Vis Comput (2011) 27: 717–727
DOI 10.1007/s00371-011-0578-7

ORIGINAL ARTICLE

Attention-based high dynamic range imaging

Wen-Chieh Lin · Zhi-Cheng Yan

Published online: 19 April 2011

Abstract Many tone mapping algorithms have been proposed based on the studies in Human Visual System; however, they rarely addressed the effects of attention to contrast response. As attention plays an important role in human visual system, we proposed a local tone mapping method that respects both attention and adaptation effects. We adopt the High Dynamic Range (HDR) saliency map to compute an attention map, which predicts the attentive regions and nonattentative regions in an HDR image. The attention map is then used to locally adjust the contrast of the HDR image according to attention and adaptation models found in psychophysics. We applied our tone mapping approach to HDR images and videos and compared with the results generated by three state-of-the-art tone mapping algorithms. Our experiments show that our approach produces results with better image quality in terms of preserving details and chromaticity of visual saliency.

Keywords High dynamic range imaging · Visual saliency · Attention and adaptation

1 Introduction

The luminance in the real world ranges widely from 10^5 cd/m^2 to 10^{-1} cd/m^2 [9]; however, conventional digital images can only represent a limited dynamic range of intensities. To faithfully describe the dynamic range of the real-world luminance, High Dynamic Range Image (HDRI), which typically stores physical values of luminance, was proposed [26]. HDRI plays an important role in image acquisition and processing since it makes images look more realistic. In computer graphics, HDRI is also used to represent the results of global illumination or to render a scene through an HDR environment map or HDR textures. Although HDR has important applications in computer graphics, it cannot be shown on conventional displays or digital image output devices since the dynamic range of these devices is limited. Thus, a tone mapping function that converts the dynamic range of an HDRI to the dynamic range of a low dynamic range image (LDRI) while retaining realistic color and contrast is needed. Although some HDR display devices [28] have emerged recently, they are still under development and very expensive. In fact, even with the availability of HDR devices, a tone mapping function can still play an important role for HDR devices to be readily accessible since HDR data require larger storage. A tone mapping function can be utilized for compressing HDR data effectively [34, 35].

As tone mapping is related to visual perception, several tone mapping algorithms were proposed based on the studies in Human Visual System (HVS) [9, 22]. These tone mapping algorithms usually apply an identical mapping function to the entire image; however, vision studies suggest that attention plays a crucial role in the information processing of human visual system. Attention enhances the performance on hyperacuity, visual search, orientation detection, and discrimination and localization tasks [3]. In particular, it is found in many vision literatures that the contrast sensitivity in our neuron response would increase for attentive stimuli while decreasing for nonattentative stimuli [4, 20, 24]. It has also been shown that the contrast sensitivity of HVS decreases once human eyes adapt to the luminance of a region. According to these findings, it is apparent that we need to

W.-C. Lin (✉) · Z.-C. Yan
Department of Computer Science, National Chiao Tung University, Hsinchu, Taiwan
e-mail: wclin@cs.nctu.edu.tw

Z.-C. Yan
e-mail: peter12.cs97g@g2.nctu.edu.tw

take the attention and adaptation effects into account when designing a tone mapping algorithm.

In this paper, we propose a local tone mapping algorithm that respects the attention and adaptation effects of human visual system. We adopt the HDR saliency map proposed by Brémond et al. [2] to obtain an attention map, which can predict the attentive and nonattentative regions of an HDRI. Once the attentive and nonattentative regions are detected, we then locally adjust the tone mapping function according to the contrast response model found by Pestilli et al. [24]. This model was obtained from a series of psychophysical experiments, which explore the influences of the attention and adaptation effects on the contrast sensitivity of HVS. We also extend our tone mapping algorithm to handle HDR videos. Our experiments on HDR images and videos demonstrate the effectiveness of our algorithm.

The contributions of this paper are introducing the visual attention and adaptation effects to the research field of tone mapping and proposing a local tone mapping method that respects both the attention and adaptation effects of human visual system.

2 Related work

Existing tone mapping methods can be roughly categorized into two types. One is the global tone mapping method, which applies the same mapping function to all pixels of an image. The other is the local tone mapping method, which transforms the luminance of each pixel differently according to the luminance values of its neighborhood pixels. We briefly review several tone mapping methods that are closely related to our work in this section. Global tone mapping methods are usually more efficient in computation than local tone mapping methods.

