PARIPEX - INDIAN JOURNAL OF RESEARCH
Volume-8 | Issue-2 | February-2019 | PRINT ISSN - 2250-1991

ORIGINAL RESEARCH PAPER

Engineering

REALISTIC HDR TONE-MAPPING BASED ON CONTRAST PERCEPTION MATCHING

KEY WORDS: Realistic Tone-Mapping, Contrast Perception, Contrast Reaction Time

Dr. A.Senthil Arumugam
Associate Professor,Department of Information Technology,Sri Ram Nallamani Yadava College of Arts and Science,Manonmaniam Sundaranar University, Tirunelveli-12,Tamilnadu, India

D. Paulin Diana Dani*
Reserch Scholar, Department of Information Technology,Sri Ram Nallamani Yadava College of Arts and Science, Manonmaniam Sundaranar University,Tirunelveli-12, Tamilnadu, India *Corresponding Author

ABSTRACT

In this paper, a new method to measure human's reaction time on local contrast is introduced. It is believed the reac- tion time is highly correlated with the response strength on local contrasts in human visual system. Using the proposed method, a large set of subjective viewing data were then col- lected, and a full-range (from near-threshold to suprathresh- old) contrast perception model on low dynamic range dis- play is built. Based on the proposed model and Plainis' reac- tion time model [17], a new realistic tone-mapping operator, which minimizes the perception difference of local contrasts between HDR and LDR images, is proposed. Experimental results confirmed the performance of the proposed operator.

1. INTRODUCTION

The topic of tone-mapping has been extensively studied in the last decade and many algorithms have been developed. How- ever, developing realistic tone-mapping is still a very chal- lenging task and to achieve a good processed image is only achievable via interactive methods. The reason is that tone- mapping operators should not only ensure that the resulting pixel values are in the restricted range, but also at the same time minimize the perception difference between input HDR and output LDR images. The perception includes both global tone perception and local contrast perception. Obviously, ac- curate perception models are needed to achieve the realistic requirement. Between the two perception mechanisms, it was found that human visual system (HVS) relies more on the contrast perception than tone perception [1], which can be proved by the observation that humans can easily detect and recognize objects under wide luminance range. So, most studies were focusing on the contrast perception.

Majority of tone mapping operators adopt Contrast Sen- sitivity Function (CSF) or visibility threshold models, which represent the relationship among visibility of local contrast, and the characteristics of itself and surroundings. Visibility models were first applied in image enhancement. In 1994, Ward [8] proposed an adaptive scale-factor algorithm to im- age enhancement. The algorithm makes use of Blackwell's

luminance difference threshold model [9], which describes the relationship between the background luminance and the visibility threshold in the luminance. In the same year, Ji et al [10] also developed an adaptive contrast enhancement method based on a similar contrast visibility model of Black- well's. Subsequently, a visibility matching based HDR tone- mapping technique is proposed by Ward Larson et al [11] that ensures the visibility of the contrasts in tone-mapped LDR images is close to those in HDR images. In Volevich's work [13], Daly's Visual Difference Predictor (VDP) [12], which is another visibility model, is embedded. However, CSF and other visibility threshold models are built based on the detectability of contrasts. They are effective in near- threshold range, but most contrasts in natural images cover up to suprathreshold range. So, in recent Mantiuk's paper [6], a visibility threshold model is extended from near-threshold to suprathreshold contrast range using Wilson's 'S' curve [14].

