OPTIMIZED DISPLAY ALGORITHM FOR NIGHT TIME IMAGES

J. Grave
LCPC, Paris, France
LIRIS, Lyon, France
email: justine.grave@lcpc.fr

R. Bremond
LCPC, Paris, France
email: roland.bremond@lcpc.fr

ABSTRACT
A fine computation model may lead to poor image display, due to the limitations of the display device. This well known problem has a specific intensity for night time scenes, first because of the high luminance dynamic between dark areas and light sources, secondly because of the image quantification of the dark levels. We propose a specific TMO devoted to this kind of images, and compare it to usual TMO of the literature, both in terms of visual appearance and performance in the visual scene.

KEY WORDS
Tone mapping, luminance, visualization, image quality.

1 Introduction

The visual quality of images depends on the quality of the underlying computational models, but also on the quality of the display. The visualization of night time images, either from a synthetic model or from a video is a challenge for two reasons: first the high dynamic luminance range (HDR), from the black sky to the artificial light sources, cannot be rendered with a display device. Second, item discrimination among low luminance levels faces a specific bias in most usual display devices: as luminance images are rendered with 8 bits (that is 256 grey levels), quantification effects may lead to objects disappearing, or non significant contrasts appearing. Tone Mapping Operators (TMO) are designed in order to manage this problem, and the night time situation is among the harder they have to face. Another points concerns the adaptation luminance, the luminous level where the visual system adapts itself, which depends on the display conditions. Visual performances strongly depend on this adaptation level. In this paper, we only consider the image visualization in controlled conditions, that is, in the case of night time images, in a dark room where no daylight or artificial light can add a veiling luminance to the displayed image.

Two main streams have developed TMO in order to cope with image display limitations: a first set of operators use a number of algorithmic tricks in order to give a better subjective impression of the displayed images (e.g. [1]); another growing approach tries to model at least some aspects of the optical and psycho-visual properties of image vision and image display in order to designed the operators (e.g. [2, 3, 4]). [5] propose a State of the art in the field of TMO, where most display issues are addressed; however, in this paper, some of them (as color and real time) are not considered, because we preferred to focus on the two main issues with night time images: dynamic compression and quantization at low luminance levels. This paper follows the main stream of TMO developments, trying to take into account vision science data in order to include a vision model and a display device model in our TMO. In the field of vision science, among others, the International Commission on Illumination publishes recommendations and state of the art reports, for instance on visual performances, display devices calibration [6] and adaptation luminance [7]. The idea of improving the image quality soon led to the idea that a subjective evaluation of the displayed images cannot be enough, and that evaluation experiments were necessary, following the psycho-visual standard of experiments in visual sciences. A number of experimental paradigms have been proposed, mostly on visual appearance criteria [8], but also on visual performance criteria [9].

In this paper, we propose a TMO designed in order to improve the display of night time images. This operator is limited to luminance images, because we deal mainly with mesopic light levels. It is also limited to low photopic, high and medium mesopic light level, which means that we do not consider luminance levels under $0.1\ cd.m^{-2}$ which are very difficult to render. Our model is based on both vision models (including a visual adaptation model) and a display model (taking into account the photometric characteristics of the display device, especially the quantification effects). We have compared our TMO to standard TMO from the computer graphics literature [1, 2, 3, 4] with both appearance and performance indexes for two kind of night time images: with and without light sources in the field of view.

2 Night time images and vision

The Human Vision System (HVS) has 3 different behaviors, depending on the luminous level: photopic under daylight conditions, when the rod photoreceptors of the retina are saturated; scotopic, under dark night, when the cones photoreceptors cannot perceive any light (they are not sensitive enough, and no color vision available); and mesopic, were both cones and rod are active. Usual night time scene (urban night scene, road night with the lights on) typically belong to the mesopic conditions. In this paper, we focus on visualization of mesopic images. However, we follow the results of the MOVE project [10] in considering that there is no need to use a specifically mesopic luminance definition above 0.1 $cd.m^{-2}$.