Global tone mapping methods Tumblin and Rushmeier [30] proposed two observer models based on humans' visual responses to the light in television and film systems. One is for the real world and the other is for display devices. Under the assumption that a real-world observer should be the same as a display observer, they proposed a tone mapping operator for grey-level HDRIs. Ward [33] adopted a simple linear operator in his global tone mapping method that can preserve apparent contrast and visibility. Ferwerda et al. [9] applied psychophysical studies to their tone mapping operator, which can capture the image properties of color appearance, visual acuity and light/dark adaptation. Ward et al. [15] filtered an input HDRI to obtain a foveal sample image and then performed histogram adjustment for tone mapping according to the foveal image. Ward et al. assume that all pixels would participate in the adaptation process; however, in the human visual system, eye movements are critical for acquiring and processing visual information [10]. Therefore, it should be more reasonable to compute the foveal sample image based on a viewer's fixation positions. Pattanaik et al. [22] used a simplified adaptation model proposed by Hunt [12] to simulate the retinal-response process for high contrast scenes. Mantiuk et al. [19] formulated tone mapping as an optimization problem that minimizes the visible contrast distortions between a human visual system and the display. Van Hateren [32] used the human cone model to perform tone mapping, in which all components are represented as temporal kernels to transform HDR images/videos into LDR images/videos.

Local tone mapping methods Global tone mapping is usually used to compress an HDRI for its computational efficiency [26]; however, it may cause some loss of details since all pixels are transformed using the same mapping function. It is desirable to adaptively adjust tone mapping function at different locations to preserve the details. Chiu et al. [6] defined a scaling function and used it as guideline to scale the pixel values. Inspired by artists who usually compress the contrast of large features and add the details in their drawings process, Tumblin and Turk [31] proposed a detail-preserving contrast reduction method for tone mapping. Fattal et al. compressed large magnitude of gradient and solved a Poisson equation to obtain LDR images [8]. Their method avoids some noisy appearance in Tumblin and Turk's method. Durand and Dorsey [7] decomposed an HDRI as product of a base image and a detail image. The base image is created by bilateral filtering the HDRI so that high luminance is preserved. They performed contrast reduction on the base image and multiplied the result with the detail image to produce LDRI. Reinhard et al. [27] adopted the concepts of the zone system and the dodging-and-burning technique in photography to perform the dynamic range compression. Chen et al. [5] defined different tone mapping functions for different objects in an image, in which objects are detected using the Earth Mover's Distance. Recently, there are also interactive tools developed for local tone mapping. Lischinski et al. [18] used brush and stroke to set constraints, which are propagated to form a completely adjusted tone map under an image-guided energy minimization framework. Liang et al. [16] used a simple touch screen to set constraints and modified the stroke-based algorithm [18] so that the local tone mapping can be executed efficiently on mobile devices.

Despite the success of existing tone mapping methods, none of them addressed both the attention and adaptation effects on contrast response. We demonstrate in this paper that a tone mapping function that respects and exploits visual attention and adaptation effects can produce better results.

3 Attention and adaptation effects

We introduce the attention and adaptation effects to the contrast response of the human visual system and the visual attention model used in our tone mapping algorithm.

3.1 Contrast response under attention and adaptation

Attention influences the performance of many visual tasks, such as visual search, localization and discrimination, and contrast response [4, 20, 23, 24]. There are two types of attention effects in human visual system [11]. One is transient or exogenous attention and the other is sustained or endogenous attention. Transient attention is an unconscious and bottom-up process, which is affected by salient stimuli in the scene. Sustained attention is a task-relevant and top-down process. As sustained attention is voluntary and related to cognition, it is more difficult to build a computational model of sustained attention. Therefore, we only consider the transient attention effect in our tone mapping function.

Figure 1 describes how transient attention and adaptation affect the contrast response function in visual neurons [23, 24]. The black curve represents the situation when there are no attention and adaptation effects. When transient attention occurs, the contrast response function would move from the black curve toward the red curve. This means contrast at attentive regions appears more intense as attention stimulates the visual neurons causing stronger response. That is, visual neurons are more sensitive to contrast difference at the attentive regions (the slope of the response function is larger). Sustained attention also affects the contrast sensitivity. It has a similar effect as transient attention for a short duration but the effect of adaptation takes over afterward [17]. In contrast to transient attention, adaptation induces an opposite effect to the contrast response of visual neurons. When our eyes adapt to the luminance of a scene, our visual neurons would be less sensitive to contrast. As illustrated in Fig. 1, the contrast response function would shift from the black curve toward the green curve when adaptation occurs. One can also observe that the perceiving contrast range after adaptation is smaller than that of neutral or attentive conditions. These findings suggest that we should increase the contrast at attentive regions and decrease the contrast at the regions where adaptation occurs.

