Gaze-Dependent Tone Mapping

Radosław Mantiuk and Mateusz Markowski

West Pomeranian University of Technology, Szczecin,
Faculty of Computer Science,
Żołnierska 52, 71-210, Szczecin, Poland,
`rmantiuk@wi.zut.edu.pl`,
WWW home page: `http://rmantiuk.zut.edu.pl/`

Abstract. In this paper we model the process of temporal adaptation of the human visual system to varying luminance conditions. An eye tracker is used to capture the location of an observer's gaze in a high dynamic range image displayed on the screen. We apply a novel technique of eye tracker data filtering to avoid flickering caused by incorrect gaze estimation. Temporary adaptation luminance is then determined in the area surrounding the gaze point. We use its value to compress the high dynamic range image and display it on the low dynamic range display. The applied tone mapping technique uses a global compression curve in which location is shifted along the luminance axis according to a value of the adaptation luminance. This technique models the natural process of adaptation occurring in the human eyes, also taking into account the time-dependent visual adaptation to dark and bright backgrounds.

Keywords: gaze-dependent tone mapping, eye tracking, luminance adaptation, tone mapping operators, high dynamic range images

1 Introduction

Human visual system (HVS) is sensitive to contrast in luminance range from a $10^{-6}cd/m^2$ (objects viewed under illumination from the stars) to $10^{8}cd/m^2$ (objects viewed on a bright sunny day) [7]. However, momentary dynamic range is limited to 4 orders of magnitude. In this range of luminance, a human can see details and at the same time she/he experiences lower luminance as noise and higher luminance as over-saturated areas. To extend the dynamic range, the HVS has the ability to adapt to changes in the ambient luminance and move the "detailed vision" window along the luminance range of a scene. Interestingly, it mainly adapts to an area covering approximately one degree of the viewing angle around the gaze direction [14]. Other areas of the scene, observed not in foveal but in para-foveal and peripheral regions, have significantly less impact on the adaptation level, although, a human frequently changes his gaze direction (even a hundred times per second) and tries to adapt to different regions. As the process of the luminance adaptation is slower than changes of gaze direction, the HVS is permanently in the maladaptation state, in which the adaptation luminance is changing towards a target value but never reaches this value because in the meantime the target is changed.

In this work we dynamically reproduce the process of maladaptation using the scene data stored in a high dynamic range (HDR) image. We use an eye tracker to capture the gaze direction of a human observer. Then, we compute the temporal adaptation luminance and use its value to display the HDR image on the low dynamic range (LDR) display. The global tone mapping parametrized by the temporal adaptation luminance can be applied to HDR images to achieve perceptually correct and plausible tone compression results. For every pixel in the HDR image we use the same compression curve which changes over time, following the new values of the adaptation luminance. We called this technique a gaze-dependent tone mapping operator (GDTMO).

Advanced local tone mapping algorithms apply different compression curves for every pixel in an HDR image. These techniques preserve visibility of details in the output LDR image, although the excess of details can be distracting for an observer and interpreted as unnatural. The GDTMO technique reveals all the details in the HDR image and at the same time retains noisy and over-saturated areas at the peripheral region of vision.

A practical contribution of this paper is a novel technique for gaze data filtering. Eye trackers introduce significant inaccuracies during estimation of the gaze direction, leading to incorrect computation of the adaptation luminance and causing flickering of the compressed LDR image. We propose a technique which maximises the probability of choosing a correct luminance level based on a statistical analysis of luminance values in gaze point surrounding.

In Sect. 2 a short information on the eye tracking is presented together with a review of the previous work on gaze-dependent tone mapping. In Sect. 3 we explain details of our tone mapping model and present its real time implementation. In Sect. 4 a real time implementation of GDTMO and the results of user studies are presented. We conclude the paper in the last section.

2 Background and Related Works

Advanced *tone mapping techniques* were developed to minimise the deterioration of brightness, contrast and/or perceptual appearance of the HDR images displayed on the LDR displays (see [7] for comprehensive reviews). However, most of the TMOs do not take into account the temporary adaptation to luminance caused by changes of view direction. In this work we concentrate on techniques that reproduce this characteristic behaviour of the HVS [5].

