A Tone-Mapping Operator for Road Visibility Experiments

JUSTINE GRAVE
Laboratoire Central des Ponts et Chaussées, and
Laboratoire d'InfoRmatique en Image et Systèmes d'information
and
ROLAND BREMOND
Laboratoire Central des Ponts et Chaussées

One may wish to use computer graphic images to carry out road visibility studies. Unfortunately, most display devices still have a limited luminance dynamic range, especially in driving simulators. In this paper, we propose a tone-mapping operator (TMO) to compress the luminance dynamic range while preserving the driver's performance for a visual task relevant for a driving situation. We address three display issues of some consequences for road image display: luminance dynamics, image quantization, and high minimum displayable luminance. Our TMO characterizes the effects of local adaptation with a bandpass decomposition of the image using a Laplacian pyramid, and processes the levels separately in order to mimic the human visual system. The contrast perception model uses the visibility level, a usual index in road visibility engineering applications. To assess our algorithm, a psychophysical experiment devoted to a target detection task was designed. Using a Landolt ring, the visual performances of 30 observers were measured: they stared first at a high-dynamic range image and then at the same image processed by a TMO and displayed on a low-dynamic range monitor, for comparison. The evaluation was completed with a visual appearance evaluation. Our operator gives good performances for three typical road situations (one in daylight and two at night), after comparison with four standard TMOs from the literature. The psychovisual assessment of our TMO is limited to these driving situations.

Categories and Subject Descriptors: I.3.3 [**Computer Graphics**]: Picture/Image generation - Display algorithms; I.3.8 [**Computer Graphics**]: Applications

General Terms: Experimentation, Human Factors, Measurement

Additional Key Words and Phrases: HDR images, road visibility, visual performance, psychophysics

ACM Reference Format:

Grave, J. and Bremond, R. 2008. A tone-mapping operator for road visibility experiments. ACM Trans. Appl. Percpt. 5, 2, Article 12 (May 2008), 24 pages. DOI = 10.1145/1279920.1361704 http://doi.acm.org/10.1145/1279920.1361704

1. INTRODUCTION

Using computer graphic images for road visibility studies implies that the images are realistic enough to lead to the same behavior as in a driving situation, which is called an "ecological" situation [Hoc 2001]. Thus, driving simulations and psychovisual experiments using computer graphic images need to

Authors' address: Justine Grave, Sagem Securité, Paris; email: justine.grave@sagem.com. Roland Brémond, Laboratoire Central des Ponts et Chaussées, Lyon; email: roland.bremond@lcpc.fr

be validated with regard to a number of visual tasks. Among them, the ability to detect a target in the road scene is a key task in driving [Hills 1980], which is known to be poorly achieved in usual driving simulators [Viénot et al. 2002]. It seems, then, that road visibility studies could benefit from the use of computer graphic images, through driving simulation and psychovisual experiments.

Unfortunately, the visual environment of the driver is far more complex than any display device is able to render in terms of luminance dynamic range, minimum and maximum luminance values, quantization, color gamut, color values and spatial resolution, etc. In a daytime driving scene, the luminance can be as high as 10^6 cd.m^{-2}. In a nighttime driving scene, the luminance can be as low as 10^{-2} and as high as 10^5 cd.m^{-2} at the same time, because of the headlights of on-coming traffic. A standard LCD monitor is unable to display luminances below 0.5 and beyond 200 cd.m^{-2}. The video projectors usually used in driving simulators are even worse in displayable luminance dynamic range: because of the size of the projection area, the maximum luminance can be as low as 10 cd.m^{-2}. Thus, the main limitation of display devices for road vision applications is the need to compress the luminance dynamic range of images to fit the display device characteristics. This is what a TMO actually does: to map real world luminances to target display luminances in order to reproduce the overall impression of brightness and contrast of the real world, despite the limited range of the display device.

In this paper, we propose a methodology that aims to validate the use of computer graphic images for road visibility applications, using a detection task to build a quality index. The ecological situation of interest is a target detection while driving, and the experimental protocol takes into account the main variables of this task. Hence, we can compare several tone-mapping operators (TMO). Thus, we propose a specific TMO designed in order to fulfill the requirements of road vision experiments and we assess our TMO in terms of a target detection task.

If one wants to use real-world luminance images to carry out road visibility studies, an appropriate TMO should preserve the main visual cues of the road scene. TMOs found in the computer graphic literature mainly address visual appearance issues rather than visual performances: they try to reproduce the subjective impression of observers. In this paper we focus on visual performance in a detection task (i.e., the detection threshold), which is the main issue in road visibility applications. The difference between appearance and performance indexes is that appearance measurements record subjective feeling on a topic where there is no objective answer (e.g., *in your opinion, which of these images is nicer*?), whereas performance measurement refer to objective situations, so that the answer may be right or wrong (e.g., *what letter do you read here?*).

Considering the state of the art in computer graphics, we designed a TMO based on vision science literature, which tries to maintain the observer's visual performances through keeping the visibility level of objects (see Section 3.2.1). A psychovisual experiment allows to test our algorithm in three situations with known display limitations for road scenes: two night scenes (with and without oncoming cars with headlights on) and a daylight scene in fog (headlights on).

After a short overview of road visibility key issues, TMO design and evaluation protocols, we develop our TMO in terms of underlying vision models and in terms of algorithmic implementation. We then describe the specific psychophysical experiments we designed to assess a target detection task. This protocol is used to compare our TMO to four well known operators. Finally the results are discussed and future work is proposed.

2. RELATED WORK

2.1 Road Visibility

Target detection is a basic ability, which is critical in a number of driving tasks, as visual cues are the main source of information for the driver [Hills 1980]. Thus, an experimental way to assess the ecological

validity of a detection task on displayed images is of great importance in the field of perception studies in driving.