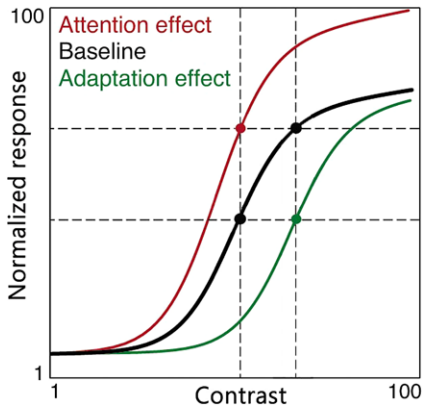

Fig. 1 Contrast response function of a visual neuron summarized by Pestilli et al. [24] (image curtesy of Franco Pestilli). *The black, red*, and *green curve* is the neural response in neutral, attention and adaptation condition, respectively

3.2 Computational model of visual attention

Many computational models of visual attention have been proposed based on the feature integration theory [29]. The theory says that fairly simple visual features are computed over the entire scene in the first step of visual processing, but only those attentive features will be further processed to form a united object representation. Itti et al. [14] proposed the saliency map, which modeled the salient positions where primates would pay attention in an image.

We adopt the HDR saliency map [2] as our transient attention model. HDR saliency map is developed based on the saliency map [14], in which Itti et al. construct Gaussian pyramids $I(\sigma)$ with nine spatial scales $\sigma \in [0..8]$ to simulate the bottom-up process of transient attention in the human visual system. An image is decomposed into intensity, color, and orientation components. Then the center-surround difference operation and across-scale combination are applied to each component to simulate the feature integration theory [29]. This mimics the function of visual receptive field as visual neurons are more sensitive to the stimuli in the center of a visual space than those in the peripheral of the visual space.

The intensity component of an image I is computed using

$$I = \frac{r + g + b}{3}, \tag{1}$$

where r, g, b is the red, green, and blue channel of the input image, respectively. The six feature maps of intensity are defined as follows:

$$I(c, s) = \left| I(c) \ominus I(s) \right|, \tag{2}$$

where $c \in \{2, 3, 4\}$ and $s = c + \delta$ with $\delta \in \{3, 4\}$ denote the feature maps of intensity at different scales (larger scale for higher resolution). The center-surround operator $\ominus$ is an across-scale difference obtained by interpolating the coarser scale to the finer scale and then subtracting them pixel by pixel. The feature maps of color are defined according to the

opponent process theory:

$$R = r - \frac{g+b}{2}, \qquad G = g - \frac{r+b}{2},$$
$$B = b - \frac{r+g}{2}, \qquad Y = \frac{r+g}{2} - \frac{|r-g|}{2} - b. \tag{3}$$

As the responses excited by the opponent channel of one color would inhibit those by the other color, Itti et al. define the 12 color feature maps as

$$RG(c,s) = \big|\big(R(c) - G(c)\big) \ominus \big(R(s) - G(s)\big)\big|$$
$$BY(c,s) = \big|\big(B(c) - Y(c)\big) \ominus \big(B(s) - Y(s)\big)\big|, \tag{4}$$

where $c \in \{2,3,4\}$ and $s = c + \delta$ with $\delta \in \{3,4\}$ denote different scales. The feature maps of orientation are computed using the Gabor pyramid $O(\sigma,\theta)$, where σ represents the scale and $\theta \in \{0°, 45°, 90°, 135°\}$ is the orientation of the Gabor filter

$$O(c,s,\theta) = \big|O(c,\theta) \ominus O(s,\theta)\big| \tag{5}$$

All feature maps of intensity/colors/orientation are normalized separately using $N(\cdot)$, which includes a series of operations that find the location of the global maximum M and the average $\bar{m}$ of all local maxima in the map, normalize the values in the map to $[0, M]$, and multiply the map by $(M - \bar{m})^2$. The feature maps of each channel are combined into a conspicuity map using the across-scale addition operator $\bigoplus$, which downsamples the feature map of different scales to scale-four ($\sigma = 4$) and sums them up pixel by pixel,

$$\overline{I} = \bigoplus_{c=2}^{4} \bigoplus_{s=c+3}^{c+4} N\big(I(c,s)\big),$$
$$\overline{C} = \bigoplus_{c=2}^{4} \bigoplus_{s=c+3}^{c+4} \big[N\big(RG(c,s)\big) + N\big(BY(c,s)\big)\big],$$
$$\overline{O} = \sum_{\theta\in\{0°,45°,90°,135°\}} N\Big(\bigoplus_{c=2}^{4} \bigoplus_{s=c+3}^{c+4} N\big(O(c,s,\theta)\big)\Big).$$

Finally, the conspicuity maps of intensity, color, and orientation are normalized and added to obtain the saliency map

$$S = \frac{1}{3}\big(N(\overline{I}) + N(\overline{C}) + N(\overline{O})\big). \tag{6}$$

