Tone Reproduction: A Perspective from Luminance-Driven Perceptual Grouping

Hwann-Tzong Chen   Tyng-Luh Liu   Tien-Lung Chang
Institute of Information Science, Academia Sinica
Nankang, Taipei 115, Taiwan
{*pras, liutyng, tino*}*@iis.sinica.edu.tw*

Abstract

We address the tone reproduction problem by integrating local adaptation with global-contrast consistency. Many previous works have tried to compress high-dynamic-range (HDR) luminances into a displayable range in imitation of the local adaptation mechanism of human eyes. Nevertheless, while the realization of local adaptation is not theoretically defined, exaggerating such effects often causes unnatural global contrasts. We propose a luminance-driven perceptual grouping process to derive a sparse representation of HDR luminances, and use the grouped regions to approximate local properties of luminances. The advantage of incorporating a sparse representation is twofold: We can simulate local adaptation based on region information, and subsequently apply piecewise tone mappings to monotonize the relative brightness over only a few perceptually significant regions. Our experimental results show that the proposed framework gives a good balance in preserving local details and maintaining global contrasts of HDR scenes.

1. Introduction

The attempt to reproduce the visual perception of the real world is at the heart of painting and photography. Artists have long been endeavoring to develop skills in simulating actual reflected light within the limitation of the medium, since our world generally delivers a much wider range of luminances than pigments can reflect. Apart from artistic concern, recreating real-scene impressions on limited media is also inevitable in many vision and graphics applications. For example, the highest contrast of today's LCD monitors is around $1,000$ to 1; however, we may still need to display a sunset scene whose contrast exceeds $10,000$ to 1.

By a tone reproduction problem, we focus on establishing an efficient method to faithfully reconstruct the high-dynamic-range (HDR) radiance on a low-dynamic-range (LDR) image. The *dynamic range* of a digital image is simply the contrast ratio in intensity between its brightest and darkest parts. In [20], Ward introduces a floating-point picture format to record HDR radiance in 32 bits per pixel, and designs a graphics rendering system that outputs images in that format. Debevec and Malik have shown that HDR radiance of *real* scenes may be captured using regular SLR or digital cameras [5]. They propose a method to combine a series of pictures with different exposure settings into a single HDR image, which is called a *radiance map*, with contrast of about $250,000$ to 1. In this context, the aim of our research can be stated as solving the tone reproduction problem of radiance maps, that is, generating a displayable standard RGB image that preserves perceptual properties of the original HDR radiance map.

1.1. Related Work

Several works have been devoted to producing HDR images of real scenes [2], [5], [11], [12], including those that are designed to capture HDR luminances simultaneously under multiple exposures [2], [11]. Inherently, panoramic imaging can also be extended to carry HDR luminances by adding spatially varying optical filters to a camera [1], [16]. Different from static HDR imaging, Kang et al. [9] propose to generate HDR videos by changing the exposure of each frame and then by stitching consecutive frames.

On displaying HDR images, tone reproduction addresses *visibility* and *impression* through finding an appropriate mapping to compress high contrasts into a visible (displayable) range, accounting for perceptual fidelity. With a global mapping, pixels are mapped *uniformly* regardless of their spatial or local properties, and hence details are often smeared. The main advantage of using a global mapping is its efficiency. Ward et al. [21] describe a more sophisticated method to adjust contrast globally based on luminance histograms. Nevertheless, the approach still smooths out the details in areas of flat histograms. To improve visual fidelity, a number of tone reproduction methods have explored *nonuniform* (local) mappings, as human visual system operates more likely this way [3], [6], [7], [8], [14], [19]. In particular, visual cells are organized in a center-surround manner so that we can see a wide range of luminances by discriminating locally. In simulating the center-surround organization, Reinhard et al. [14] calculate, for each pixel, the average intensity of a *proper* circular region, and then use the information to adjust a mapping function.

Decomposing a radiance map into layers is another popular choice for preserving image details. Methods of this kind often separate an image into an *illumination layer* and a *reflectance layer*. The illumination layer carries the luminance information of the original image and thus has a wider dynamic range, while the reflectance layer keeps the textures and is of low dynamic range. Consequently, the dynamic range of an HDR image can be reduced by compressing its illumination layer. For images of natural scenes, Land's retinex theory [10] can be used to estimate illumination and reflectance. Indeed center-surround based and layer-decomposition based methods are closely related. Both aim to preserve details and exploit local adaptation to match human perception. However, overemphasizing local contrasts may produce halos, which are defects of reversal contrasts. A number of new methods have been introduced to resolve halos by incorporating more appropriate local averaging schemes, e.g., *bilateral filtering* [18] used in [6], [7], multi-scale Gaussians [3], and dodging-and-burning [14], or, by directly working on the gradient domain according to derived PDE formulations, e.g. *anisotropic diffusion* [19] and the Poisson equation [8].