A number of rendering models, either from ray tracing techniques or from global illumination techniques, allow now to compute luminance images. On the other hand, video-photometry today allows to capture real images with photometric accuracy, in a metrological sense. Both techniques lead to a increasing need of accurate image display of luminance images. Unfortunately, even if their quality is increasing, no display device allow the dynamic range, the color range and sensibility of natural or computed images; what is more, the best display device technology is not available everywhere (for instance, on driving simulators), and algorithmic strategies are developed in order to keep the image quality despite of the display biases. A number of display device techniques are nowadays available : CRT, LCD, DLP, plasma, etc. The common framework to take these systems into account in the image display procedure is to perform a photometric (and if necessary colorimetric) characterization of the display device [6], in order to include these photometric data in the TMO display model.

The typical luminance range of a CRT display (when used in a dark room) is [1:100] $cd.m^{-2}$, which is very far from what would be necessary for an urban night scene rendering: car light may be up to 20 000 $cd.m^{-2}$, dark areas under 0.1 $cd.m^{-2}$, road surface between 0.1 and a few $cd.m^{-2}$. Another limitation is the quantization effect in the dark areas: with 256 grey levels, if one wish to render light sources, very few grey levels are available to render the low level luminances (between 0.1 and 2 $cd.m^{-2}$). This results in artificial contrasts which should not appear.

Light sources lead to another set of display issues. Actual glare is almost impossible with current display technology, while glare effects should be rendered if one wish to display images with the same visual appearance as the real scene. Spencer et al. [11] address this problem with a specific algorithm, based on vision science data, which computes a veiling luminance depending on the angle between the light source and the gaze axis, and adding it to the luminance image. The veiling luminance lowers the contrast discrimination capacities for people looking at the displayed image, as true glare would do. As light source are a major issue in night time images, we use Spencer's technique as a pre-processing step in our algorithm.

3 A night time TMO

We want to design a TMO which aims at improving the visualization of night time images. As we explained in our introduction, the images are luminance maps. The technical limitations of display devices (luminance dynamic and quantification) highly modify luminance and contrasts in images. Our objective is to build an operator which modifies the luminance and the contrasts of the image according to the device characteristics, preserving the contrast perception. The HVS is sensitive to a large range of luminance managed by visual adaptation. But perception depends on the adaptation levels; in particular, contrast perception. Our HVS is sensitive differently to different range of spatial frequencies. Our operator imitates the HVS on those aspects by using vision models function of adaptation luminance and spatial frequency.

3.1 Theoretical background of the operator

In this section, we develop the vision model we use to build our operator.

To mimic the sensitivity of the HVS to the spatial frequency, we decompose the image into different frequency bands. Burt and Adelson [12] proposed a fast pyramidal decomposition of the image into 7 frequency bands. They first build a Gaussian pyramid in which each image is the Gaussian filtered and sub sampled of the image at previous level. Then they compute the laplacian pyramid by subtracting the up sampled image of next upper level to the current level. The laplacian pyramid has 7 levels. The first 6 are images containing luminance contrasts. In the rest of this article we will call them contrast images. This is how we get the contrast values of objects with different size. The last level contains the low frequencies of the original image.

To model the contrast perception, we use the visibility level. It is the ratio between the luminance contrast between the object and the background, and the smaller increment of luminance detectable from the background luminance [13]:

$$VL = \frac{\Delta L}{\Delta L_t} \tag{1}$$

In our case, the ΔL is the value of luminance contrast given by the pyramidal decomposition. The contrast threshold ΔL_t is given by the Threshold Versus Intensity (TVI), a vision model that gives a contrast threshold depending on the adaptation luminance. The TVI is available for rods and for cones [14]. We use an equation system proposed by Larson et al. [3] that merges the two functions called Just Noticeable Differences (JND). The JND depend on the adaptation luminance L_a which is different for each level of the pyramid.

Few definitions exist for the adaptation luminance. The experiments carried out by Ishida [7] show that it can be computed by the mean of the luminance if the variance is small. We use this definition in our operator. As the adaptation is mainly achieved at the fovea, which size is around one degree, we will compute the adaptation luminance in a 1^o area.