Itti et al. [13] further extended the original saliency map to handle videos. They added flicker and motion features into the saliency map. Flicker is calculated from the absolute difference between the luminance I_n of current frame and that I_{n-1} of the previous frame, yielding a flicker pyramid $F_n(s)$, where n and s denote the frame number and scale, respectively. Motion is computed from spatially-shifted differences between Gabor pyramids from the current and previous frames, where the shifted pyramids $S_n(\sigma,\theta)$ is obtained by shifting each Gabor pyramid $O_n(\sigma,\theta)$ one pixel orthogonal to the Gabor orientation. For computational efficiency, motion feature maps can also be computed from intensity pyramids instead of Gabor pyramids at the cost of losing orientation selectivity. Intensity pyramids are used in this paper and a motion pyramid $R_n(\sigma,\theta)$ is computed using the Reichardt model [25],

$$R_n(\sigma,\theta) = \big|O_n(\sigma,\theta) * S_{n-1}(\sigma,\theta) - O_{n-1}(\sigma,\theta) * S_n(\sigma,\theta)\big|,$$

where the subscripts n and $n-1$ denote the current frame and previous frame. The symbol $*$ represents a pixel-wise product.

Similar to the feature and conspicuity maps of other feature channels, the feature and conspicuity maps of flicker and motion are defined as

$$F_n(c,s) = \big|F_n(c) \ominus F_n(s)\big|,$$
$$R_n(c,s,\theta) = \big|R_n(c,\theta) \ominus R_n(s,\theta)\big|,$$
$$\overline{F_n} = \bigoplus_{c=2}^{4} \bigoplus_{s=c+3}^{c+4} N\big(F(c,s)\big), \tag{7}$$
$$\overline{R_n} = \sum_{\theta} N\Big(\bigoplus_{c=2}^{4} \bigoplus_{s=c+3}^{c+4} N\big(R_n(c,s,\theta)\big)\Big),$$

where $c \in \{2,3,4\}$ and $s = c + \delta$ with $\delta \in \{3,4\}$.

4 Our approach

Our tone mapping approach consists of attention map computation and tone mapping adjustment. Given an HDRI, our approach computes its attention map to find the attentive regions and then locally adjusts the tone mapping function to account for the transient attention and adaptation effects [24]. We propose *attention tone mapping function* to enhance the contrast at attentive regions while reducing the contrast at inattentive regions. On the other hand, we also develop *adaptation tone mapping function* to reduce the contrast of an HDRI. This exploits the adaptation effect of HVS since the contrast response of visual neurons decreases when adaptation occurs. Finally, these two tone mapping functions are linearly combined as attention and adaptation effects act independently.

4.1 Attention map computation

To find the attentive and nonattentative regions in an HDR image, we adopt the HDR saliency map proposed by Brémond et al. [2] since it can better predict the attentive regions

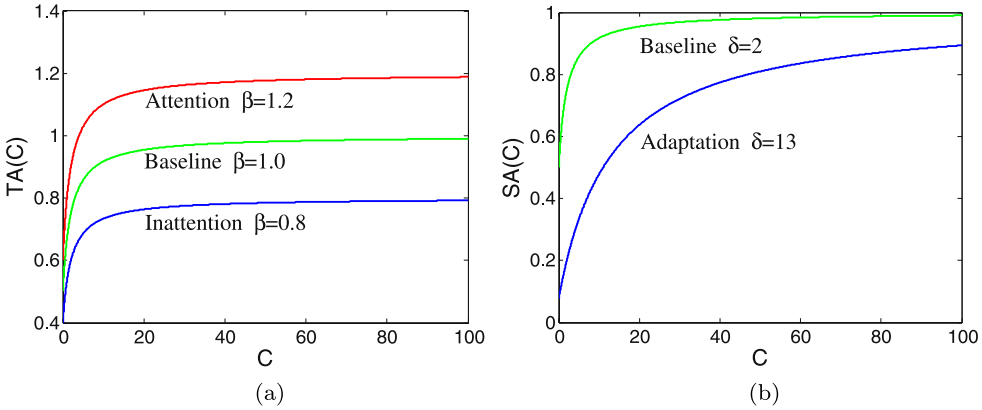

Fig. 2 (**a**) Plot of Eq. 11. We treat $\beta = 1$ as the baseline. *The red* and *blue curve* corresponds to $\beta = 1.2$ and $\beta = 0.8$, respectively. (**b**) Plot of Eq. 13. We set $\delta = 2$ as *the baseline curve* that represents the neutral condition. $\delta = 13$ denotes the curve in adaptation condition

in an HDRI than the saliency map [14]. The HDR saliency map modifies the intensity and orientation channels in the saliency map as follows:

$$I(c,s) = \frac{|I(c) \ominus I(s)|}{I(s)} \quad \text{and} \quad O(c,s,\theta) = \frac{O(c,\theta)}{I(s)}.$$

We then obtain the attention map of an HDRI by normalizing the values of the HDR saliency map to $[-1, 1]$. Figure 3(a) shows the attention map of an HDR image.