Visual performances depend on the task one considers (driving and reading, etc.) as well as on photometric and colorimetric data (luminance, acuity, color, and visual adaptation, etc.). [CIE 1981] considers that "free" visual detection (with no secondary task) is a key parameter in constrained visual detection (as in a driving or reading task). This implies that a quantitative task-dependent relation can be found between visual performances in a detection task and in a specific task (in the following, target detection during the driving task). This result, even modulated by [CIE 1992], allows to define a reference situation limited to a detection task in a laboratory, even for road vision applications.

Considering the human visual system (HVS), the main parameters, in terms of target detection, are the luminance contrast between object and background, the adaptation luminance, and the angular size of the object [Adrian 1989]. At first order of approximation, color and object motion can be disregarded. This hierarchy among object parameters leads to an experimental protocol devoted to static observation of greyscale images (see Section 4.1).

This hierarchy can be translated in terms of image display devices limitations, in order to choose which parameters have to be taken into account first. Considering the detection task, color display and spatial resolution are second-order issues, while the critical parameters are luminance dynamic range and quantization. Image quantization problems appear mostly in nighttime road images, where contrasts in dark areas may be badly rendered when light sources are in the field of view. Another point is the high luminance level (1 cd/m^2 and more), which can be found as minimum displayable luminance value, especially on driving simulator display devices, to be compared to real world values far under 0.1 cd/m^2 for a road at night.

2.2 Tone Reproduction Operators

The problem of tone reproduction is not a recent one. It was first tackled by photographers who needed to process their pictures so that they fit the visual appearance of the photographed scenes. For computer graphic images, many TMOs have been proposed to compress the dynamic range of an image so that it can be effectively displayed [Reinhard et al. 2005]. There are two main categories of such operators: spatially uniform (also known as global) and spatially varying (also known as local). Global operators apply the same transformation to every pixel of the image regardless of their position in the image. Local operators apply different transformations to different parts of the image, depending on their local properties. As this is not the place for an exhaustive state of the art (see [Reinhard et al. 2005]), we focus on four operators, which may illustrate some among the main tendancies of TMO design, with respect to local versus global operators and vision science versus heuristic based strategies. These operators are used in section 4 for comparison.

Among global TMOs, Ward's operator [Ward 1994] is one of the first to have a psychovisual theoretical background. It is a fast linear operator which computes a scale factor using the visual performance model of Blackwell [CIE 1981]. It aims at preserving the threshold of perceived contrasts while compressing the luminance range, which is relevant for road vision applications. However, for low luminance distributions, the scale factor expands the luminance dynamic so that nighttime images may appear as daytime images. Larson et al.'s operator [Larson et al. 1997] uses a visual adaptation map. They propose an histogram adjustment technique, including a visual detection threshold model in order to constraint the histogram slope.

Pattanaik et al. [1998] proposed a local operator. Their TMO is divided in two parts: a vision model and then a display model. It is based on a multiscale representation of patterns, on luminance and color processing, and addresses the problem of perception of scenes at threshold and suprathreshold levels. As most multiscale approaches, this method unfortunately introduces artefacts known as "halos"

(contrast inversion). Even if this artifact has a counterpart in the HVS [Peli 1990], it is not wanted in general-purpose image display. However, as the multiscale image decomposition is still considered as a relevant way to tackle the tone-mapping problem, solutions were proposed to minimize those halo artefacts [Li et al. 2005]. Also local, but with a completely different approach, is the TMO proposed by Reinhard et al. [Reinhard et al. 2002]. Instead of relying on vision science models, they proposed a photographic tone reproduction operator, based on a model of ad hoc know-how among photographers practice, known as the zone system technique.

A further aspect to tone reproduction is time. Some operators include temporal adaptation (to light and to darkness) [Ferwerda et al. 1996; Pattanaik et al. 2000; Ledda et al. 2004]. We do not consider temporal issues in our TMO, because paradoxically most vision models in road engineering (road lighting, road sign design, etc.) are static ones [CIE 1988].

2.3 TMO Evaluation

There are two ways to assess the quality of a TMO, both being explored in the computer graphic literature. One can either rely on a perception-based image metric (e.g., Smith et al. [2006]), or use subjective experiments (e.g., Ledda et al. [2005]), showing visual stimuli to human observers, and concluding about the TMO quality from their answers to a specific question. The first approach is a classical computational approach, whereas the second approach enters the field of experimental psychology.

Following the psychological approach, Some subjective experiments have been proposed in the recent years to evaluate TMOs. Most of them are based on visual appearance indexes. Drago et al. [2003] ask their subjects to perceptually judge and to indicate their preference over a panel of images processed by different TMOs. They analyze these preference data to determine a preference point in the stimulus domain used as a reference to compare algorithms. In another study subjects were asked to rate tone-mapped images, compared by pairs [Kuang et al. 2004]. Ledda et al. [Ledda et al. 2005] conducted experiments using a high-dynamic range (HDR) image as a reference. They compared TMOs two by two and asked the observers to choose, between two different tone-mapped images, the one they felt closer to the reference image. On the other hand, Viénot et al. [Viénot et al. 2002] presented an evaluation paradigm using a physical reference scene instead of a HDR image and performed TMO comparisons with psychophysical experiments based on both visual performance (a detection task) and visual appearance (gratings comparison). Yoshida et al. [Yoshida et al. 2006] use a more sophsiticated kind of physical reference, using a HDR photograph of the physical scene as an input of the TMO.

3. TMO

3.1 Motivations and Design Framework

Our objective is to build an operator, which modifies the luminance and the contrast in images according to the characteristics of the display device and still preserves target detection performances. We are not looking for a new TMO concept, but for concepts relevant for road vision. For this reason, the resulting TMO may somehow look like a wall of old stones put together in a new arrangement. We feel comfortable with this, as long as this new arrangement allows to extend the validity domain of computer graphic images.