The two technical limitations of the visualization devices that highly damage night time image rendering are luminance dynamic and quantification. To limit the damage, we want to anticipate the loss of information due to luminance compression and quantification. To reckon on the contrast and luminance modifications, we need to know some of the device characteristics such as the minimum and maximum of displayable luminance and the Look-Up-Table (LUT) between the addressing value and the displayed luminance. The LUT is measured with a photometer, following the method proposed in report 122 of the CIE [6]. We focus on contrast perception rendering. The VL, that quantifies the contrast perception, is used to make sure that the original contrast and the displayed contrast are perceived the same way. If we want the VL to remains constant while changing the contrast, the contrast threshold have to change as well. The contrast threshold is given by the JND, computed for an adaptation luminance. The adaptation luminance is processed, to fit the luminance dynamic of the display device in order to maintain a constant VL, despite the quantification.

3.2 Computational model

Spatial frequency decomposition The original image I is decomposed into a laplacian pyramid according to the method of Burt and Adelson [12]. We first build a Gaussian pyramid using a 5×5 gaussian filter w. Each level of the pyramid represents a low-pass image, cut at a frequency half the one of the next higher level. The Gaussian pyramid has 7 levels to cover the sensitivity domain of the HVS. The image at level l, $l \in \{1, 2, 3, 4, 5, 6, 7\}$, of the pyramid, denoted as L^l_w, w for world, is computed from level $l-1$:

$$L^l_w(i,j) = \sum_{m=-2}^{2} \sum_{n=-2}^{2} w(m,n) L^{l-1}_w(2i+m, 2j+n) \quad (2)$$

$$L^1_w(i,j) = I(i,j) \quad (3)$$

The filtering process downsamples the image. To compute a pyramid of difference-of-Gaussian images, we take the image at level l, L^l_w, and we subtract the expanded image at level $l+1$ which has been upsampled, denoted L^{l+1}_{wExp}:

$$L^{l+1}_{wExp}(i,j) = \sum_{m=-2}^{2} \sum_{n=-2}^{2} g(m,n) L^{l+1}_w\left(\frac{i}{2}+m, \frac{j}{2}+n\right) \quad (4)$$

$$C^l_w(i,j) = L^l_w(i,j) - L^{l+1}_{wExp}(i,j) \quad (5)$$

This results in a 7 levels pyramid, the first 6 levels are band-pass images and the last higher level is a low-pass image. We denote C^l_w the contrast images.

Contrast perception modelling The contrast perception is modelled by the Visibility Level (VL), defined by:

$$V^l(i,j) = \frac{C^l_w(i,j)}{\Delta L_t(L^l_{aw}(i,j))} = \frac{C^l_w(i,j)}{JND(L^l_{aw}(i,j))} \quad (6)$$

where V^l is the image of VL values at level l, ΔL_t is the perception threshold of luminance contrast which depends on an adaptation luminance denoted L^l_{aw} and is computed using the JND.

The adaptation luminance values L^l_{aw} are reckoned from the L^{l+1}_{wExp} images. We compute the mean of luminance over 1 degree in the visual field.