4.2 Tone mapping function adjustment

Our tone mapping function is locally adjusted according to the adaptation and attention effects, which are computed from the attention map. The contrast information that is needed in the tone mapping function is computed using the Weber contrast,

$$C(x,y) = \frac{L(x,y) - L_w}{L_w}, \tag{8}$$

where $L(x, y)$ is the luminance at the pixel (x, y) obtained by converting the HDRI from the RGB color space to XYZ space and L_w is the background luminance defined by

$$L_w = \exp\left(\frac{1}{N}\sum_{\forall x,y} \log\big(\varepsilon + L(x,y)\big)\right), \tag{9}$$

where N is the total number of pixels of the image and ε is a small value to avoid the singularity causing by black pixels. For the simplicity of notation, we will omit (x, y) of $C(x, y)$ and $L(x, y)$ in the following text.

To take advantage of the visual neural response to contrast, we adopt the Naka–Rushton function [21] as the base model of our tone mapping function. The Naka–Rushton function is widely used in vision neuroscience. It describes the compressive nonlinear relation between neural response and contrast [1, 23]:

$$R = \frac{a_1 R_{\max} C^{\alpha}}{C^{\alpha} + a_2 (C_{50})^{\alpha}}, \tag{10}$$

where R is the neuron response, C is the contrast of stimulus, $R_{\max}$ is the maximal firing rate of population, α is the slope of the contrast response function, and C_{50} is the contrast required to produce 50% of the neuron's maximal response. Pestilli et al. [23] observed from psychophysical data that transient attention modulates the response gain via the variable a_1 while adaptation modulates the contrast gain through a_2.

To account for the transient attention effect that increases contrast sensitivity at the attentive regions, we define the *attention tone mapping function* as follows:

$$TA(C) = \frac{\beta \cdot (C+1)}{2+C}, \qquad \begin{cases} \beta > 1 & \text{if } A(x,y) > 0 \\ \beta = 1 & \text{if } A(x,y) = 0 \\ \beta < 1 & \text{if } A(x,y) < 0 \end{cases} \tag{11}$$

where β adjusts the tone mapping function according to the value of attention map $A(x, y)$ so that the contrast at the attentive regions is increased. β plays a similar role as a_1 in Eq. 10. Figure 2(a) illustrates the behavior of our attention tone mapping function $TA(C)$. We set $\beta = 1$ as the baseline, which corresponds to the contrast response function in the neutral condition in Fig. 1. The tone mapping function would shift toward the red curve when β gradually increases. This simulates the curve under the attention effect in Fig. 1. If β gradually decreases, the function would move toward the blue curve. This equation behaves like the experimental result in Pestilli et al. [23]. To avoid over or under exposure, we restrict $\beta_{\min} \le \beta \le \beta_{\max}$ by linearly mapping $A(x, y)$ from $[-1, 1]$ to $[\beta_{\min}, \beta_{\max}]$:

$$\beta = 0.5\big((\beta_{\max} - \beta_{\min})A(x,y) + \beta_{\max} + \beta_{\min}\big). \tag{12}$$

Furthermore, as adaptation decreases the contrast response of visual neurons, we can reduce the contrast when adaptation occurs. Specifically, we model the tone mapping function under the adaptation effect as follows:

$$SA(C) = \frac{C+1}{\delta + C}, \tag{13}$$

(a) Attention map (b) Only attention (c) Only adaptation (d) Both ($\delta = 13$) (e) Both ($\delta = 51$)

Fig. 3 Illustration of our tone mapping adjustment: (**b**) is produced only by the attention tone mapping function $TA(C)$; (**c**) is generated only by the adaptation tone mapping function $SA(C)$. (**d**) and (**e**) are results with both tone mapping functions applied. The brightest region (at the ceiling) is clearer in (**e**), which simulates when our eyes adapt to high-luminance light

where δ is used to adjust the tone mapping function based on the adaptation effect. It mimics the function of a_2 in Eq. 10. We call Eq. 13 the *adaptation tone mapping function*. As adaptation is a global response that does not vary much across an image, we set δ a constant in our approach. In fact, our adaptation model is similar to the static response model of adaptation used in [22]. Figure 2(b) plots the function used in our model, where $\delta = 2$ and $\delta = 13$ denotes the curve in the neutral and adaptation condition, respectively. Note that when $\delta = 2$ and $\beta = 1$, Eqs. 11 and 13 are the same baseline function that corresponds to the contrast response at the neutral condition.