Driving environments generate complex images in terms of luminance distribution (for instance, a nighttime road scene may include very dark areas with important semantic informations and, at the same time, glaring light sources). It is well known that target detection mostly relies on luminance contrast [CIE 1981], so that the images we are dealing with, in the following, are luminance maps. The MOVE project [MOVE 2005] showed that it is possible to rely on the photopic definition of luminance

down to 0.1 cd/m^2, which we considered low enough to cope with most driving environments (see [Ferwerda 2001] for an overview of early vision).

Within all the display device limitations, we chosed to focus on two major constraints that can have strong effects on a detection task in a driving context, luminance dynamic range, and quantization. The compression of the luminance dynamic from HDR to LDR images modifies contrast and adaptation luminance, which are key parameters in detection models. The quantization problem is because of the fact that most display devices (including those in driving simulators) are limited to 256 luminance values for each channel. For a grey-level image, the truncation of floating-point luminance values may lead to significant errors in terms of contrast values—at least this is what we observed with nighttime road images. To our knowledge, this issue was not previously addressed in the TMO literature. The quantization drawback is emphasized when one tries to display low luminances with a display device with a high minimum display value (e.g., around 1 cd/m^2, which is typical of CTR display devices).

As pointed out by [Peli 1990], if one wishes to keep some perceptual meaning in the contrast definition, a specific approach is necessary for complex images, taking into account local adaptation and the multiple bandpass behavior of the HVS. We follow Peli in its implications by designing a local TMO in order to deal with such a complexity. This starting point gives a design framework for our TMO. In road visibility literature, the contrast perception is defined using the VL index, which is detailed later (see Section 3.2.1). The HVS is sensitive to a large range of luminance values, thanks to visual adaptation [Shapley 1991]. This complex physiological process tunes the perception characteristics (pupil diameter and neural sensitivity, etc.) in order to give optimal performances around the current level of illumination. However, perception is not achieved identically for all adaptation levels. The HVS is differently sensitive depending on the range of spatial frequencies [Campbell and Robson 1968]. Our operator decomposes the image into consecutive spatial frequencies bands and separately processes them to mimic the HVS behavior.

3.2 Vision Models

In this section, we detail a vision model that meets our objectives, taking into account the three main parameters, which seem to be the most relevant for object detection: detection threshold and suprathreshold, visual adaptation, and spatial frequency decomposition. The quantization issue is addressed, then, as a specific problem.

3.2.1 *Visibility Level.* In road lighting literature [Adrian 1989], the visibility level (VL) is the usual index to rate the visibility of an object over its background. It is the ratio between the luminance difference $\Delta L = L - L_b$ between the object and its background (L_b is supposed to be uniform), and the smaller increment of luminance ΔL_t (t for threshold) detectable from the same background luminance:

$$V = \frac{\Delta L}{\Delta L_t(L_b)} \tag{1}$$

The VL is a quantitative index for contrast perception, which allows to describe the contrast detection threshold ($V = 1$) as well as suprathresholds ($V > 1$).

Several models in vision science allow to compute the contrast detection threshold ΔL_t. The contrast sensitivity function (CSF) gives ΔL_t as a function of the spatial frequency, but this model was developed for a visual task, which is not a detection task (the discrimination of sinusoidal bands in a grid) and is only available for a few adaptation luminance values. The threshold versus intensity function (TVI) gives ΔL_t as a function of the adaptation luminance. It explicitly addresses the discrimination of a target on a uniform background. Thus, the TVI model corresponds better to our objectives of modeling a target detection at various adaptation levels than the CSF.

Table I. JND Equations from [Larson et al. 1997], with $B = \log(L_a)$ for Easy Reading

Luminance domain	$\log(D(L_a))$
$B < -3.94$	-2.86
$-3.94 \leq B < -1.44$	$-2.86 + [1.6 + 0.405B]^{2.86}$
$-1.44 \leq B < -0.0184$	$B - 0.395$
$-0.0184 \leq B < 1.9$	$-0.72 + [0.65 + 0.249B]^{2.7}$
$1.9 \leq B$	$B - 1.255$

[a]The JND appears as D in the equations. TVI data in [Ferwerda 1996] comes from vision science literature.

Two TVI models are needed, for rods and for cones photoreceptors. To merge these two functions, we use the same just noticeable differences (JND) equation system as Larson et al. [Larson et al. 1997] for consistency with the TVI functions from Ferwerda et al. [Ferwerda et al. 1996] (see Table I).

The JND depend on the adaptation luminance L_a, which has to be computed locally. If L_a can be assumed to be L_b, the VL of a contrast ΔL with adaptation level L_a can be rewritten from Eq. (1) as:

$$V = \frac{\Delta L}{D(L_a)} \tag{2}$$

3.2.2 *Visual Adaptation.* Visual adaptation depends both on temporal (to darkness or to light) and on spatial adaptation. In the following, as we consider static visual performances, we only address the spatial adaptation issue.

Several definitions of the adaptation luminance can be found in the vision science literature (see [CIE 1999] for a short review). When a target is seen on a uniform background, the adaptation luminance can be set to the background luminance. However, usual scenes have rich and nonuniform luminance distribution. In the general nonuniform case, even if there is no scientific consensus [CIE 1999], the experiments carried out by Ishida [Ishida and Iriyama 2003] showed that it can be computed as the mean luminance if the variance is small. As we are interested in local adaptation, we assume that this definition is suitable to our operator. As the adaptation is mainly achieved in the fovea [Moon and Spencer 1945], the angular size of which is around 1°, we computed the adaptation luminance as the mean over a square of 1° width (this implies that our TMO needs to know the geometrical conditions of the display, which is easy in usual psychovisual experiments).