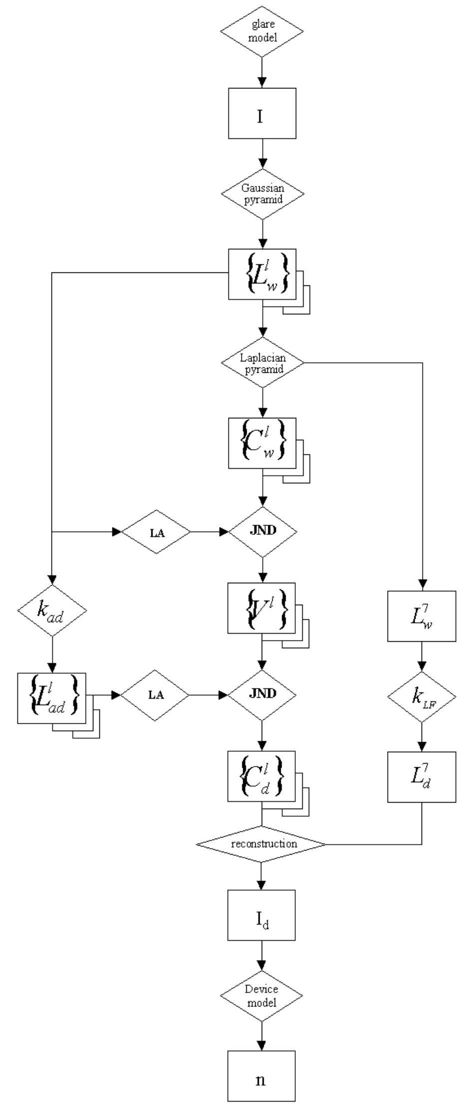

Figure 1. Framework of our algoritm

The VL images are used to compute new contrast images C^l_d (d for display) adapted to the display device characteristics on which the rebuilt image shall be displayed. The new contrast images are computed using equation 6 and the JND, to inverse the model. The new adapted contrast is computed in order to maintain a VL constant:

$$C^l_d(i,j) = V^l(i,j).\Delta L_t(L^l_{ad}(i,j)) \quad (7)$$

Adaptation luminance processing To adapt the contrast to the display device, we need then to modify the adaptation luminance taking the device characteristics into account. New adaptation images L^l_{ad} are computed from the old ones L^l_{aw}, using a scale factor k_{ad} (one for each spatial frequency band, in order to fit the HVS behaviour).

$$L^l_{ad}(i,j) = k_{ad}.L^l_{aw}(i,j) \quad (8)$$

Ward [2] proposes a scale factor that can be used for each level of the Laplacian pyramid to process the adaptation images. But this scale factor expands the luminance dynamic when the average luminance is small and night time images appear like daytime images.We could adapt Ward's scale factor by maximizing the scale factor by one. However, this simple solution does not hold for night time images. First because the minimum of displayed luminance L_{dmin} of most display devices is higher than the lower luminance to be displayed L_{wmin}. Secondly, because of the quantification, many contrasts in those scenes are destroyed. So the scale factor should be higher than one but in a reasonable way. We propose in this paper to choose the smallest value of the scale factor allowing that the smallest perceptible contrast ΔL_t would be perceived. Let's note A_w an adaptation level in the original image and A_d the one in the displayable image.

$$A_d = k_{ad} * A_w \tag{9}$$

$$A_d^{'} = A_d + \Delta L_s^d = k_{ad} * A_w + JND(k_{ad} * A_w) \tag{10}$$

We call $LtoAV$ is the function, characteritic of the display device, that gives an 8-bit numeric value from a luminance value and $AVtoL$ the function that gives a luminance value from an 8-bit numeric value ($LtoVA$ is a quantification function so $Z = LtoAV \circ AVtoL \neq Id$).

$$A_{LCD} = AVtoL(LtoAV(A_d)) = Z(A_d) \tag{11}$$

$Z(A_d)$ gets the actual adaptation level which is displayed when we want to display A_d.

$$A_{LCD}^{'} = A_{LCD} + \Delta L_s^{LCD} \tag{12}$$

$$\Delta L_s^{LCD} = A_{LCD}^{'} - A_{LCD} = Z(A_d^{'}) - Z(A_d) \tag{13}$$

Since we want $\Delta L_s^{LCD} = \Delta L_s^d$:

$$Z(A_w + k_{ad} * A_w) - Z(k_{ad} * A_w) = JND(k_{ad} * A_w) \tag{14}$$

Let's consider:

$$\alpha(k_{ad}, A_w) = \frac{Z(A_w + k_{ad} * A_w) - Z(k_{ad} * A_w)}{JND(k_{ad} * A_w)} \tag{15}$$

For a given value A_w, we choose the smallest value of k_{ad} such that $\alpha(k_{ad}, A_w) = 1$. If $\alpha(k_{ad}, A_w)$ remains lower than 1, we choose k_{ad} such that $\alpha(k_{ad}, A_w) = \alpha_{max}$. To make sure that k_{ad} is optimal for every adaptation level in the image, we chose $A_w = L_{wmin}$. Finally, we clip the adaptation luminance by the two extreme displayable luminances, L_{dmin} and L_{dmax}.

Image reconstruction The image is reconstructed, following the inverse process of the pyramidal decomposition. The last level of the pyramid is first scaled with the factor k_{LF} [2], clipped to 1. It is then upsampled and added to the next lower level. The result is upsampled and added to the next lower level and so on until the image is rebuilt. We finally apply the LUT of the display device to convert luminance values into addressing values.