As transient attention and adaptation effects act independently and change contrast response function simultaneously [24], the attention tone mapping function $TA(C)$ and the adaptation tone mapping function $SA(C)$ are weightedly combined using

$$R(C) = \omega \cdot TA(C) + (1 - \omega)SA(C). \tag{14}$$

We give the adaptation effect a larger weight by setting $\omega = \frac{1}{3}$ since adaptation affects contrast response function longer than transient attention, which is short-lived. After we obtain $R(C)$, the intensity value of each color channel in an HDRI is mapped to that of an LDRI as follows,

$$[R_d, G_d, B_d] = \frac{R(C)}{L}[R_w, G_w, B_w], \tag{15}$$

where R_d, G_d, and B_d are the RGB value of a pixel in the LDRI; R_w, G_w, and B_w are the RGB value of the corresponding pixel in the HDRI. L is the luminance of the pixel. Figure 3 shows the results when only attention or adaptation tone mapping function is used. The effect of these two tone mapping functions are complementary. The overall contrast is better presented by $TA(C)$ while the luminance at the brightest region are clearly adjusted by $SA(C)$. This demonstrates that both attention and adaptation are needed for optimal visual performance [24]. In particular, the effects of adaptation tone mapping function can be well illustrated in Fig. 3(c) and (e). This mimics the adaptation effect of human eyes as one can see the brightest part more clearly after adaptation but less discrimination for middle-range contrast (details lost).

Our tone mapping algorithm can also be applied to an HDR video by modifying two components. First, we adopt the video saliency map [13], which includes flicker feature map and motion feature map to compute the attention map in an HDR video. Second, we modify the weighting function to address the change of contrast sensitivity due to the temporal property of videos.

$$R(C_t) = \frac{3 + \omega_t}{6} TA(C_t) + \frac{3 - \omega_t}{6} SA(C_t), \tag{16}$$

where C_t is the contrast at a pixel location (x, y) at frame t and ω_t is a time-varying weighting function that is obtained by accumulating the temporal decay of the attention map:

$$\omega_t = \xi A_t + (1 - \xi)\omega_{t-1}, \qquad \omega_0 = A_0, \quad \forall t > 0. \tag{17}$$

A_t is the attention map at frame t. Equation 16 is inspired by the study of Pestilli et al. [24]. They found transient attention plays a more important role than adaptation when people watch a video since transient attention would alter the contrast sensitivity function, which had been optimized by adaptation. Therefore, we design the weights of $TA(C)$ and $SA(C)$ to vary with the time course of transient attention such that the contrast gain decreases gradually when a pixel is no longer attended.

(a) Histogram adjustment (b) Optimization tone mapping

(c) Photographic tone reproduction (d) Our result

Fig. 4 Memorial Church in Stanford University. Radiance map courtesy of Paul Debevec

(a) (b) (c) (d)

Fig. 5 From *left* to *right* are the results of (**a**) histogram adjustment, (**b**) optimization tone mapping, (**c**) photographic tone reproduction, and (**d**) our approach. Radiance map courtesy of Paul Debevec

(a) Histogram adjustment (b) Optimization tone mapping

(c) Photographic tone reproduction (d) Our result

Fig. 6 MtTamWest.hdr (radiance map courtesy of ILM). Our method produces higher contrast in the region of plants and the forest. The contrast of the rocks at front is also maintained

5 Experimental results

HDR images We compare our results with three state-of-the-art tone mapping algorithms: histogram adjustment [15], photographic tone reproduction [27], and optimization tone mapping [19]. We use the code provided in [26] for histogram adjustment and photographic tone reproduction, and the pfstmo package for optimization tone mapping. The parameters of these approaches are tuned to ensure their results are the best performance of these approaches. We set

(a) Histogram adjustment (b) Optimization tone mapping

(c) Photographic tone reproduction (d) Our result

Fig. 7 DaniBelgium.hdr (radiance map courtesy of Karol Myszkowski). The result by histogram adjustment looks washed out. The leaf veins in our result are clearer than those in other results

$\gamma = 2.0$ for those approaches that require gamma correction. In all of our experiments, we set $\beta_{\min} = 0.8$, $\beta_{\max} = 1.2$, $\delta = 13$ and $\omega = \frac{1}{3}$. Figures 4–7 show the comparison results. We highlight the differences of these results in thumbnails below or next to each image.

In Fig. 4, the histogram adjustment operator preserves more details but produces a slightly overexposed image. The result produced by the optimization tone mapping operator places more emphasis on high frequency features but the overall intensity is too dark such that some details are barely visible. The photographic tone reproduction operator produces a brighter result than optimization tone mapping but provides less details and contrast than our method. Figure 5 shows the tone mapping results of the radiance map of a bathroom. Histogram adjustment does not preserve chromaticity well. Optimization tone mapping performs better around the lamp region but produces a dimmer scene. Photographic tone reproduction produces a similar result as ours; however, one can observe that its results around the lamp region are slightly overexposed and the color of wood frame seems too dark.