From a segmentation viewpoint, Schlick [17] has proposed to divide an image into zones of similar values, and then compute the average intensity of each zone. The average intensity map can be used to constitute the spatially nonuniform tone mapping function. Yee and Pattanaik [22] develop a multi-layer partitioning and grouping algorithm to compute the local adaptation luminance. In each layer, pixels are partitioned based on a given intensity resolution (bin-width), and pixels that are partitioned into the same bin form a group. Each pixel's local adaptation luminance value is thus computed by averaging over all pixels of the groups (in different layers) to which the pixel belongs.

1.2. Our Approach

We describe an HDR tone reproduction method based on perceptual grouping in luminance. Since most tone reproduction techniques are intended to construct appropriate average luminances for local adaptation, we believe that it is worth investigating the perceptual grouping for approximating the local characteristics of luminance. In our approach, the grouping process gives rise to a sparse representation for an HDR image so that it is possible to estimate *local adaptation luminances* [3] based on perception-based region information. As a result, we are able to consider the tone reproduction problem by taking account of local details and global perceptual impression. In the following sections, we describe first how to construct the local adaptation luminances by perceptual grouping, and then explain how to use the region information to modulate the mapping functions for tone reproduction.

2. A Sparse Representation for HDR Images

The luminance values of an HDR image can be computed from its R, G, and B channels by $L(x,y) = 0.2126R(x,y) + 0.7152G(x,y) + 0.0722B(x,y)$. To reduce $L(x,y)$ into a low dynamic one, we propose a *sparse representation* that decomposes an HDR image into regions through a perceptual grouping process. The dynamic-range compression is then carried out *region-wise*, where the advantages of using the perceptually significant region information will be explained in the next section. Suffice it to say now that working on an adequate number of regions, we can perform the HDR compression without incurring excessive overheads in patching together the results across different regions. On deriving such a decomposition, it takes two steps: *adaptive block partition* and *perceptual grouping*, detailed in what follows.

2.1. Adaptive Block Partition

Previous experience on exploring perception has suggested that the human visual system senses the contrast of light based on *intensity ratio* rather than *intensity difference* (e.g., see the Weber's law discussed in [13], p.672). Following this observation, we consider the decomposition of an HDR image by examining its luminance property in the logarithmic domain. More precisely, the luminance L is transformed into log-luminance by $\tilde{L}(x,y) = \log L(x,y)$.

Understandably, the outcome of grouping depends on the choice of the *basic element* of an image partition. While working on pixel level is both time-consuming and sensitive to noise, we also find partitioning with blocks of uniform size often leads to unsatisfactory segmentation results. Though the situation could be improved by using small-size blocks, such a tactic again has the drawback of inefficiency. We thus design an adaptive scheme to partition the image with blocks of two different sizes. The smaller blocks are placed in the areas of strong log-luminance gradients; the larger ones are in the areas of less log-luminance variation.

To partition the log-luminance $\tilde{L}$ adaptively, we use Canny edge detector [4] to obtain the edge information, and then divide $\tilde{L}$ into blocks of larger size $b_\ell \times b_\ell$. For those blocks containing Canny edges, they are further split into blocks of smaller size $b_s \times b_s$. An example to illustrate these steps is given in Figure 1, where block sizes $b_\ell = 8$ and $b_s = 2$ are used in all our experiments too.

2.2. Perceptual Grouping

Among the many possible ways to group the blocks, we are interested in finding a sparseness one, driven by the log-luminance factor. For that, we look at two important matters: 1) how to define an appropriate distance function