Figures 3 and 2 show images processed by different TMOs. We cannot judge or compare the TMOs based on those images because they have been calculated to be displayed on a specific LCD display device, under specific visual conditions, and because we cannot display the HDR image. In the next section, we detail the experiments we carried out to evaluate the image quality rendered by our operator.

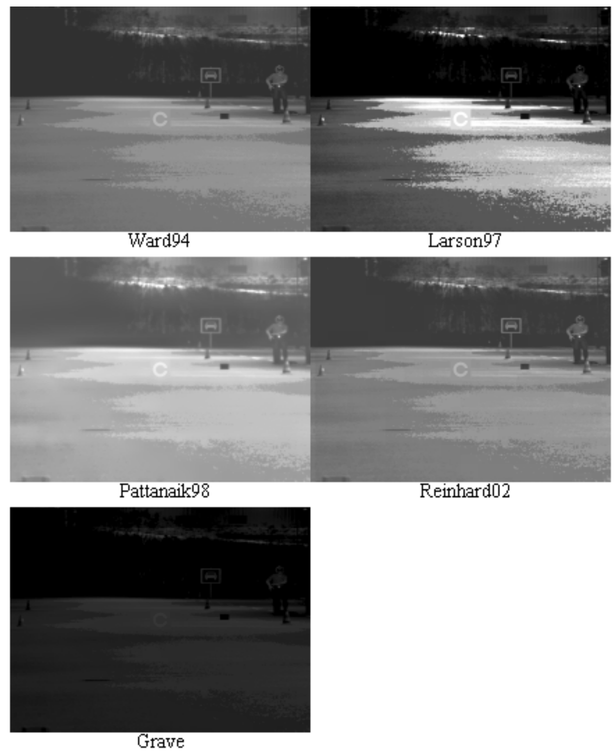

Figure 2. Image LN150W, L_{wmin} = 0.25 $cd.m^{-2}$ and L_{wmax} = 5.1 $cd.m^{-2}$, processed by 5 different TMOs: W94, L97, P98, R02 and the one we present in this paper

4 Quality evaluation

Two main indexes are used in general purpose perception evaluation: appearance and performance cues. In the case of image quality evaluation, these two psycho-visual indexes difference lies in the subjective vs. objective record. An appearance test concerns the subjective evaluation of the images quality (or image comparisons) by a panel of observers, while a performance test concern the objective performance of a panel of observers in a visual task (object detection, object shape discrimination, reading, etc.). As both criteria are important, we test our TMO with both kind of psycho-visual experiments. A performance and an appearance tests are carried out with two night time images: a nighttime urban driving scene without light source in the

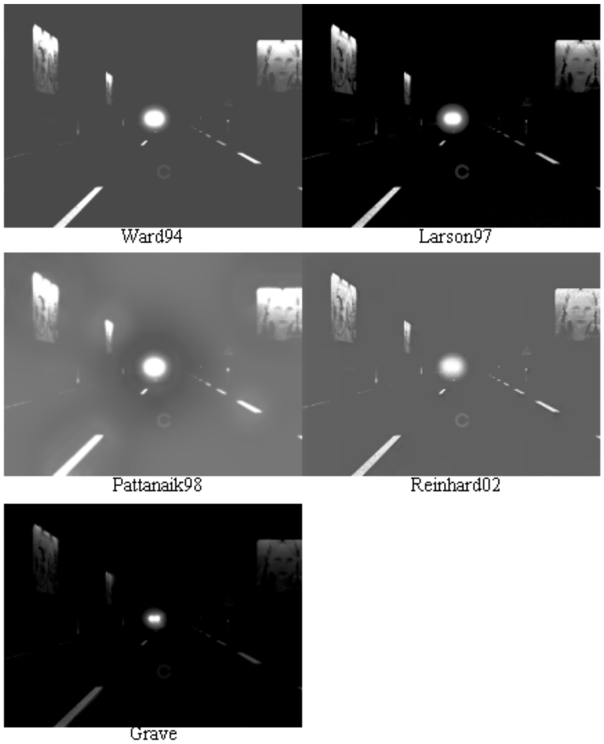

Figure 3. Image NIGHTDRIVE, L_{wmin} = 0.5 $cd.m^{-2}$ and L_{wmax} = 478.4 $cd.m^{-2}$, processed by 5 different TMOs: W94, L97, P98, R02 and the one we present in this paper

visual field and a nighttime country driving scene with light source in the visual field. 10 subjects take the tests for each scene. For the NIGHTDRIVE image, there are 2 women and 8 men. 7 subjects were between 25 and 35 years old and 3 were between 35 and 55. For the LN150W image, there are 1 woman and 9 men. 3 subjects were between 25 and 35, 2 were between 35 and 55 and 5 were over 55.