Figure 6 shows the tone mapping results of an outdoor scene. The result of histogram adjustment demonstrates clear contrast for the rocks at front, but the color of plant, forest, and the lake is not preserved well. Our method produces higher contrast in the region of plants and the forest while the contrast of the rocks at front is also maintained (see the bottom thumbnails). Figure 7 shows the tone mapping results of DaniBelgium.hdr. The result produced by histogram adjustment operator looks washed out. The leaf veins in our result are clearer than those in other results; however, the plant region (the bottom thumbnail) in optimization tone mapping shows higher contrast.

HDR videos We tested our tone mapping algorithm on HDR videos captured by Grzegorz Krawczyk. Adaptation does not affect the contrast response obviously when humans watch a video since it needs longer time to take effect. Therefore, we set $\delta = 3$ rather than $\delta = 13$ in HDR video tone mapping. This makes the adaptation tone mapping function closer to the contrast response at the baseline condition. $\xi = 0.5$ is used in all videos. In the accompanying video, we show a side-by-side comparison of the results produced by photographic tone reproduction and our method. The video result of photographic tone reproduction is generated frame by frame independently. One can find that our

approach demonstrates better contrast while photographic tone reproduction generates a dimmer scene. Besides, the chromaticity is better presented in our result, especially the color of traffic signs, road trees and sunlight on the road.

6 Conclusion and future work

We present a local tone mapping method that considers the attention and adaptation effects in the HVS. In particular, we adopt the HDR saliency map and video saliency map to estimate the attentive regions and propose two contrast adjustment models for transient attention and adaptation based on studies in psychophysics and neuroscience. Our experiments show that our approach can present better contrast and chromatic details than three state-of-the-art tone mapping methods: histogram adjustment [15], photography tone mapping operator [27], and optimization tone mapping operator [19]. Although we use the HDR saliency map and video saliency map to determine the attentive regions, our approach can still be applied if different visual attention models are used. A more sophisticated computational model of attention may further improve our results. In the future, we would like to apply our approach to HDR compression [34, 35].

Acknowledgements This work was supported in part by National Science Council under grant NSC-99-2628-E-009-178- and the UST-UCSD International Center of Excellence in Advanced Bioengineering sponsored by the Taiwan National Science Council I-RiCE Program under Grant Number: NSC-99-2911-I-010-101.