As mentioned, an accurate suprathreshold model is necessary for realistic tone-mapping. However, study on suprathreshold contrast perception is very limited. The rea- son is that it is difficult to design a method that can reliably used to study human's contrast perception in suprathreshold range. The common methods are Masking [14], Contrast- matching paradigm [15] and Reaction time [16, 17, 18]. The former two methods are both indirect methods, which means the human's response strength of a contrast stimulus is not measured directly. In masking experiments, subjects are typically required to discriminate between two overlapped sine-waves. Comparatively, Contrast-matching paradigm method is a very popular method used to model HVS. In Peli's subjective experiments [15], viewing subject is required to adjust the physical contrast of one patch against the other at different background luminance and spatial frequencies, until the subject believes the two patches are perceptually the same. High-level subjective comparison activities are involved in these two methods. The high-level activities introduce personal variations on test results, which limit the accuracy of the above-mentioned models. The last method measures the reaction time of HVS on contrast stimulus, which can be regarded as a direct measurement. Statistically, it is believed the reaction time has a high correlation with re- sponse strength, or saliency of a stimulus. Short reaction time means the contrast is easy to detect, and has a high percep- tual strength, and long reaction time refers to low perceptual strength. Using manual measuring, a number of reaction time models are generated [16, 17, 18]. Pie´ron's model is only subjective to the intensity of stimulus and is represented as a power function. Vassilev et al [18] proposed a model that takes in both intensity of contrast and spatial frequency, and Plainis's model [17] showed a model on intensity, spatial frequency and background luminance in real-world measure- ments. Manual measuring introduce variations, which reduce the accuracy of models.

In this work, we use eye-tracker to measure the reaction time automatically by saccadic eye-movement detection, in which the outside influences are minimized. Because the detection of contrast stimulus only triggers the bottom-up mechanisms of visual attention in human brain, no high-level recognition or other brain activity is involved. The subjective influence is also minimized. A contrast reaction time model on LDR display is constructed using the method. Utilizing Plainis's model, the contrast perception difference between HDR image and tone-mapped LDR image can be matched and minimized.

This paper is organized as follow: Section 2 introduces the design the subjective experiment and the proposed con- trast perception model. The proposed tone-mapping operator and experimental results are presented in Section 3 and 4, re- spectively. Section 5 gives the conclusions.

2. CONTRAST PERCEPTION MODEL

The subjective experiment is designed to be a simple vi- sual search task. It uses a Tobii-1750 eye-tracker system (http://www.tobii.com) to record eye-movements. Basically the monitor can be regarded as a general low-cost LDR (0-255 grayscale) display device.

In the experiment, the viewing subject is required to detect and locate a local single contrast stimulus by his/her eye-gaze

in the fastest way, and stays there for at least 2 seconds. The contrast stimuli have a sinusoidal waveform in a fixed size. The viewing test sequences are pre-generated in a spatial res- olution of 1280 × 1024 (display resolution) in domain of $Y\ C\ bCr$ color space, which means only luma-contrast spatial masking effect is taken into consideration in the experiment. The frame rate of viewing test sequence is set at 5Hz. Every test sequence contains 10 test points, and each test point is 10 seconds long. Before each test point, there is a blank with $0 \le \triangle t \le 1$ second, which is randomly assigned to reduce the influence of periodical appearance of contrast stimuli in a test sequence. The single stimulus lasts 9 – seconds. And after that, there is a one-second blank to remove the af- terimage of the stimulus in HVS. Ten seconds blanks are also left before and after the 10 test points, so each sequence is 2 minutes long. The characteristics of contrast stimulus are assigned randomly from 7 pre-defined physical contrasts, 5 frequencies, 4 orientations and 8 background brightness val- ues. The location of the stimulus is also assigned with a ran- dom position to the previous stimulus with a minimal con- straint on distance. The background brightness value in each test sequence is fixed to prevent the influence of brightness change. In the experiment, eye-display calibration tasks were performed before every test group to ensure higher accuracy of eye-tracking. Each group contains 3-5 test sequences. Be- tween every test sequences, the subject would have a 1-2 min- utes' on-site rest; and between every test groups, the subject would have a 10-15 minutes' rest. Every day, each subject would view at most 10 test groups to prevent tiredness, which may have a serious influence on detection efficiency. All these constraints were used to ensure good accuracy of the experi- mental results.