Psychophysical data suggest that the processing of retinal images may be described as including a set of bandpass filters in the frequency domain, which process the visual signal in parallel [Campbell and Robson 1968]. This corresponds to the different processing channels in the neural architecture [Ginsburg 1986]. The contrast detection may then be described as the result (envelope) of the combined activity of a number of bandpass channels. These results are of some consequencies on visual adaptation, as it means that adaptation not only happens in a spatial surround, but also should be considered on each channel. This idea lead to Peli's local contrast definition [Peli 1990], well suited to perception issues in image representation, because it takes local adaptation into account. Peli considers a bandpass filtered image $a_{i,j}$ and defines the contrast at this frequency level as:

$$C_{i,j} = \frac{a_{i,j}}{l_{i,j}} \tag{3}$$

where l is the low-pass filtered image containing all energy bellow the band. The link between this contrast definition and the VL comes from the fact that the bandpass image $a_{i,j}$ contains luminance differences, homogeneous to ΔL. This is made clear, considering a pyramidal decomposition of the

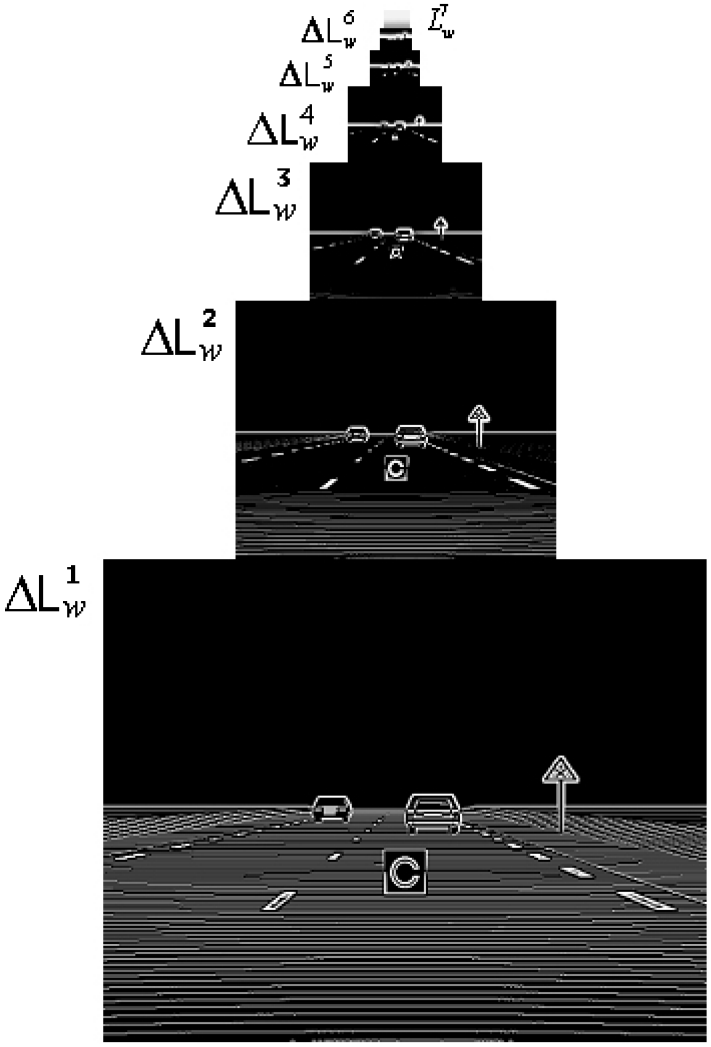

Fig. 1. Pyramidal decomposition of image DayFog: Contrast images ΔL_w^k (k =1–6) and low-frequency image L_w^7.

image, in which bandpass images are composed of differences of filtered images (with increasing filter size). On the other hand, the local adaptation luminance L_a may be estimated as the mean luminance around pixel (i, j) in the next bandpass level.

In short, the TVI data allows to compute luminance contrasts with the VL index in terms proposed by Peli [Peli 1990], taking into account the differential HVS behavior depending on the spatial frequency. The implementation of such a model leads to a spatial frequency decomposition of the images. Figure 1 shows such a decomposition, with six bandpass images and 1 low-pass image.

3.2.3 *Quantization.* We denote *LtoAV* (for *luminance to addressing value*) the function, characteristic of the display device, which gives an 8-bit numeric value from a floating-point luminance value. *AVtoL* is the reverse function that gives a luminance value from an 8-bit addressing value. *LtoAV* is a quantization function so that $Z = LtoAV \circ AVtoL \neq Id$. $Z(L_d)$ can be understood as the actual

Table II. Gaussian Filter g Used to Build the Laplacian Pyramid.

0.0025	0.0125	0.02	0.0125	0.0025
0.0125	0.0625	0.1	0.0625	0.0125
0.02	0.1	0.16	0.1	0.02
0.0125	0.0625	0.1	0.0625	0.0125
0.0025	0.0125	0.02	0.0125	0.0025

luminance, which is displayed on the monitor when it is asked to display L_d. The quantization issue is the fact that $Z \neq Id$. We address it by considering an adaptation luminance L_{ad} in the low-dynamic range (LDR) image and the smallest visible luminance increment ΔL_{ad}^{JV} (JV for just visible). We use the display device photometric properties at low luminance levels in order to be sure that ΔL_{ad}^{JV} shall actually be visible (see Section 3.3.3).

3.3 Computational Model

3.3.1 *A Pyramidal Decomposition.* Figure 2 shows the general framework of our operator. For the spatial decomposition of the HDR luminance image I, we use the Laplacian pyramid described by [Burt and Adelson 1983]. We first build a Gaussian pyramid using a 5×5 Gaussian filter g (see Table II).