References

1. Albrecht, D.G., Hamilton, D.B.: Striate cortex of monkey and cat: contrast response function. J. Neurophysiol. **48**(11), 217–237 (1982)
2. Brémond, R., Petit, J., Tarel, J.P.: Saliency maps of high dynamic range images. In: Media Retargeting Workshop in Conjunction with ECCV'10 (2010)
3. Cameron, E.L., Tai, J.C., Carrasco, M.: Covert attention affects the psychometric function of contrast sensitivity. Vis. Res. **42**(8), 949–967 (2002)
4. Carrasco, M., Penpeci-Talgar, C., Eckstein, M.: Spatial covert attention increases contrast sensitivity across the csf: support for signal enhancement. Vis. Res. **40**(10–12), 1203–1215 (2000)
5. Chen, H.T., Liu, T.L., Chang, T.L.: Tone reproduction: A perspective from luminance-driven perceptual grouping. CVPR **2**, 369–376 (2005)
6. Chiu, K., Herf, M., Shirley, P., Swamy, S., Wang, C., Zimmerman, K.: Spatially nonuniform scaling functions for high contrast images. In: Graphics Interface, pp. 245–253 (1993)
7. Durand, F., Dorsey, J.: Fast bilateral filtering for the display of high-dynamic-range images. In: Proceedings of SIGGRAPH, pp. 257–266 (2002)
8. Fattal, R., Lischinski, D., Werman, M.: Gradient domain high dynamic range compression. In: Proceedings of SIGGRAPH, pp. 249–256 (2002)
9. Ferwerda, J.A., Pattanaik, S.N., Shirley, P., Greenberg, D.P.: A model of visual adaptation for realistic image synthesis. In: Proceedings of SIGGRAPH, pp. 249–258 (1996)
10. Henderson, J.M.: Object identification in context: the visual processing of natural scenes. Can. J. Psychol. **46**(3), 319–341 (1992)
11. Hickey, C., van Zoest, W., Theeuwes, J.: The time course of exogenous and endogenous control of covert attention. Exp. Brain Res. **201**(4), 789–796 (2010)
12. Hunt, R.W.G.: The Reproduction of Colour. Wiley, New York (2004)
13. Itti, L., Dhavale, N., Pighin, F.: Realistic avatar eye and head animation using a neurobiological model of visual attention. In: Proc. SPIE, vol. 5200, pp. 64–78 (2003)
14. Itti, L., Koch, C., Niebur, E.: A model of saliency-based visual attention for rapid scene analysis. IEEE Trans. Pattern Anal. Mach. Intell. **20**(11), 1254–1259 (1998)
15. Larson, G.W., Rushmeier, H., Piatko, C.: A visibility matching tone reproduction operator for high dynamic range scenes. IEEE Trans. Vis. Comput. Graph. **3**(4), 291–306 (1997)
16. Liang, C.K., Chen, W.C., Gelfand, N.: TouchTone: Interactive local image adjustment using point-and-swipe. In: Eurographics, pp. 253–261 (2010)
17. Ling, S., Carrasco, M.: When sustained attention impairs perception. Nat. Neurosci. **9**(10), 1243–1245 (2006)
18. Lischinski, D., Farbman, Z., Uyttendaele, M., Szeliski, R.: Interactive local adjustment of tonal values. SIGGRAPH **25**(3), 646–653 (2006)
19. Mantiuk, R., Daly, S., Kerofsky, L.: Display adaptive tone mapping. ACM Trans. Graph. **27**(3), 1–10 (2008)
20. Martínez-Trujillo, J.C., Treue, S.: Attentional modulation strength in cortical area mt depends on stimulus contrast. Neuron **35**(2), 365–370 (2002)
21. Naka, K.I., Rushton, W.A.H.: S-potentials from luminosity units in the retina of fish (cyprinidae). J. Physiol. **185**(3), 587–599 (1966)
22. Pattanaik, S.N., Tumblinand, J., Yee, H., Greenberg, D.P.: Time-dependent visual adaptation for fast realistic image display. In: Proceedings of SIGGRAPH, pp. 47–54 (2000)
23. Pestilli, F., Ling, S., Carrasco, M.: A population-coding model of attention's influence on contrast response: Estimating neural effects from psychophysical data. Vis. Res. **49**(10), 1144–1153 (2009)
24. Pestilli, F., Viera, G., Carrasco, M.: How do attention and adaptation affect contrast sensitivity? Journal of Vision **7**(7) (2007)
25. Reichardt, W.: Evaluation of optical motion information by movement detectors. J. Comp. Physiol. **161**(4), 533–547 (1987)
26. Reinhard, E.: High Dynamic Range Imaging.: Acquisition, Display, and Image-Based Lighting. Kaufmann, Los Altos (2005)
27. Reinhard, E., Stark, M., Shirley, P., Ferwerda, J.: Photographic tone reproduction for digital images. In: Proceedings of SIGGRAPH, pp. 267–276 (2002)
28. Seetzen, H., Heidrich, W., Stuerzlinger, W., Ward, G., Whitehead, L., Trentacoste, M., Ghosh, A., Vorozcovs, A.: High dynamic range display systems. SIGGRAPH **23**(3), 760–768 (2004)
29. Treisman, A.M., Gelade, G.: A feature-integration theory of attention. Cogn. Psychol. **12**(1), 97–136 (1980)
30. Tumblin, J., Rushmeier, H.: Tone reproduction for realistic images. IEEE Comput. Graph. Appl. **13**(6), 42–48 (1993)
31. Tumblin, J., Turk, G.: LCIS: a boundary hierarchy for detail-preserving contrast reduction. In: Proceedings of SIGGRAPH, pp. 83–90 (1999)
32. Van Hateren, J.H.: Encoding of high dynamic range video with a model of human cones. ACM Trans. Graph. **25**(4), 1380–1399 (2006)
33. Ward, G.: A contrast-based scalefactor for luminance display. In: Graphics Gems IV, pp. 415–421. Academic Press, San Diego (1994)

34. Ward, G., Simmons, M.: JPEG-HDR: a backwards-compatible, high dynamic range extension to jpeg. In: Proceedings of the 13th Color Imaging Conference (2005)
35. Xu, R., Pattanaik, S.N., Hughes, C.E.: High-dynamic-range still-image encoding in jpeg 2000. IEEE Comput. Graph. Appl. **25**(6), 57–64 (2005)

Wen-Chieh Lin received the BS and MS degrees in control engineering from the National Chiao Tung University, Hsinchu, Taiwan, in 1994 and 1996, respectively, and the PhD degree in robotics from Carnegie Mellon University, Pittsburgh, in 2005. Since 2006, he has been with the Department of Computer Science and the Institute of Multimedia Engineering, National Chiao Tung University, as an assistant professor. His current research interests include computer graphics, computer animation, and computer vision. He is a member of the IEEE and the ACM.

Zhi-Cheng Yan received the BS and MS degrees from the Department of Management Information Systems, National Chengchi University, and the Institute of Multimedia Engineering, National Chiao Tung University, Taiwan, in 2008 and 2010, respectively. He is interested in Computer Graphics, Image Processing, Computer Vision, and Software Engineering. He has been a software engineer in ASUSTeK Computer Inc. since 2010.