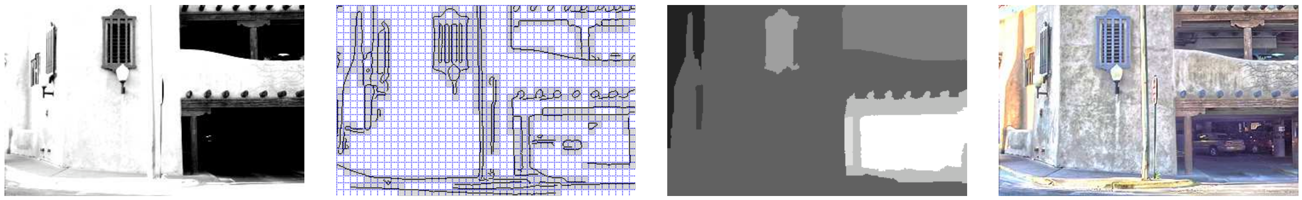

Figure 1: *Garage.* From left to right, the corresponding log-luminance $\tilde{L}$; the adaptive block partition of $\tilde{L}$, where blocks of the smaller size are located in those shaded areas on top of the Canny edges; a sparse representation of 8 regions for $\tilde{L}$; and the displayable *garage* image derived by our method.

to measure the degree of similarity between two regions, and 2) how to proceed with a reasonable grouping process to derive a sparse representation. These two issues can be properly addressed through integrating a *perceptual distance* with a *luminance-driven grouping process*.

2.2.1 Perceptual Distance

We use the *earth mover's distance* (EMD) to evaluate the perceptual similarity between two image regions [15]. Applied mostly in image retrieval, EMD has been proven to be a useful measurement to perceptually correlate two images. In fact, finding the EMD between two distributions is equivalent to solving a *minimum cost flow* problem.

The objects involved in the calculation of EMD are often represented in the form of *signature*. A signature is a set of clusters of which each cluster comprises a pair of feature and weight. In our formulation, the signature of a region (could be just a block or a region of many blocks) is computed as follows. We equally divide the dynamic range of the region into three bins. The mean s_i and the number h_i of the pixels in each bin are then calculated. It also directly implies the weight of each bin is $w_i = h_i / \sum_j h_j$. Thus, the signature $\mathbf{p}$ of each region contains three clusters $\{(s_1, w_1), (s_2, w_2), (s_3, w_3)\}$ that accordingly represent the *bright*, the *middle-gray*, and the *dark* part of that region. Written explicitly, the perceptual distance between two regions $\mathcal{R}_1$ and $\mathcal{R}_2$ is defined by

$$D(\mathcal{R}_1, \mathcal{R}_2) = \mathrm{EMD}(\mathbf{p}_1, \mathbf{p}_2), \tag{1}$$

where $\mathbf{p}_i$ is the signature of region $\mathcal{R}_i$.

2.2.2 Luminance-Driven Grouping

How to optimally decompose an image is a cognitive problem. While analytic arguments are difficult to establish, we prefer a compact/sparse representation to decompose an image into *few* regions (see Figure 1). Specifically, we adopt a greedy approach to grow a new region *as large as possible*, starting each time from the location of the *brightest* log-luminance value in the unvisited areas. That is, the algorithm follows a *brightest-block-first* rule to determine the seeds and to merge image blocks. Since each block's signature includes three clusters, i.e., the bright, the middle-gray, and the dark parts, the brightest block can be simply identified as the one with the largest s_1.

All image blocks are initially marked as *unvisited*. Later on as the grouping process iterates, the number of unvisited blocks decreases. At iteration k, we pick the unvisited and brightest block, say, block $\mathcal{B}_{i^*}$, and start to grow the region $\mathcal{R}_k$ from it. These steps of region grouping are summarized in Algorithm 1. Upon termination, the process will yield a decomposition that each derived region consists of connected blocks of similar luminance distributions. Indeed, our algorithm works by balancing the local and global similarity in a region. Similar blocks are pulled into the same region if the EMD between two neighboring blocks is smaller than δ. On the other hand, a region will stop growing when all blocks right beside the region boundary are not close enough to the whole region within θ. (See Algorithm 1 for further details.) The typical values of EMD threshold θ are from 1.5 to 2.0, and those of δ are between 0.5 and 1.0.

3. Region-Based Tone Mapping

Our method to compress the high dynamic range relies on region-wise constructing suitable tone mapping functions, based on the estimations of *local adaptation luminances*. It is also critical that the resulting *piecewise tone mappings* could be smoothly pieced together to produce a good-quality and displayable image without violating the overall impression of the original HDR radiance map. In this section, we will show that all these issues can be satisfactorily addressed by considering the region information encoded in a luminance-driven sparse representation.