Devices called *eye trackers* capture two types of eye movements, *saccades* and *smooth pursuits* [2, 11]. A smooth pursuit is active when eyes track a moving target and are capable of matching its velocity. A saccade represents a rapid eye movement used to reposition the fovea to a new location, which lasts from 10 ms to 100 ms. The main goal of the eye tracking is to capture a single location an observer intends to look at. This process is known as a *visual fixation*. A point of fixation can be estimated as a location where saccades remain stable for a 200-400 msec on the most significant areas of an image (called Region-of-Interest, ROI) [4]. A typical remote eye tracker is capable of estimating the gaze point on a display screen with accuracy close to 1 degree of the viewing angle (a circular region of roughly 80 pixels in diameter observed on a 22 inch display of 1960x1050 pixel resolution from a 60 cm distance). However, the accuracy is prone to higher temporal errors, even when using the state-of-the-art fixation techniques like Velocity-Threshold Identification (I-VT) or Dispersion Threshold Identification (I-DT). Hence, we propose a technique of gaze data filtering leading to estimation of the actual fixation region, or more precisely permitting accurate computation of the adaption luminance for this region (see Sect. 3.1).

To the best of our knowledge, the concept of tone compression using an eye tracker was first shown during the SIGGRAPH Asia 2009 emerging technologies exhibition [3], however details have not been published. In [1] a gaze-dependent approach based on Reinhard's photographic operator [18] is proposed. The authors examine the use of ROI subtended 2, 4, and 10 degrees of the viewing angle to compute the local adaptation luminance. This value is then replaced with the logarithmic average from the original compression equation. The results show that this concept is scene-referred [10] and cannot be applied as a general tone mapping model. Contrary to this solution, we estimate adaptation luminance based on a statistical analysis of the eye tracker data and use a perceptual maladaptation model of luminance adaptation (see Sect. 3). A similar gaze-dependent approach applied to different image analyses and visualisation techniques, was proposed to determine the location of the accommodation plane in the depth-of-field simulation [9] or to reduce computations in the real time ambient occlusion algorithm [8].

3 Gaze-Dependent Tone Compression

The overview of the proposed gaze-dependent TMO system is shown in Fig. 1. For every pixel in the input HDR image, a map of the background adaptation luminance (L_{ba}) is computed. We accomplish this task by applying a Gaussian filter with the kernel size covering one degree of the viewing angle to the HDR image luminance map. In the main processing loop, the gaze direction is captured by tracker, filtered (see Sect. 3.1) and then used to compute the temporary adaptation luminance (L_a) (see Sect. 3.2). This adaptation luminance is used to compute the tone compression curve and compress the HDR image (see Sect 3.3). Finally, the R, G, and B colour values are computed, desaturated using the exponential rule [17] and the LDR image is displayed on the screen.

3.1 Gaze data filtration

The L_{ba} map is analysed in the area subtended by one degree of the viewing angle around the temporary gaze direction. We have taken that value because it is close to the average accuracy of our eye tracker. The luminance range occurring in the area is divided into a fixed number of levels, then the number of pixels

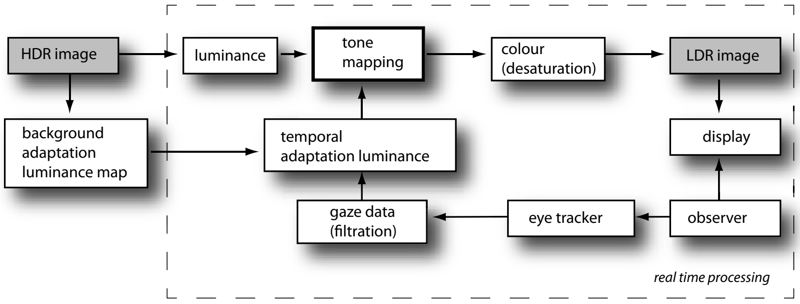

Fig. 1. The diagram of the GDTMO system.

belonging to each level is computed. It is assumed that an observer has adapted to the most numerous level in a relevant area and this level is recognised as the adaptation luminance.

As the eye tracker error creates a normal distribution around the gaze point [6], the distance from a pixel to the gaze point could also be taken into account. However in the conducted pilot studies there was no noticeable improvement in performance using the distance-weighted sum.

