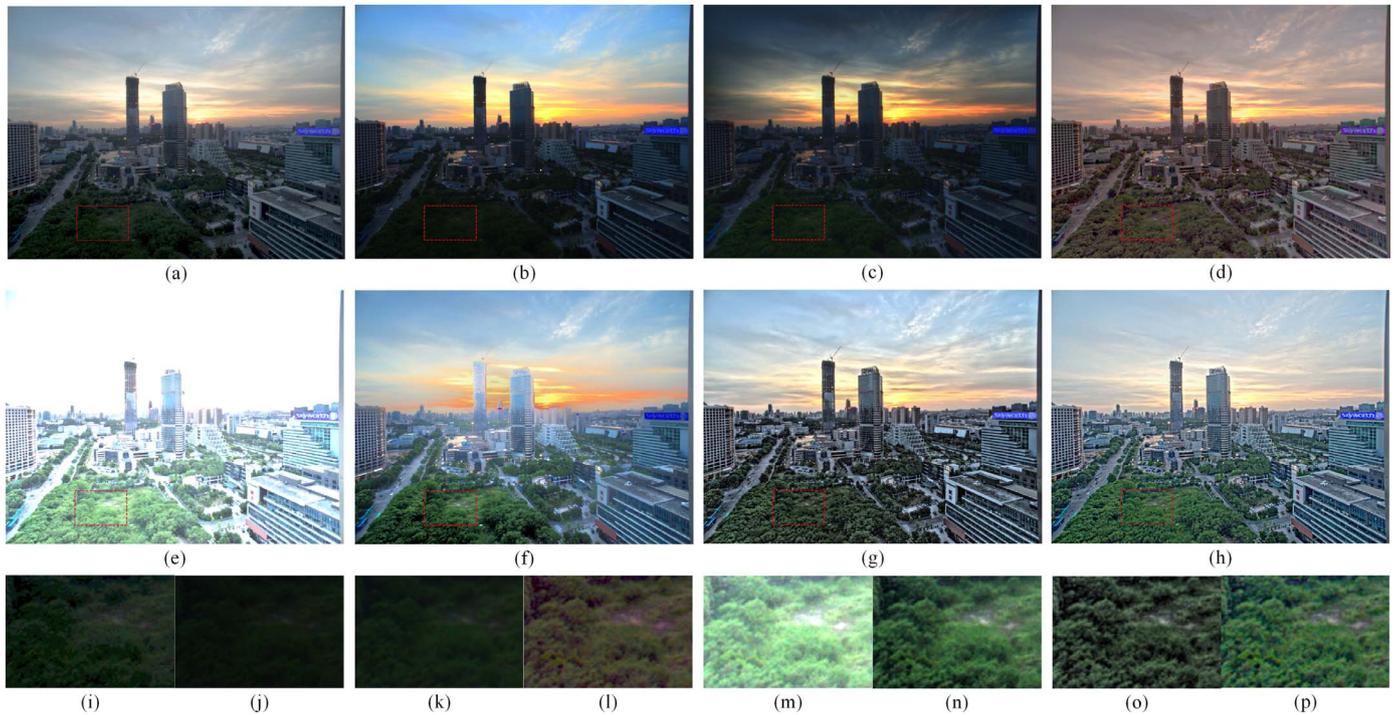


Fig. 17. The tone mapping results on image 18 (refer to Fig. 10) by competing tone mapping operators. From (a) to (h): results by “Mantiuk” (Mantiuk et al., 2008), “Drago” (Drago et al., 2003), “Fattal” (Fattal et al., 2002), “Kuang” (Kuang et al., 2007), “Farbman” (Farbman et al., 2008), “Shan” (Shan et al., 2010), “Shibata” (Shibata et al., 2016), and ours. From (i) to (p): the close-ups of (a)–(h). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article).

Acknowledgment

We would like to thank the editors and reviewers for constructive suggestions and detailed comments which helped improve the algorithm and manuscript. This work is supported by Hong Kong RGC GRF (general research fund) grant (PolyU 5313/13E).