3.1. Local Adaptation Luminances

Human visual system attains the HDR perception by locally adapting to different levels of luminance to ensure a proper dynamic range can be recreated by the responses of

Figure 2: *Stanford memorial.* From left to right, the respective results derived by *bilateral filtering* [7], *photographic tone reproduction* [14], *our method*, and *gradient domain* [8]. Our method not only reveals fine details but maintains a global impression similar to that of the photographic [14].

visual cells. For tone reproduction, the local adaptation effect is often simulated by computing the local adaptation luminance. Thus it is important to have a reliable way for pixel-wise estimating the local adaptation luminance of an HDR image. And that in turn can be done by investigating the average log-luminance of a *suitable neighborhood* about each pixel.

Let $\tilde{V}(x, y)$ be the local adaptation log-luminance at pixel (x, y). To compute $\tilde{V}$, we consider a generalized version of *bilateral filtering* [18] by constructing a region-dependent scheme such that the computation of $\tilde{V}(x, y)$ takes account of bilateral effects from different regions. Particularly, for each pixel (x, y) in region $\mathcal{R}_k$, we have

$$\tilde{V}(x,y) = \frac{1}{\tilde{Z}_{x,y}} \Big\{ \sum\nolimits_{(i,j)\in\mathcal{R}_k} \tilde{L}(i,j)\, G_{x,y}(i,j)\, K_{x,y}(i,j) + \sum\nolimits_{(i,j)\notin\mathcal{R}_k} \tilde{L}(i,j)\, G_{x,y}(i,j)\, K'_{x,y}(i,j) \Big\}, \tag{2}$$

where

$$\begin{aligned} G_{x,y}(i,j) &= \exp\Big\{-\big((i-x)^2+(j-y)^2\big)/2\sigma_s^2\Big\} \\ K_{x,y}(i,j) &= \exp\Big\{-\big(\tilde{L}(i,j)-\tilde{L}(x,y)\big)^2/2\sigma_r^2\Big\} \\ K'_{x,y}(i,j) &= \exp\Big\{-\big(\tilde{L}(i,j)-\tilde{L}(x,y)\big)^2/2\sigma_{r'}^2\Big\} \\ \tilde{Z}_{x,y} &= \sum\nolimits_{(i,j)\in\mathcal{R}_k} G_{x,y}(i,j)\, K_{x,y}(i,j) \\ &\quad + \sum\nolimits_{(i,j)\notin\mathcal{R}_k} G_{x,y}(i,j)\, K'_{x,y}(i,j)\,. \end{aligned}$$

Eq. (2) includes two parts of bilateral filtering. The first part calculates the averaging in the same region, and the second evaluates the contributions from other regions. Note that we have used the same spatial-domain filter $G_{x,y}$ to regulate the effects to pixel (x, y) from all regions. Such a choice makes it possible to apply the fast bilateral filtering implementation described in [7]. On the other hand, we have $\sigma_r \geq \sigma_{r'}$ to ensure a flatter and more expanded range-domain filter $K_{x,y}$ for pixels in the same region of (x, y), and to lessen the influences from pixels of different regions with ${K'}_{x,y}$. If $\sigma_r = \sigma_{r'}$, the proposed scheme in (2) is reduced to the one used in [7]. With the region-based filtering, the effects of local adaptation inside perceptually related regions can be enhanced by using a larger σ_r. A suitable value for σ_s can be set to 4% of the image size, and $\sigma_r = 0.4$ and $\sigma_{r'} = 0.5 \times \sigma_r$. So far, we have worked in log-luminance domain. The local adaptation luminance, denoted as V, can be recovered by $V = \exp(\tilde{V})$.

3.2. Piecewise Tone Mappings

A handy choice of simple functions for compressing the high luminances into the displayable range $[0, 1]$ is the non-linear mapping $\varphi(x) = x/(1 + x)$. If φ is applied to the whole luminance map, i.e., $L' = \varphi(L) = L/(1 + L)$, we will actually obtain a displayable but smoother image. This type of compression scheme is called *global mapping* or *spatially uniform mapping*. Another tone mapping method is to extract from L the detail layer H by $H = L/V$. Then, only the local adaptation luminance is compressed by $V' = \varphi(V)$. Recombining the detail layer H with the compressed V', we have $L' = H \times V' = (L/V) \times (V/(1 + V)) = L/(1 + V)$, a *local mapping* for preserving details.

Even though a global mapping like φ often has the drawback of losing the details in brighter areas, its monotone property ($d\varphi/dx > 0$) is desirable for preventing halos

Algorithm 1: Luminance-Driven Grouping with EMD.

Input : The adaptive block partition $\mathcal{B}$ of $\tilde{L}$.
Output: A sparse representation of $\tilde{L}$.