To compute the temporary adaptation luminance in real time we analyse a 40x40 pixels area (corresponding to 1° of the viewing angle) around the gaze position captured by the eye tracker. For every luminance level occurring in this area, a separate texture of a size equal to this area is created. GPU shader program sets all the pixels in this texture corresponding to the same luminance level (read from the L_{ba} map) to a value of one, while the remaining pixels are set to zero. After further procession the textures are down-sampled to the size of one pixel using bilinear interpolation. The value of this pixel multiplied by the number of down-sampling steps indicates the number of pixels belonging to a luminance level.

3.2 Temporal adaptation to light and dark

The luminance conditions can change drastically over time. The HVS reacts to such changes through the temporal adaptation process, which differs depending on whether we adapt to light or to darkness, and whether we perceive mainly using rods (during night time) or cones (during day time). In the tone mapping algorithm, the adaptation can be modelled by varying the adaptation luminance value over time. Ferwerda et al. [16] presented a computational model of changes in luminance sensitivity over time using Ward's scaling tone mapping approach [15]. Pattanaik et al. [12] proposed photoreceptor-based, time dependent, retinal adaptation mechanisms for both cones and rods.

In this work, we apply a simplified model using an exponential function (following [13]):

$$\widehat{L}_a^{new} = \widehat{L}_a + (\widehat{L}_{HDR} - \widehat{L}_a)(1 - e^{-\frac{T}{\tau}}), \tag{1}$$

where L_a and L_a^{new} is the current and new adaptation luminance respectively, L_{HDR} denotes the temporary luminance in the area surrounding the gaze point, $\widehat{L} = log_{10}(L)$. T is the time step between the display of two frames, and the τ constant describing the speed of the adaptation process. This speed is different for cones and rods; to account for this, we use the following interpolation:

$$\tau(L_a) = \sigma(L_a) * \tau_{rod} + (1 - \sigma(L_a)) * \tau_{cone}, \tag{2}$$

where $\tau_{rod} = 0.4\ sec$, $\tau_{cone} = 0.1\ sec$, and σ models the sensitivity of rods with the following equation:

$$\sigma(L_a) = \frac{0.04}{0.04 + L_a}. \tag{3}$$

The presented model describes adaptation to light, in which the observer adapted to a dark area turns his eyes to the bright areas of the HDR image. For practical reasons, and following [13], we simulate the reverse process - adaptation to the dark as the linear decrease of adaptation luminance. Additionally, we accelerate the process by two orders of magnitude. In the real world the adaptation to dark takes up to tens of minutes and preserving perceptual correctness would be impractical in our tone mapping application. After quick adaptation to light (about a second), one would wait for more than a minute to see any changes caused by the adaptation to dark [12].

3.3 Tone compression

For HDR tone compression we apply the modified Naka-Rushton equation [22], which models the intensity-response function of distal retinal neurones:

$$L_{LDR} = \frac{L_{HDR}}{L_{HDR} + s}, \tag{4}$$

where L_{LDR} denotes the output low dynamic range luminance, L_{LDR} is a luminance of HDR pixels. We replace the semi-saturation constant s from the original equation with the the gaussian function:

$$s = \mathit{\Gamma}(L_a) = 1.002 * exp(\frac{0.949 - L_a}{6.806})^2, \tag{5}$$

which approximates the "curve-shifting" of the response [19]. A curve is chosen for which half of the maximum response relates to $\mathit{\Gamma}(L_a)$ luminance value. We determined this function by applying the experimental data of the rods' and cones' sensitivity presented in [12].

The mechanism of visual adaptation is responsible for constant contrast perception in varying luminance conditions and in a direct way can be adapted for compression of HDR values. The Naka-Rushton equation correctly mimics the perceived brightness in a maladaptation state because it was modelled in an experiment with pulsing stimuli [20]. The same situation occurs while viewing real-world scenes, as we frequently use saccadic movements to look for and track objects. The stimuli fields are not steady but pulsed.