Up to now, two subjects, who are also the authors of this paper, attended the experiment. They are both physically and mentally healthy males, in their ages of 37 and 40, re- spectively. They both have normal or correct-to-normal eye- sight. In total, 24120 test points were collected. Among them, 21032 test points are valid experimental results. The others are discarded because: (1) eye blinks during the visual search stage; or (2) initial eye location is too close to stimulus (distance below a predefined value); or (3) no saccadic eye movement is detected. To achieve a comfortable viewing, the maximal luminance of the monitor is set to about $300 cd/m^2$ and gamma value is set to 1.8. Both are maximized. The test was conducted in an enclosed environment with normal and comfortable fluorescent ceiling light. The viewing distance is about 1.5× of screen height, which is within the effective range of eye-tracking. In reaction time measurement, the last saccadic eye movement before it reaches the stimulus is used as the timeline of human to locate the target stimulus. In this work, saccadic eye movement is defined as a big direction change and/or a big acceleration of eye movement.

By data fitting, the proposed equation for contrast reaction time model is given by:

$$RT_{LDR} = a_1 \cdot\ e^{a2} \quad {}^{3} \quad {}^{4} \quad {}^{5}\ {}_{index} + a_6 \qquad (1)$$

where $\{\alpha_i\ : i = 1, \cdots, 6\}$ are model parameters. RT_{LDR} de- notes reaction time in the measurement of second. denotes the background brightness in 256 grayscale. C = B_{MAX} – denotes physical contrast, in which B_{MIN} and B_{MAX} de- note the minimal and maximal brightness value on stimulus, respectively. B_{INDEX}= | B– 103 | denotes a brightness index, and the number 103 is estimated from the response curve of the display. $\in (0, 8]$ denotes frequency. In the experiment, frequency index are obtained using a 16 × 16 2D DFT. For example, = 1 represents a frequency of 16 pixel/cycle, and F = 8 represents a frequency of 2 pixel/cycle. denotes the distance between the stimulus and eye-gaze location when the stimulus appears. It is worth noting that equation 1 is mono- tonic to contrast , which means a single physical contrast value can be recovered from a given by the inverse function of equation 1.

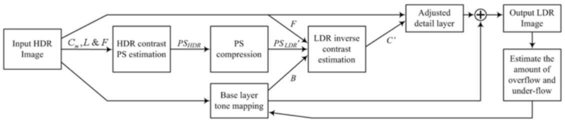

Fig. 1. Diagram of the proposed tone-mapping operator.

As mentioned above, contrast perception strength has an inverse relationship to reaction time, so we have:

$$P\ S_{LDR} = -\ log(RT_{LDR} + C_0) \qquad (2)$$

where C_0 = 0 is a constant in the current work.

3. REALISTIC TONE-MAPPING

The diagram of the proposed tone-mapping operator is pre- and , the target physical contrast value can be sented in Fig. 1. First, the contrast perception values of the input HDR image is obtained via Plainis's reaction time model [17], which is expressed as:

$$\underset{1}{RT} \quad = RT\ + b\ \cdot\ F - b\ \cdot\ \log L + \cdot \quad \cdot\ 10^{(b_3 + b_4 \cdot\ F - b_5} \quad \cdot\ \log\ L) \quad C_m$$

where D_i : i = 1, ⋯ , 5 are parameters. $R1_{HDR}$ denotes reaction time in the measurement of millisecond. RT_0 is a constant. denotes real-world luminance. and denotes local Michelson contrast:

$$C_m = \frac{L_{max} - L_{min}}{L_{max} + L_{min}}$$

where L_{MAX} and L_{Min} denote maximum luminance and min- imum luminance, respectively.

Similar, we have $P\ S_{HDR}$ = - log(RT_{HDR}). Because of the different subjective experiment setups, measure meth- ods, fitting algorithms, and maybe some unknown reasons, values obtained from Plainis's model need to be compressed to achieve a pleasant view in LDR images. A linear compres- sion function is used to maintain the relative distribution of the values.