Performance rendering To evaluate the quality rendering of our operator in terms of performance, we record the visual performance of observers with a reference scene and with a comparison scene. The reference scene is a High Dynamic luminance Range (HDR) image displayed with a DLP (Digital Light Processing videoprojector). The comparison scene is the image processed by a TMO to fit the low dynamic range of the display device, a LCD (Liquid Crystal Device). The closer the visual performance with the comparison scene are to the one with the reference scene, the better is the TMO. The visual performance is measured with a Landolt ring (see figure 4) in a uniform square. The whole is inserted in a night time image (see figures 2 and 3). The ring is full during 1 s, then a gap appears in one of the four possible positions (left, right, top, bottom), during 100 ms. The subjects indicate, with a gamepad, the position of the gap. The test is carried out 200 times spread over eight different contrasts. The luminance of the square, the size and the position of the ring remain constant. Figure 5 shows the visual performance for both

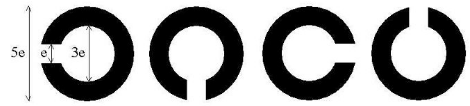

Figure 4. The Landolt ring for the performance test

scenes, measured with the reference scene, and with the comparison scene computed using different TMO: W94, L97, P98, R02 and our operator. For the image LN150W, 2 operators provide good results, P98 and our operator. W94, L97 and R02 over estimate the contrast perception and improve the visual performance of observers. For image NIGHTDRIVE, the visual performance with our operator are closer to the visual performance with the reference than with the other TMO, yet the contrast perception is under estimated.

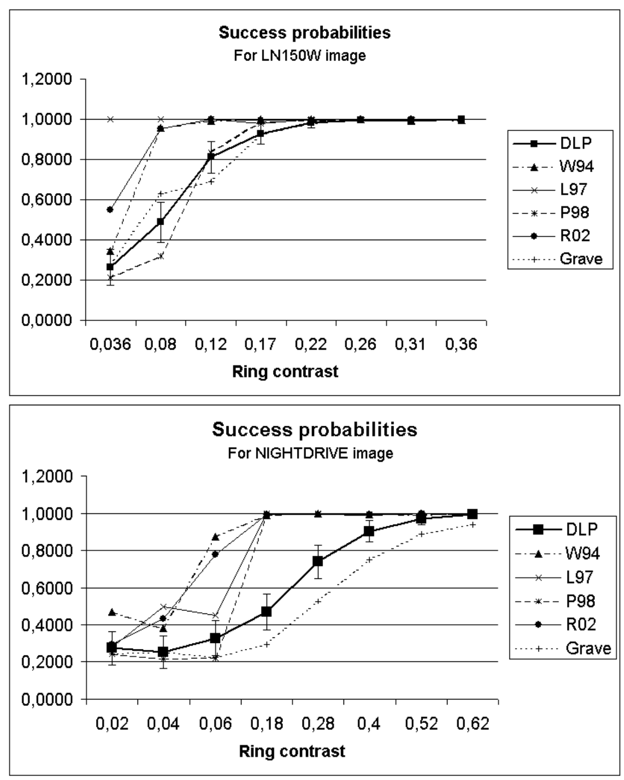

Figure 5. Visual performance measured with the image processed by 5 TMO, including our, for images LN150W (top) and NIGHTDRIVE (bottom).