Each level of the pyramid represents a low-pass image, cut at a frequency one-half the one of the next higher level. The Gaussian pyramid has seven levels to cover the sensitivity domain of the HVS (see Figure 1). The image at level l of the pyramid, denoted as L_w^l (w for world) is computed from level $l-1$:

$$L_w^l(i,j) = \sum_{m=-2}^{2}\sum_{n=-2}^{2} g(m,n) L_w^{l-1}(2i+m, 2j+n) \tag{4}$$

$$L_w^1(i,j) = I(i,j) \tag{5}$$

The filtering process downsamples the image. To calculate a pyramid of difference-of-Gaussian images (luminance difference images, denoted ΔL_w^l), the image at level $l+1$ is upsampled into L_{wExp}^{l+1} and substracted to the image at level l:

$$L_{wExp}^{l+1}(i,j) = \sum_{m=-2}^{2}\sum_{n=-2}^{2} g(m,n) L_w^{l+1}\left(\frac{i}{2}+m, \frac{j}{2}+n\right) \tag{6}$$

$$\Delta L_w^l(i,j) = L_w^l(i,j) - L_{wExp}^{l+1}(i,j) \tag{7}$$

This results in a seven-level pyramid: the first six levels are bandpass images ΔL_w^l and the higher level is a low-pass image L_w^7.

3.3.2 *Contrast Detection.* We aim at preserving the contrast perception, which means that the observer should detect in the same way both the contrast of a pixel in the HDR image and the contrast of the same pixel in the processed and displayed LDR image. To quantify the contrast perception, we use the VL index for each level of the pyramid at every pixel (i,j), as:

$$V^l(i,j) = \frac{\Delta L_w^l(i,j)}{D\left(L_{aw}^l(i,j)\right)} \tag{8}$$

where V^l is the image of VL values at level l. The D function depends on the adaptation luminance L_{aw}^l (see Section 3.2.1), which is computed from the expanded upper-level image L_{wExp}^{l+1} as the mean

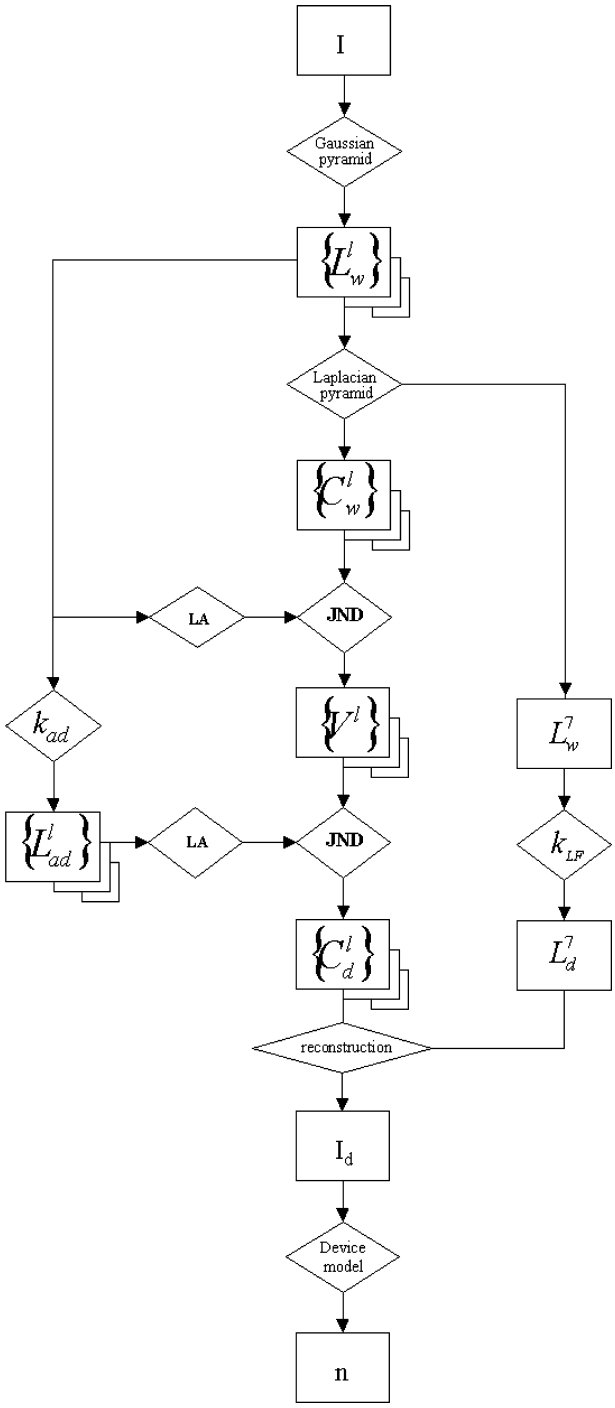

Fig. 2. Framework of the algorithm.

luminance over 1° in the visual field (S_p is half the number of pixels in 1°):

$$L^l_{aw}(i,j) = \sum_{m=-S_p}^{S_p} \sum_{n=-S_p}^{S_p} L^{l+1}_{wExp}(i+m, j+n) \tag{9}$$

The VL images are used to compute ΔL^l_d images (d for display) adapted to the display device on which the rebuilt image shall be displayed. The new contrast images are computed using Eq. (8) and the JND equation, to inverse the model: instead of L^l_{aw}, the new contrast values are computed with an estimation of the display adaptation luminance (see next section), in order to keep a constant VL in the HDR and LDR images. Displayed values are computed as:

$$\Delta L^l_d(i,j) = V^l(i,j) D\big(L^l_{ad}(i,j)\big) \tag{10}$$

3.3.3 *Adaptation Luminance.* The display adaptation luminance L^l_{ad} is computed from the world adaptation images L^l_{aw} of each channel l with a linear operator: a scale factor k^l_{ad} is used, one for each spatial frequency channel, in order to fit the HVS behavior.

$$L^l_{ad}(i,j) = k^l_{ad} L^l_{aw}(i,j) \tag{11}$$

Ward [Ward 1994] proposes such a scale factor k_W, intended to keep the detection threshold:

$$k_W = \left[\frac{1.219 + ((L_{dMax} - L_{dMin})/2)^{0.4}}{1.219 + (L_{aw})^{0.4}} \right]^{2.5} \tag{12}$$

where L_{dMax} and L_{dMin} are the extreme luminances that can be displayed by the device and L_{aw} is the log mean luminance of image I. $(L_{dMax} - L_{dMin})/2$ is an estimation of the adaptation luminance on the display device.