Initialization: Create an empty priority queue $\mathcal{Q}$ on D as in (1); Compute the signatures of all blocks; Label all blocks *unvisited*; Let $k \leftarrow 1$; Choose $\theta, \delta > 0$.

while $\exists$ *unvisited block* **do**
    Create a new region $\mathcal{R}_k \leftarrow \emptyset$, and let $\mathcal{Q} \leftarrow \emptyset$;
    Select the brightest block $\mathcal{B}_{i^*}$ and add it into $\mathcal{Q}$;
    Let $D(\mathcal{B}_{i^*}, \mathcal{R}_k) \leftarrow 0$;
    while $\mathcal{Q} \neq \emptyset$ **do**
        Retrieve $\mathcal{B}_i$ from $\mathcal{Q}$ with the smallest $D(\mathcal{B}_i, \mathcal{R}_k)$;
        if $D(\mathcal{B}_i, \mathcal{R}_k) < \theta$ **then**
            Add $\mathcal{B}_i$ into $\mathcal{R}_k$ and label $\mathcal{B}_i$ *visited*;
            Update $\mathcal{Q}$ by recomputing the Ds;
            Find $\mathcal{B}_i$'s *unvisited* neighbors $\{\mathcal{B}_j\}$;
            foreach $\mathcal{B}_j$ *satisfies* $D(\mathcal{B}_j, \mathcal{B}_i) < \delta$ **do**
                Insert $\mathcal{B}_j$ into $\mathcal{Q}$;
            Update the signature of region $\mathcal{R}_k$.
        else
            Goto 1.
1   $k \leftarrow k + 1$.
Output all $\mathcal{R}_k$s.

and other artifacts. It would be favorable if the monotonicity can be incorporated into a local tone mapping method. Nonetheless, one still needs to figure out a reasonable way to *monotonize* a local mapping and determine an appropriate "neighborhood" for each such a monotonization.

We argue here that the sparse representation does provide useful hints for solving the foregoing problems. Perceptually, the derived region decomposition correlates with an overall visual impression about the scene. Maintaining this impression after compressing the dynamic range should be a good criterion. We thus consider a *piecewise tone mapping* scheme that region-wise performs the monotonization, and globally retains the relative brightness among different regions, i.e., those $\mathcal{R}_k$s derived in Algorithm 1. The whole idea of such tone mappings is realized by the following four steps:

1. *Design a local mapping* ψ. As pointed out in Sec. 3.1, the local adaptation luminance V plays an important role in compressing the luminance L. We define a local mapping ψ of the following form.

$$\psi(L, V; \rho, \gamma) = (\frac{L}{V})^{\rho} \varphi^{\gamma}(V) = (\frac{L}{V})^{\rho} (\frac{V}{1+V})^{\gamma}, \tag{3}$$

where $0 < \rho < 2$ and $0 < \gamma \leq 1$ are spatial-dependent parameters to adjust the image quality resulting from the HDR compression. More precisely, when $\gamma < 1$, dark areas in an HDR radiance map will be compressed into larger brightness levels, compared with the case $\gamma = 1$. On the other hand, since L/V is the detail layer, the ρ values would have direct impacts on preserving the image details after the dynamic-range reduction. (In all our experiments, $\gamma = 0.3$.)

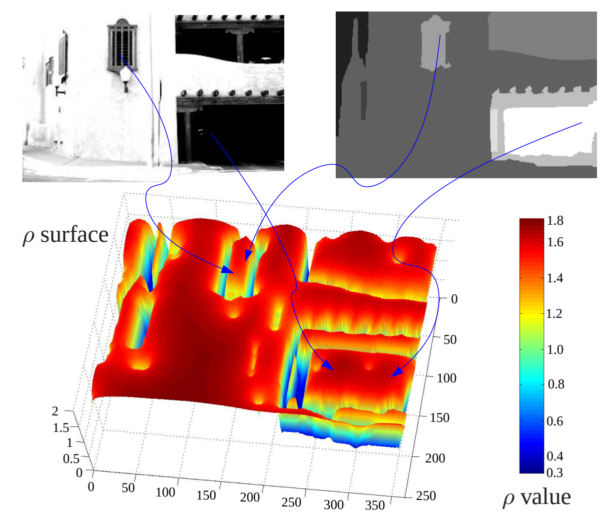

Figure 3: The ρ values after kernel smoothing. Examples of correspondences to the log-luminance and to the region decomposition are illustrated.