4 Results

We implemented our GDTMO system in Matlab using a set of GPU shaders controlled by the Psychtoolbox [1] library. The shaders are responsible for tone compression, colour desaturation, and gaze point filtration. The software is executed on a 2.8 GHz Intel i7 930 CPU equipped with an NVIDIA GeForce 580 GTI 3072MB graphics card. To capture gaze direction we use a professional 250 Hz P-CR RED250 eye tracker controlled by the proprietary SMI iViewX software (version 2.5) running on a dedicated PC. RED250 eye tracker was mounted under a 22 inch Dell E2210 LCD display with the screen dimensions 47.5x30 cm, and the native resolution 1680x1050 pixels (60Hz).

Fig. 2. Images viewed by an observer adapted to different areas of the scene indicated by the white crosses. In the left image the adaptation luminance is close to 1790 cd/m^2, in the right to 5 cd/m^2.

In Fig. 2 example screenshots from the GDTMO session are presented. In the conducted pilot user study we asked observers about their impression of the HDR viewing. The general conclusion is that observers appreciate fidelity to reality and a ”joy to use” of the system. The limited length of this article does not allow us to describe the results of perceptual experiments so we address this evaluation to future work.

In Fig. 3 we present example results of gaze-dependent estimation of the adaptation luminance. Using the raw gaze data results in frequent changes of adaptation values (the solid line in the plot) and causes image flickering. The gaze data filtration together with the temporary adaptation smooth the data (the dashed line). Steep parts of the curve denote the adaptation to bright areas while gentler slopes show the adaptation to dark. The curve is consistent with the behaviour of HVS [12].

---

[1] http://www.psychtoolbox.org/HomePage

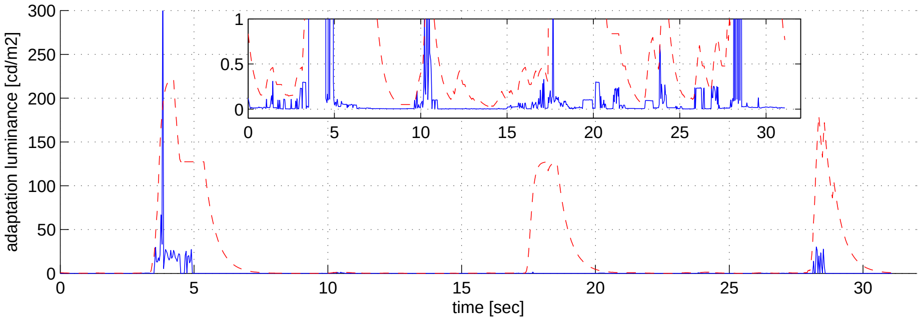

Fig. 3. Plot of the adaptation luminance values determined based on the raw gaze points (solid line) and our filtration technique (dashed line). The inset shows values close to zero. The data was collected during the actual eye tracking session.

5 Conclusions and Future Work

In this work we have shown how using information about gaze direction can be applied to a real time tone mapping of HDR images. The proposed model is based on temporary adaptation to luminance and perceptual compression of the luminance range. It is notable that the presented GDTMO algorithm filters gaze data to counteract inaccuracies introduced by the eye tracking hardware. In future work, we will introduce our technique into the tone mapping of HDR video, considering optimised processing of the HDR data [21]. We also plan to explore the spatial properties of the mechanism of luminance adaptation.

Acknowledgements This work was supported by the Polish Ministry of Science and Higher Education through the grant no. N N516 508539, and partially supported by the COST Action IC1005.

References

1. Mikami, T., Hirai, K., Nakaguchi, T., Tsumura, N.: Real-time tone-mapping of high dynamic range image using gazing area information. In: Proc. International Conference on Computer and Information (2010)
2. Duchowski, T. A.: Eye Tracking Methodology: Theory and Practice (2nd edition). Springer, 2007
3. Rahardja, S., Farbiz, F., Manders, C., Zhiyong, H., Ling, J.N., Khan, I.R., Ping, O.E., Peng, S.: Eye HDR: gaze-adaptive system for displaying high-dynamic-range images. In ACM SIGGRAPH ASIA 2009 Art Gallery & Emerging Technologies, pp. 6868 (2009).
4. Salvucci, D.D., Goldberg, J.H.: Identifying fixations and saccades in eye-tracking protocols. In: Proceedings of the 2000 symposium on Eye tracking research & applications (ETRA), New York, pp. 7178 (2000)