Appearance rendering Our operator is compared to the same other TMOs, with an experimental protocol close to Ledda et al. [8]. Two images are displayed at the same time, with the same angular size: the reference image, displayed by the DLP, and the comparison image, displayed by the LCD. The reference image remains the same during the whole test. On the LCD device, two images processed by two different TMO are displayed successively. The subjects are asked to choose between the two processed images the one that seems, according to them, the "closest" to the reference scene. The observers may go back and forth from one image to another as many time as needed to make their choice. The 5 TMO are compared, which results in 10 different image couples. Each couple is displayed 3 times, which means that each TMO appears in 12 couples. Table 1 shows the average number of time an operator was chosen, over the 10 subjects. The results show that our operator is the first chosen in the night time with direct source light image, and is the first, equal with Ward's operator, in the night time without direct light source image.

	W94	L97	P98	R02	G	total
nightdrive	5.9	8.9	0.8	2.9	11.5	30.0
LN150W	9.4	0.8	3.1	6.8	9.9	30.0

Table 1. Appearance tests results: Ward94, Larson97, Pattanaik98, Reinhart02, Grave.

5 Conclusion

We propose a new tone mapping algorithm designed for night time image visualization. Our main objective is to preserve the observer's contrast perception at low luminance levels despite the luminance dynamic range compression and the quantification of images. To evaluate our algorithm, we carried out two psychophysical experiments in order to evaluate the quality of the processed images with appearance and performance tests. We performed the tests with 2 night time images: a night time without direct light source and a night time with direct light sources scenes. Over the 2 scenes, our operator gives better results than the 4 other TMO used for comparison. This means that our TMO fulfills our objectives. It allows to display images preserving the contrast perception and the general brightness. Our operator only process luminance map. Even if colour is not a critical parameter for the mesopic domain, the ecological validity of our operator would need to expand our operator to colour images. Our operator deals with still images. We could develop our algorithm to video. Then with a real time implementation and a reconfiguration of some aspects, we could integrate our algorithm into the visual loop of a driving simulator and simulate night time driving.

References

[1] E. Reinhard, M. Stark, P. Shirley and J. Ferwerda, Photographic tone reproduction for digital images, *In Proceedings of ACM SIGGRAPH*, 2002.

[2] G. Ward, A Contrast-Based Scalefactor for Luminance Display, *Graphics Gems IV* (Boston, Academic Press, 1994), 415-421.

[3] G. W. Larson, H. Rushmeier and C. Piatko, A Visibility Matching Tone Reproduction Operator for High Dynamic Range Scenes, *EEE Transactions on Visualization and Compter Graphics*, 1997.

[4] S. Pattanaik, J. Ferwerda, M. Fairchild and D. Greenberg, Multiscale Model of Adaptation and Spatial Vision for Realistic Image Display, *Proc. SIGGRAPH*, 1998.

[5] K. Devlin, A. Chalmers, A. Wilkie and W. Purgathofer, Tone Reproduction and Physically Based Spectral Rendering, *State of the Art Reports, Eurographics*, 2002.

[6] CIE 122, The relationship between digital and colorimetric data for computer-controlled CRT displays, *Commission Internationale de l'Eclairage*, 1996.

[7] T. Ishida and K. Iriyama, Estimating Light Adaptation Levels for Visual Environments with Complicated Luminance Distribution, *CIE Proceedings the 25th session*, 2003.

[8] P. Ledda, A. Chalmers, T. Troscianko and H. Seetzen, Evaluation of Tone Mapping Operators using a High Dynamic Range Display, *Proc. SIGGRAPH*, 2005.

[9] F. Vienot, C. Boust, E. Da Costa, R. Bremond and E. Dumont, Psychometric assessment of the look and feel of digital images, *Driving Simulator Conference*, 2002.

[10] *http://lightinglab.fi/CIETC1-58*

[11] G. Spencer, P. Shirley, K. Zimmerman and D. Greenberg, Physically-based Glare Effects for Computer Generated Images, *Proceedings ACM SIGGRAPH*, 1995.

[12] P. J. Burt and E. H. Adelson, The Laplacian Pyramid as a Compact Image Code, *IEEE Transactions on Communications*, 1983.

[13] W. Adrian, Visibility of targets : model for calculation, *Lighting Research and Technology*, 1989.

[14] J. A. Ferwerda, S. N. Pattanaik, P. Shirley and D. P. Greenberg, A Model of Visual Adaptation for Realistic Image Synthesis, *Proceedings of ACM SIGGRAPH*, 1996.