2. *Globally reshape* φ^{γ} *by* $\tilde{\varphi}^{\gamma} = \alpha\varphi^{\gamma} + \beta$. Let L_{max} and L_{min} be the maximum and the minimum luminance values of a given radiance map, respectively. To make sure φ^{γ} will take up the complete displayable range $[0, 1]$, we solve the following linear system to obtain the proper scaling and shifting parameters α and β:

$$\begin{bmatrix} \varphi^{\gamma}(L_{max}) & 1 \\ \varphi^{\gamma}(L_{min}) & 1 \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}. \tag{4}$$

The values of α and β will be needed in monotonizing the local mapping of each region $\mathcal{R}_k$.

3. *Estimate* ρ *by kernel smoothing.* For each $\mathcal{R}_k$, we construct a grid D_k within $\mathcal{R}_k$ such that it is the largest grid of resolution $\epsilon \times \epsilon$ with all its grid points at least ϵ-pixel away from the region boundary $\partial\mathcal{R}_k$. We then assign some preliminary ρ_n value to each pixel $n \in D_k$ by

$$\rho_n = \begin{cases} \gamma\,, & \text{if } \log(L_n/V_n) \leq -1, \\ (\gamma + \rho_{max})/2\,, & \text{if } \log(L_n/V_n) \geq 1, \\ \rho_{max}\,, & \text{otherwise.} \end{cases} \tag{5}$$

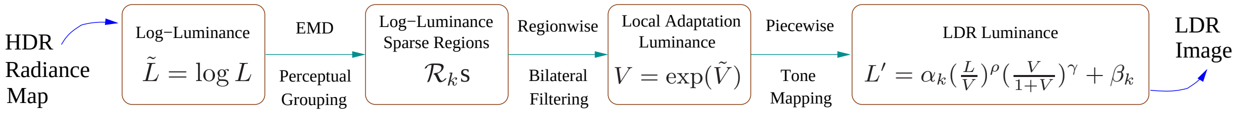

Figure 4: The steps of our tone reproduction method.

The above rules simply reflect that if L_n is already "very different" from its local adaptation luminance V_n, using a larger ρ_n may only cause an inconsistent overemphasis that further amplifies the difference. In addition, for all pixels on $\partial\mathcal{R}_k$, their ρ_n values are also set to γ, and those of the pixels between D_k and $\partial\mathcal{R}_k$ can be computed by interpolation. It will later become clear that putting such constraints near $\partial\mathcal{R}_k$ will be vital in constructing an overall smooth ρ surface across regions. With all these ρ_n values pre-defined, we can now apply *kernel smoothing* to adjust and to derive the ρ values for all pixels in $\mathcal{R}_k$. Notice that, throughout this work, we have $\rho_{max} = 1.8$ and $\epsilon = 4$.

4. *Monotonize local tone mappings.* A monotonization, say for $\mathcal{R}_k$, is to estimate α_k and β_k so that a local mapping ψ in (3) can be elevated to $\tilde{\psi} = \alpha_k\psi + \beta_k$, and to account for both local and global factors in compressing the luminances. We begin by sampling N pixels from $\partial\mathcal{R}_k$ according to the sorted $|\log(L_n/V_n)|$ values in ascending order. In our experiments, using the first 5% of boundary pixels will be sufficient to give good results. The values of α_k and β_k can now be obtained by calculating the least square solution of

$$\begin{bmatrix} \psi_1 & \cdots & \psi_N \\ 1 & \cdots & 1 \end{bmatrix}^T \begin{bmatrix} \alpha_k \\ \beta_k \end{bmatrix} = \begin{bmatrix} \tilde{\varphi}_1^{\gamma} & \cdots & \tilde{\varphi}_N^{\gamma} \end{bmatrix}^T.$$

Finally, we apply $\tilde{\psi}(L, V; \rho, \gamma) = \alpha_k\psi(L, V; \rho, \gamma) + \beta_k$ to each pixel $(x, y) \in \mathcal{R}_k$ to compute its displayable luminance value $L'(x, y)$.

The least square fitting to reshape ψ_k into $\tilde{\psi}_k$ can be justified by the facts that the monotonization uses only pixels on region boundary and their ρ values after kernel smoothing are still close to γ. Furthermore, with this property, we can efficiently derive a globally smooth ρ surface (see Fig. 3). We conclude this section with a remark that the main characteristic of our method is to perform the dynamic-range compression by emphasizing the local details of each region without breaking the global visual consistency.