It should be noted that a fixed distance of D = 400 pixels is used on equation 1 in the scheme to reduce the computa- tional cost. The prediction of eye scanning order in the ap- pearance of an image is very costly, subjective and complex. And research reveals that people generally only locate their eye fixations to a number of high attentional regions on an image, and comparatively ignore the other parts. D = 400 pixels is a distance in a level between foveal and peripheral vision, thus it is selected.

Like most current tone-mapping operators, the input HDR image is separated into a smooth base layer and a detail layer for contrast compression. Farbman's edge-preserving smoothing operator [7] is adopted in the work with parameters λ = 2.0 and = 1.75. The base layer is compressed using Reinhard's global operator [3], where the background brightness values are obtained from. It should be noted that other global tone management techniques also can be adopted. The comparison and selection of the techniques will be included in the next step work. With the compressed

$P\ S_{LDR}$ and F , the target physical contrast value C can be calculated using the inverse function of equations 1 and 2. Then, the local contrast of detail layer can be adjusted by C

which is a matching of the contrast perception in HDR image. The adjusted detail layer and based layer are then combined into the output LDR image. To avoid over- and underflow of LDR image, the amount of the over- and underflow pixels are calculated and used as a feedback to adjust the range of base layer. No input parameters are required in the proposed operator.

4. EXPERIMENTAL RESULTS

A comparison is given in Fig. 2. Fig. 2(a)-© are generated by Qtpfsgui 1.8.12 (http://qtpfsgui.sourceforge.net/) using Reinhard's [2], Fattal's [4] and Mantiuk's operator [5]. Fig. 2(d) is the result generated by the proposed tone-mapping operator. Considering HDR display is not available on market, subjective comparison with original HDR image is impossible at present. However, compared with the images generated by other operators, it can be found that the local details in Fig. 2(d) are well-reserved and naturally distributed. No detail-loss can be found in the image.

5. CONCLUSIONS

A new realistic tone-mapping operator is proposed in the paper. Different to previous techniques, the proposed operator handle the details based on a contrast perception matching to original HDR images. The objective is to maintain the scene perception on the tone-mapped LDR image as close as possible to the original HDR image. To achieve the target, a new method to study full-range (from near-threshold to

6. REFERENCES

[1] E. Reinhard and K. Devlin, "Dynamic range reduction inspired by photoreceptor physiology," IEEE trans. on Visualization and Computer Graphics, vol. 11, no. 1, pp. 13–24, Jan.-Feb. 2005.
[2] R. Mantiuk, K. Myszkowski, and H.-P. Seidel, "A perceptual framework for contrast processing of high dynamic range images," ACM trans. on Applied Perception, vol. 3, no. 3, pp. 286–308, July 2006.
[3] R. Mantiuk, S. Daly, and L. Kerofsky, "Display adaptive tone mapping," ACM trans. on Graphics, vol. 27, no. 3, pp. 68, July 2008.
[4] T.-L. Ji, M. K. Sundareshan, and H. Roehrig, "Adaptive image contrast enhancement based on human visual properties," IEEE trans. on Medical
[5] V. Volevich, K. Myszkowski, A. Khodulev, and E. A. Kopylov, "Using the visual differences predictor to improve performance of progressive global illumination computation," ACM trans. on Graphics, vol. 19, no. 1, pp. 122–161, April 2000.
[6] H. R. Wilson, "A transducer function for threshold and suprathreshold human vision," Biological Cybernetics, vol. 38, no. 3, pp. 171–178, Oct. 1980.
[7] E. Peli, L. Arend, and A. T. Labianca, "Contrast perception across changes in luminance and spatial frequency," J. Opt. Soc. Am. A, vol. 13, no. 10, pp. 1953–1959, 1996.
[8] H. Pie´ron, The Sensations: Their Functions, Processes and Mechanisms, Yale University Press (Translated by Pirenne, M. H., & Abbott, B. C.), 1952. ging, vol. 13, no. 4, pp. 573 – 586, Dec. 1994.