However it is not possible to use this scale factor as it is in Eq. (11). The first reason is that when the average luminance is small enough, as in night scenes, Ward's scale factor expands the luminance dynamic range so that nighttime images may appear like daytime images. More important for road applications, visual performances may be improved instead of being preserved.

In a previous version of our TMO [Grave and Brémond 2005], we adapted Ward's scale factor by clipping it to 1. However, this solution does not allways hold for nighttime 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}, and, second, because of the quantization, many contrasts in these scenes are destroyed. This leads us to accept scale factors higher than one. We choose the smallest value of the scale factor allowing the smallest perceptible contrast in the image to be displayed to be actually perceived. Let us denote L_{aw} an adaptation luminance in a HDR image and L_{ad} the corresponding adaptation luminance in the displayable image (Laplacian pyramid indexes are removed for easy reading). We consider the smallest visible luminance contrast ΔL^{JV}_{ad} (JV for just visible) in the displayed image on a background L_{ad}:

$$L_{ad} + \Delta L^{JV}_{ad} = k_{ad} L_{aw} + JND(k_{ad} L_{aw}) \tag{13}$$

We call $Z = LtoAV \circ AVtoL$ the quantization function. We can compute the actual displayed luminance difference $Z(L^{JV}_{ad}) - Z(L_{ad})$, when the device is asked to display L^{JV}_{ad} and a background L_{ad}:

$$Z\left(L^{JV}_{ad}\right) - Z(L_{ad}) = Z(k_{ad} L_{aw} + JND(k_{ad} L_{aw})) - Z(k_{ad} L_{aw}) \tag{14}$$

Assuming that this contrast has to be just perceived when L_{ad} is the adaptation luminance, in other words, assuming that $JND(L_{ad}) = Z(L_{ad}^{JV}) - Z(L_{ad})$, we can define:

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

The quantization issue is no more a problem for a given adaptation level L_{aw} if $\alpha(k_{ad}, L_{aw}) = 1$. The α function may be computed for the complete range of L_{aw} values, however, as the quantization issue emerges in road scenes in the darkest areas of nighttime images, we only consider the smallest value L_{wMin}, because it is the one which leads to the strongest constraint. Then, in our TMO, we chose the smallest value of k_{ad} such that $\alpha(k_{ad}, L_{wMin}) = 1$. If α remains lower than 1, we first explore the L_w domain and, if α is still smaller than 1, we choose k_{ad}, which result in the maximum value of α.

Finally, the adaptation luminance is computed with a k_{ad} value for each pyramid level, and then clipped by the two extreme displayable luminances, L_{dmin} and L_{dmax}.

3.3.4 *Image Recontruction.* The last level of the pyramid, being a low-pass image, is scaled with Ward's scale factor k_{LF}, which is k_{ad}^{W} clipped to 1 (see Eq. (12)). The image is then reconstructed, following the inverse process of the pyramidal decomposition: the last level is 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. Finally, a display device model *LtoVA* is applied to convert luminance values into grey levels. The device model is measured from the display device chosen for the LDR display, through a photometric calibration [CIE 1996].

Figures 3–5 show three road images processed by different TMOs from the computer graphic literature (see Section 2.2), as well as the TMO proposed in this paper. We cannot judge or compare these TMOs, based on the printed images, because they have been calculated to be displayed on a specific LCD device, under specific visual conditions. In addition, the HDR image cannot be printed, so that no comparison is possible that way. At this stage of our work, we need to evaluate our algorithm with an objective method.

4. TMO EVALUATION FOR A DETECTION TASK

Our main objective being the evaluation of a TMO for road vision applications, we focus on the experimental psychology approach instead of image metric methods, for two reasons. First, it may be unconclusive to use vision models to assess the quality of a TMO, which is already based on vision models: depending on the models included in the image metric, the results would change. Second, and more important, we are looking for a TMO evaluation as ecological as possible (in the sense of ergonomic psychology) considering an object detection task while driving, because such an evaluation would be of greater impact on the confidence one could have in the results of experiments using tone-mapped images for road visibility applications [Hoc 2001].

4.1 Experimental Setup

Unlike most experimental TMO psychometric evaluations in the computer science literature [Drago et al. 2003; Ledda et al. 2005; Yoshida et al. 2005], we assess the quality of our TMO with respect to the target detection performances of observers. We use a standard psychovisual experiment in order to measure a visual performance in terms of object detection with a "Landolt ring" task (see Fig. 6: the broken ring of Landolt is the simplest optotype, that is to say, a visual test, which leads to a certain point to form recognition). The ring is presented at the center of the image on an uniform square background of 1°, and a gap in the ring is then presented during 100 ms. People are asked to say where they saw