4. Experiments and Discussions

Having detailed our approach (summarized in Fig. 4), we now describe some of our experimental results and comparisons with other related works. In all our experiments, the HDR images are downloaded from the Web and stored in radiance map format. A typical radiance map with multiple exposure values is shown in Fig. 5. (In this example, the dynamic range of *Stanford memorial* is $250,000:1$.)

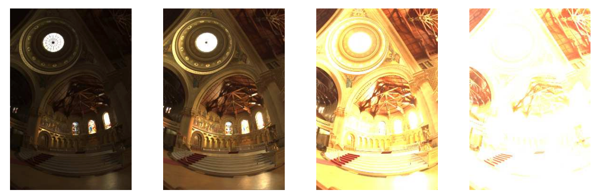

Figure 5: Radiance map. By setting a fixed displayable range and cutting off out-of-range intensities, the radiance map can only be viewed like these images.

The bottlenecks of our method lie in the steps of EMD perceptual grouping and region-wise bilateral filtering. For the *memorial* image of size 512 by 768 in Fig. 5, the elapsed time of perceptual grouping and of bilateral filtering on a 2.4GHz PC is about 3.6s and 5.3s, respectively. Clearly, the complexity of region-wise bilateral filtering depends on the number of regions. By extending fast bilateral filtering [7] to incorporate region support, we can implement a region-wise bilateral filter in a way that slowdowns are not proportional to the number of regions. For instance, ten regions are constructed for the *memorial*, but region-wise bilateral filtering ($\approx$ 5.3s) just doubles the time needed for a typical bilateral filtering ($\approx$ 2.4s).

After luminance reduction, the LDR image can be displayed by multiplying the compression ratio L'/L to each of the high dynamic range RGB channels. In Fig. 6, we show several LDR images derived by our method. Our results have two aspects of visually pleasing effects: 1) Overall impressions of luminance are maintained; 2) Details and local high contrasts are preserved. Some recent methods [7], [8], [14] on displaying HDR images also can eliminate halos and preserve details. The gradient-domain compression [8] performs well in preserving local contrasts and details. However, a noticeable difference between their compression outcomes and ours is that in their results the brighter areas may not be bright enough as they should be. A good example is the circular window of the *memorial* in Fig. 2. Note that, in the original radiance map, the area of the circular window is at least 200 times brighter than the top-right corner.

Figure 6: HDR tone-reproduction results. From left to right and top to bottom (with the number of derived regions): *Tahoe* (5), *office* (68), *Belgium* (32), *clock* (34), *designCenter* (13), *Tintern* (14), *chairs* (14), and *groveC* (17).

While our approach relates to photographic tone reproduction [14] and bilateral filtering [7] in computing the local adaptation luminance, our method generally produces better contrasts within the displayable range, and preserves more details due to γ parameter and flexible ρ values. Again, the examples in Fig. 2 well demonstrate these common phenomena. In addition, we highlight another two distinctions in Fig. 7. The first one is that although the photographic method [14] basically maintains overall impressions, sometimes the approximate local adaptation luminance decreases the brightness of a large bright area, e.g., the sun in Fig. 7a. Owing to piecewise tone-mapping, our results do not have this anomaly, e.g., Fig. 7b. The second distinction is that the bilateral filtering method [7] still generates some halos near the boundaries between bright and dark regions. The differences can be observed near the skyline and near the shadows in Fig. 7c and 7d. Via a smooth ρ-surface adjustment, our method produces a bit lower contrasts along boundaries.

In passing it is worth mentioning that, like [8], our tone-reproduction method can also be used to enhance the image quality of an LDR image. We experiment on this effect by applying the same process listed in Fig. 4 to LDR images. We obtain fairly good results, compared with those by [8] and histogram equalization functions in MATLAB toolbox. One example of the results is provided in Fig. 7e, 7f.

5. Conclusion

We have thoroughly investigated the tone reproduction problem in two aspects: 1) deriving local adaptation luminances for preserving details, and 2) region-wise compressing the dynamic range without breaking the overall impression. For a visual model, the *correctness* of the local adaptation mainly depends on the definition of *locality*. We thus introduce a luminance-driven perceptual grouping to derive a sparse representation for HDR luminances, based on the EMD perceptual distance. With the sparse image decomposition, we are able to improve both the local adaptation luminances and the tone mapping functions. Consequently, a region-wise bilateral weighting scheme can be formulated to enhance the local adaptation effect inside a region. As for piecewise tone mappings, we apply monotonizations to incorporate the global property into local tone mappings so that the brightness impression of a scene is maintained.