References

- Aydin, T.O., Mantiuk, R., Myszkowski, K., Seidel, H.-P., 2008. Dynamic range independent image quality assessment. *ACM Trans. Graph. (TOG)* 27 (3), 69.
- Badki, A., Kalantari, N.K., Sen, P., 2015. Robust radiometric calibration for dynamic scenes in the wild. *Computational Photography (ICCP)*, 2015 IEEE International Conference on. IEEE, pp. 1–10.
- Banterle, F., Artusi, A., DeBattista, K., Chalmers, A., 2011. *Advanced High Dynamic Range Imaging: Theory and Practice*. CRC Press.
- Chakrabarti, A., Xiong, Y., Sun, B., Darrell, T., Scharstein, D., Zickler, T., Saenko, K., 2014. Modeling radiometric uncertainty for vision with tone-mapped color images. *IEEE Trans. Pattern Anal. Mach. Intell.* 36 (11), 2185–2198.
- Chen, H.-T., Liu, T.-L., Fuh, C.-S., 2005. Tone reproduction: a perspective from luminance-driven perceptual grouping. *Int. J. Comput. Vis.* 65 (1–2), 73–96.
- Debevec, P.E., Malik, J., 1997. Recovering high dynamic range radiance maps from photographs. *Proceedings of the 24th Annual Conference on Computer Graphics and Interactive Techniques*. ACM Press/Addison-Wesley Publishing Co., pp. 369–378.
- Dong, W., Zhang, L., Shi, G., Wu, X., 2011. Image deblurring and super-resolution by adaptive sparse domain selection and adaptive regularization. *IEEE Trans. Image Process.* 20 (7), 1838–1857.
- Drago, F., Myszkowski, K., Annen, T., Chiba, N., 2003. Adaptive logarithmic mapping for displaying high contrast scenes. *Computer Graphics Forum*. 22. Wiley Online Library, pp. 419–426.
- Duan, J., Bressan, M., Dance, C., Qiu, G., 2010. Tone-mapping high dynamic range images by novel histogram adjustment. *Pattern Recognit.* 43 (5), 1847–1862.
- Durand, F., Dorsey, J., 2002. Fast bilateral filtering for the display of high-dynamic-range images. *ACM Transactions on Graphics (TOG)*. 21. ACM, pp. 257–266.
- Farbman, Z., Fattal, R., Lischinski, D., Szeliski, R., 2008. Edge-preserving decompositions for multi-scale tone and detail manipulation. *ACM Transactions on Graphics (TOG)*. 27. ACM, pp. 67.
- Fattal, R., Lischinski, D., Werman, M., 2002. Gradient domain high dynamic range compression. *ACM Transactions on Graphics (TOG)*. 21. ACM, pp. 249–256.
- Ferradans, S., Bertalmio, M., Provenzi, E., Caselles, V., 2011. An analysis of visual adaptation and contrast perception for tone mapping. *IEEE Trans. Pattern Anal. Mach. Intell.* 33 (10), 2002–2012.
- Gu, B., Li, W., Zhu, M., Wang, M., 2013. Local edge-preserving multiscale decomposition for high dynamic range image tone mapping. *IEEE Trans. Image Process.* 22 (1), 70–79.
- Gu, H., Wang, Y., Xiang, S., Meng, G., Pan, C., 2012. Image guided tone mapping with locally nonlinear model. *European Conference on Computer Vision*. Springer, pp. 786–799.
- Gu, K., Wang, S., Zhai, G., Ma, S., Yang, X., Lin, W., Zhang, W., Gao, W., 2016. Blind quality assessment of tone-mapped images via analysis of information, naturalness, and structure. *IEEE Trans. Multimedia* 18 (3), 432–443.
- Guarnieri, G., Marsi, S., Ramponi, G., 2011. High dynamic range image display with halo and clipping prevention. *IEEE Trans. Image Process.* 20 (5), 1351–1362.
- He, K., Sun, J., Tang, X., 2013. Guided image filtering. *IEEE Trans. Pattern Anal. Mach. Intell.* 35 (6), 1397–1409.
- Jobson, D.J., Rahman, Z.-u., Woodell, G.A., 1997. Properties and performance of a center/surround retinex. *IEEE Trans. Image Process.* 6 (3), 451–462.
- Kou, F., Chen, W., Wen, C., Li, Z., 2015. Gradient domain guided image filtering. *IEEE Trans. Image Process.* 24 (11), 4528–4539.
- Kuang, J., Johnson, G.M., Fairchild, M.D., 2007. icam06: a refined image appearance model for hdr image rendering. *J. Vis. Commun. Image Represent.* 18 (5), 406–414.
- Kundu, D., Ghadiyaram, D., Bovik, A., Evans, B., 2017. No-reference quality assessment of tone-mapped hdr pictures. *IEEE Trans. Image Process.*
- Larson, G.W., Rushmeier, H., Piatko, C., 1997. A visibility matching tone reproduction operator for high dynamic range scenes. *IEEE Trans. Vis. Comput. Graph.* 3 (4), 291–306.
- Li, Y., Sharan, L., Adelson, E.H., 2005. Compressing and companding high dynamic range images with subband architectures. *ACM Transactions on Graphics (TOG)*. 24. ACM, pp. 836–844.