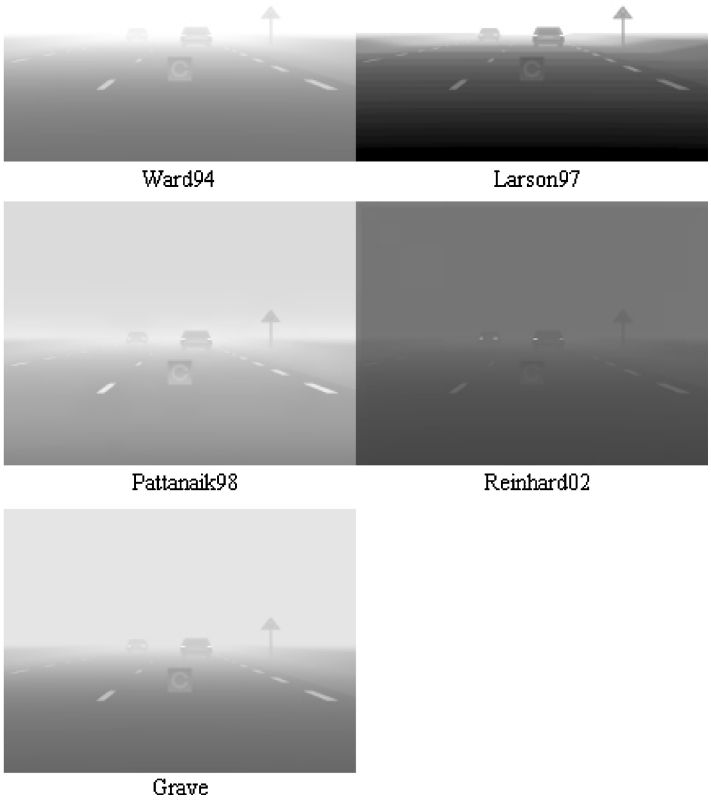

Fig. 3. Daytime image DayFog, $L_{wMin} = 113.18$ cd.m^{-2}, $L_{wMax} = 1392$ cd.m^{-2}, processed by five TMOs. The HDR image was computed following [Dumont and Cavallo 2004].

the gap (top, bottom, right or left of the ring), by means of a gamepad. The Landolt ring task is relevant for a large amount of detection tasks [CIE 1981].

We consider a reference situation as close to a driving situation as our laboratory conditions allow. It is designed in order to keep the main variables of a detection task and we make comparisons between this reference and comparison situations, in which tone-mapped LDR images are displayed. Our goal at this point is to assess the quality of the LDR displayed images with respect to the detection task. In the reference situation, we display an HDR image. In the comparison situations, the same image is displayed after having been processed by a TMO to fit the low dynamic range of the display device.

Concerning the reference situation, two experimental setup are possible: either a "physical" scene [Viénot et al. 2002; Yoshida et al. 2006], made of physical objects, among which a visual detection task would be performed, or a displayed HDR image [Ledda et al. 2005]. A "physical" scene may be closer to a driving situation than a displayed image, in particular, in terms of visual acuity, of depth vision, of visual complexity of the scene, etc. On the other hand, such a physical scene suffers from a number of problems. First, it is necessary to provide the LDR display device with images and to measure the visual performances staring at these images. Thus, the comparison between physical and displayed scenes implies an intermediary step, namely, converting the physical scene into a numerical one, and to build HDR luminance images from this numerical scene. This could be done for static images using photographic HDR acquisition techniques [Seetzen et al. 2004], but the complexity of the technical

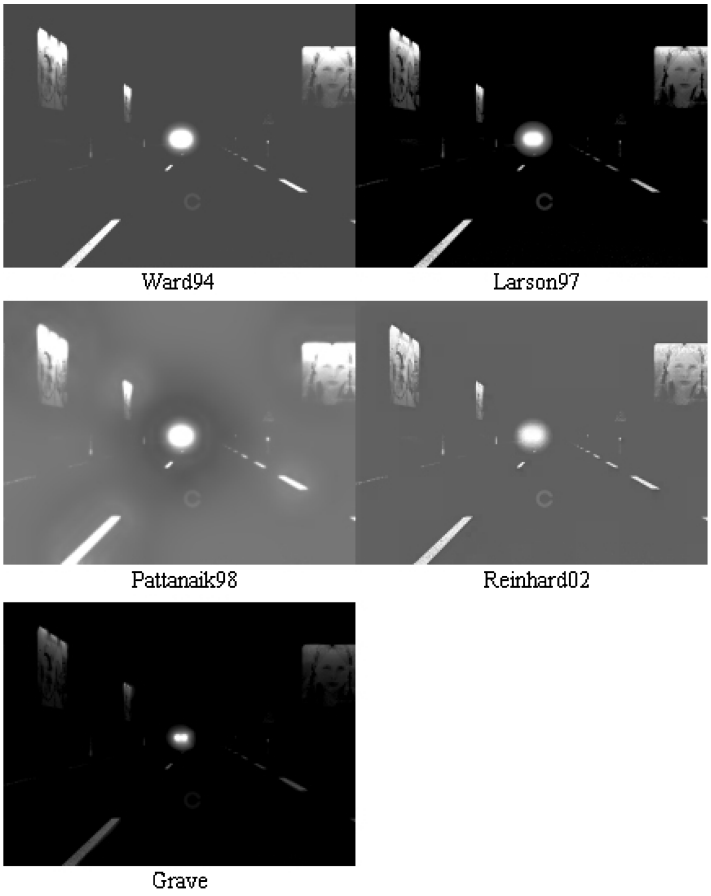

Fig. 4. Nighttime images: NightDrive ($L_{wMin} = 0.5$ cd.m^{-2}, $L_{wMax} = 478.4$ cd.m^{-2}) processed by five TMOs. The Nightdrive HDR image was computed following [Dumont and Cavallo 2004].

environment, which is necessary to perform a "physical" detection task (including a 100-ms stimulus presentation) was the main problem. On the other hand, the biases because of the use of an HDR image as a reference do not include the main independent variables, but only visual cues that were not considered as pertinent for our experimentation (see Sections 2.1 and 3.1). Therefore we decided to use HDR images as reference scenes. This seemed to us the optimal experimental framework, considering the problems raised by both approaches.