5. Mantiuk, R. K., Krawczyk, G., Mantiuk, R., Seidel, H. P.: High dynamic range imaging pipeline: Perception-motivated representation of visual content. In: Proc. of Human Vision and Electronic Imaging XII, Vol. 6492, No. 649212, (2007).
6. Mantiuk, R., Bazyluk, B., Mantiuk, R.K.: Gaze-Dependent Object Tracking for Real Time Rendering. Computer Graphics Forum (Proc. of Eurographics'13), (2013)
7. Reinhard, E., Ward, G., Pattanaik, S., Debevec, P.: High Dynamic Range Imaging. Data Acquisition, Manipulation, and Display. Morgan Kaufmann, 2005
8. Mantiuk, R. and Janus, S.: Gaze-dependent Ambient Occlusion. Lecture Notes in Computer Science (Proc. of ISVC'12 Conference), V.7431, No.1, pp. 523-532 (2012).
9. Mantiuk, R. and Bazyluk, B. and Tomaszewska, A.: Gaze-Dependent Depth-of-Field Effect Rendering in Virtual Environments. Lecture Notes in Computer Science (Proc. of SGDA 2011), V. 6944, pp.1-12 (2011).
10. Yamauchi, T. and Mikami, T. and Ouda, O. and Nakaguchi, T. and Tsumura, N.: Improvement and evaluation of real-time tone mapping for high dynamic range images using gaze information. Proceedings of the 2010 international conference on Computer vision - Volume Part I, pp. 440-449, (2011)
11. Mantiuk, R., Kowalik, M., Nowosielski, A., Bazyluk. B.: Do-It-Yourself Eye Tracker: Low-Cost Pupil-Based Eye Tracker for Computer Graphics Applications. Lecture Notes in Computer Science (Proc. of MMM'12 Conference), Vol. 7131, pp. 115-125 (2012).
12. Pattanaik, S. N., Tumblin, J., Yee, H., and Greenberg, D. P.: Time-dependent visual adaptation for fast realistic image display. In proc. of SIGGRAPH 00, pp. 4754 (2000).
13. Krawczyk, G. and Myszkowski, K. and Seidel, H.P.: Perceptual effects in real-time tone mapping. Proceedings of the 21st spring conference on Computer graphics, Budmerice, Slovakia, pp. 195-202 (2005)
14. Hood, D. C.: Lower-level visual processing and models of light adaptation. Annu. Rev. Psychology, 49, pp. 503-535 (1998)
15. Ward, G.: A contrast-based scalefactor for luminance display. Graphics Gems IV, pp. 415 421 (1994)
16. Ferwerda, J. A., Pattanaik, S. N., Shirley, P., and Greenberg, D. P.: A model of visual adaptation for realistic image synthesis. Proc. of SIGGRAPH'96, ACM, New York, NY, USA, pp. 249258 (1996)
17. Mantiuk, R., Mantiuk, R.K., Tomaszewska, A., Heidrich, W.: Color Correction for Tone Mapping. Computer Graphics Forum (Proc. of Eurographics'09), Vol. 28, No. 2, Germany, 2009, pp. 193-202
18. Reinhard, E., Stark, M., Shirley, P., Ferwerda, J.: Photographic tone reproduction for digital images. ACM Transactions on Graphics 21(3), 267276 (2002)
19. Shapley, R., Enroth-Cugell, C.: Visual adaptation and retinal gain controls Progress in Retinal Research, vol. 3, pp. 263-346 (1984)
20. Ferradans, S., Bertalmio, M., Provenzi, E., Caselles, V.: An Analysis of Visual Adaptation and Contrast Perception for Tone Mapping. IEEE Trans. Pattern Anal. Mach. Intell., vol. 33 (10), pp. 2002-2012 (2011).
21. Mantiuk, R., Pajk, D.: Acceleration of high dynamic range imaging pipeline based on multi-threading and SIMD technologies. Lecture Notes in Computer Science (Proc. of ICCS'08 Conference), Vol. 5101, pp. 780-789 (2008).
22. Naka, K.-I., Rushton, W. A. H.: S-potentials from luminosity units in the retina of fish (Cyprinidae). J. Physiol. 185, pp. 587-599 (1966).