Acknowledgments. This work was supported in part by an NSC grant 93-2213-E-001-010.

References

[1] M. Aggarwal and N. Ahuja, "High Dynamic Range Panoramic Imaging," *8th ICCV*, vol. 1, pp. 2–9, 2001.

(a) Photographic method (b) Our method (c) Bilateral filtering (d) Our method (e) Original image (f) Enhanced image

Figure 7: (a) The sun looks too dark [14]. (b) Corrected by our approach. (c) There are some halos near the skyline and the shadows [7]. (d) Our method preserves more details and prevents most halos. (e) & (f) LDR image enhancement.

[2] M. Aggarwal and N. Ahuja, "Split Aperture Imaging for High Dynamic Range," *8th ICCV*, vol. 2, pp. 10–17, 2001.

[3] M. Ashikhmin, "A Tone Mapping Algorithm for High Contrast Images," *13th Eurographics*, pp. 145–156, 2002.

[4] J.F. Canny, "A Computational Approach to Edge Detection," *PAMI*, vol. 8, no. 6, pp. 679–698, 1986.

[5] P.E. Debevec and J. Malik, "Recovering High Dynamic Range Radiance Maps from Photographs," *SIGGRAPH*, pp. 369–378, 1997.

[6] J.M. DiCarlo and B.A. Wandell, "Rendering High Dynamic Range Images," *SPIE: Image Sensors*, vol. 3965, pp. 392–401, 2000.

[7] F. Durand and J. Dorsey, "Fast Bilateral Filtering for the Display of High-Dynamic-Range Images," *SIGGRAPH*, pp. 257–266, 2002.

[8] R. Fattal, D. Lischinski, and M Werman, "Gradient Domain High Dynamic Range Compression," *SIGGRAPH*, pp. 249–256, 2002.

[9] S.B. Kang, M. Uyttendaele, S. Winder, and R. Szeliski, "High Dynamic Range Video," *SIGGRAPH*, 2003.

[10] E.H. Land and J.J. McCann, "Lightness and Retinex Theory," *J. Optical Society of America*, vol. 61, pp. 1–11, 1971.

[11] S.K. Nayar and T. Mitsunaga, "High Dynamic Range Imaging: Spatially Varying Pixel Exposures," *CVPR*, vol. 1, pp. 472–479, 2000.

[12] C. Pal, R. Szeliski, M. Uyttendaele, and N. Jojic, "Probability Models for High Dynamic Range Imaging," *CVPR*, vol. 2, pp. 173–180, 2004.

[13] S. E. Palmer, *Vision Science: Photons to Phenomenology*, MIT Press, 1999.

[14] E. Reinhard, M. Stark, P. Shirley, and J. Ferwerda, "Photographic Tone Reproduction for Digital Images," *SIGGRAPH*, pp. 267–276, 2002.

[15] Y. Rubner, C. Tomasi, and L.J. Guibas, "A Metric for Distributions with Applications to Image Databases," *6th ICCV*, pp. 59–66, 1998.

[16] Y.Y. Schechner and S.K. Nayar, "Generalized Mosaicing: High Dynamic Range in a Wide Field of View," *IJCV*, vol. 53, no. 3, pp. 245–267, 2003.

[17] C. Schlick, "Quantization Techniques for Visualization of High Dynamic Range Pictures," *5th Eurographics*, pp. 7–18, 1994.

[18] C. Tomasi and R. Manduchi, "Bilateral Filtering for Gray and Color Images," *6th ICCV*, pp. 839–846, 1998.

[19] J. Tumblin and G. Turk, "LCIS: A Boundary Hierarchy for Detail-Preserving Contrast Reduction," *SIGGRAPH*, pp. 83–90, 1999.

[20] G. Ward, "Real Pixels," *Graphics Gems II*, pp. 80–83, 1991.

[21] G. Ward, H.E. Rushmeier, and C.D. Piatko, "A Visibility Matching Tone Reproduction Operator for High Dynamic Range Scenes," *IEEE Trans. Visualization and Computer Graphics*, vol. 3, no. 4, pp. 291–306, 1997.

[22] Y.H. Yee and S.N Pattanaik, "Segmentation and Adaptive Assimilation for Detail-Preserving Display of High-Dynamic Range Images," *The Visual Computer*, vol. 19, no. 7-8, pp. 457–466, December 2003.