- Li, Z., Zheng, J., 2014. Visual-saliency-based tone mapping for high dynamic range images. *IEEE Trans. Ind. Electron.* 61 (12), 7076–7082.
- Li, Z., Zheng, J., Zhu, Z., Yao, W., Wu, S., 2015. Weighted guided image filtering. *IEEE Trans. Image Process.* 24 (1), 120–129.
- Ma, K., Li, H., Yong, H., Wang, Z., Meng, D., Zhang, L., 2017. Robust multi-exposure image fusion: a structural patch decomposition approach. *IEEE Trans. Image Process.* 26 (5), 2519–2532.
- Ma, K., Yeganeh, H., Zeng, K., Wang, Z., 2015a. High dynamic range image compression by optimizing tone mapped image quality index. *IEEE Trans. Image Process.* 24 (10), 3086–3097.
- Ma, K., Zeng, K., Wang, Z., 2015b. Perceptual quality assessment for multi-exposure image fusion. *IEEE Trans. Image Process.* 24 (11), 3345–3356.
- Mahmoudabadi, H., Olsen, M.J., Todorovic, S., 2017. Detecting sudden moving objects in a series of digital images with different exposure times. *Comput. Vis. Image Understanding*.
- Mantiuk, R., Daly, S., Kerofsky, L., 2008. Display adaptive tone mapping. *ACM Transactions on Graphics (TOG)*. 27. ACM, pp. 68.
- Meylan, L., Susstrunk, S., 2006. High dynamic range image rendering with a retinex-based adaptive filter. *IEEE Trans. Image Process.* 15 (9), 2820–2830.
- Mitsunaga, T., Nayar, S.K., 1999. Radiometric self calibration. *Computer Vision and Pattern Recognition, 1999. IEEE Computer Society Conference on*. 1 IEEE.
- Nafchi, H.Z., Shahkolaei, A., Moghaddam, R.F., Cheriet, M., 2015. Fsim: a feature similarity index for tone-mapped images. *IEEE Signal Process. Lett.* 22 (8), 1026–1029.
- Reinhard, E., Devlin, K., 2005. Dynamic range reduction inspired by photoreceptor physiology. *IEEE Trans. Vis. Comput. Graph.* 11 (1), 13–24.
- Reinhard, E., Heidrich, W., Debevec, P., Pattanaik, S., Ward, G., Myszkowski, K., 2010. *High Dynamic Range Imaging: Acquisition, Display, and Image-based Lighting*. Morgan Kaufmann.
- Reinhard, E., Stark, M., Shirley, P., Ferwerda, J., 2002. Photographic tone reproduction for digital images. *ACM Trans. Graph. (TOG)* 21 (3), 267–276.
- Shan, Q., Jia, J., Brown, M.S., 2010. Globally optimized linear windowed tone mapping. *IEEE Trans. Vis. Comput. Graph.* 16 (4), 663–675.
- Shibata, T., Tanaka, M., Okutomi, M., 2016. Gradient-domain image reconstruction framework with intensity-range and base-structure constraints. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. pp. 2745–2753.
- Song, Y., Jiang, G., Yu, M., Zhang, Y., Shao, F., Peng, Z., 2016. Naturalness index for a tone-mapped high dynamic range image. *Appl. Opt.* 55 (35), 10084–10091.
- Tumblin, J., Rushmeier, H., 1993. Tone reproduction for realistic images. *IEEE Comput. Graph. Appl.* 13 (6), 42–48.
- Tumblin, J., Turk, G., 1999. Lcis: A boundary hierarchy for detail-preserving contrast reduction. *Proceedings of the 26th Annual Conference on Computer Graphics and Interactive Techniques*. ACM Press/Addison-Wesley Publishing Co., pp. 83–90.
- Wang, Z., Bovik, A.C., Sheikh, H.R., Simoncelli, E.P., 2004. Image quality assessment: from error visibility to structural similarity. *IEEE Trans. Image Process.* 13 (4), 600–612.
- Ward, G.J., 1994. The radiance lighting simulation and rendering system. *Proceedings of the 21st Annual Conference on Computer Graphics and Interactive Techniques*. ACM, pp. 459–472.
- Wu, S., Yang, L., Xu, W., Zheng, J., Li, Z., Fang, Z., 2016. A mutual local-ternary-pattern based method for aligning differently exposed images. *Comput. Vis. Image Understanding* 152, 67–78.
- Xu, J., Zhang, L., Zuo, W., Zhang, D., Feng, X., 2015. Patch group based nonlocal self-similarity prior learning for image denoising. *Proceedings of the IEEE International Conference on Computer Vision*. pp. 244–252.
- Xu, L., Lu, C., Xu, Y., Jia, J., 2011. Image smoothing via l0 gradient minimization. *ACM Transactions on Graphics (TOG)*. 30. ACM, pp. 174.
- Yeganeh, H., Wang, Z., 2013. Objective quality assessment of tone-mapped images. *IEEE Trans. Image Process.* 22 (2), 657–667.
- Zhang, L., Dong, W., Zhang, D., Shi, G., 2010. Two-stage image denoising by principal component analysis with local pixel grouping. *Pattern Recognit.* 43 (4), 1531–1549.
- Zhang, L., Zhang, L., Mou, X., Zhang, D., 2011. Fsim: a feature similarity index for image quality assessment. *IEEE Trans. Image Process.* 20 (8), 2378–2386.