Concerning the visual task, the main point in designing the experimental setup was to decide whether the scene should include semantic contents (namely, a road scene). We decided to keep some semantic contents, even if it may be considered as a bias in a vision science approach: people may give biased answers (with respect to a scene with no semantic at all), because semantic cues may lead to specific expectations. However, because of the ecological situation of interest, we are more interested in the biased performances than in the "standard" ones. Another point which led us to use semantic cues is that considering the possible use of LDR images on a driving simulator, it is a key parameter to use reference images containing the same kind of luminance histogram, luminance dynamics, and spatial frequencies as in road scenes. An unwanted drawback of our choice is that the validity of the

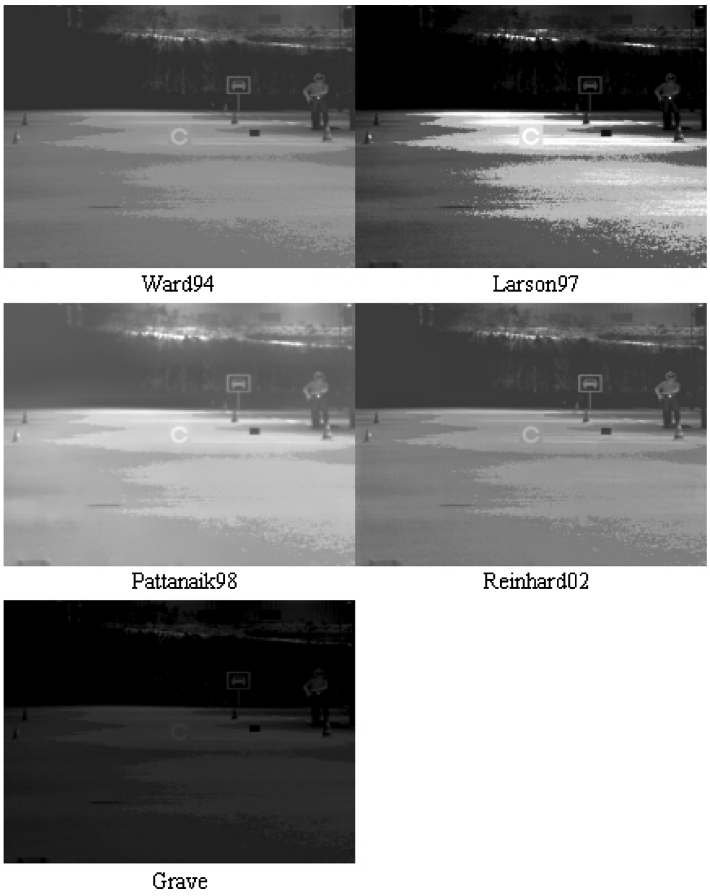

Fig. 5. Nighttime images: LN150W ($L_{wMin} = 0.25$ cd.m^{-2}, $L_{wMax} = 5.1$ cd.m^{-2}) processed by five TMOs. The LN150W HDR image was captured with a video photometer.

Fig. 6. Positions of the gap in the Landolt ring.

evaluation is restricted to driving situations, but this is understood and of little importance for road vision applications.

4.2 Apparatus and Method

Our experiments are designed in order to evaluate a TMO, together with a display device (through L_{dMin}, L_{dMax} and $LtoVA$) and a geometry for the display (through the pixels angular size). We compare the visual performances of a set of observers measured in two display configurations: first when the

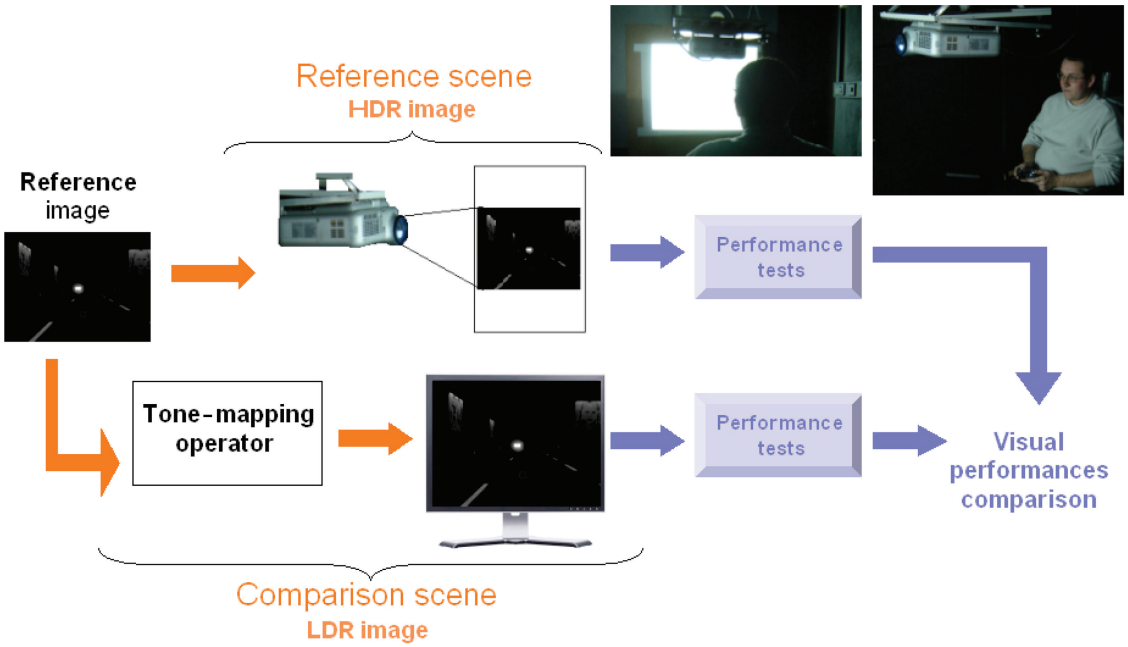

Fig. 7. Principle of the evaluation experiment.

observers are looking at a reference image (a HDR image); second, when a tone-mapped image is displayed on a LDR display device (see Fig. 7).

The performance obtained with the reference and the comparison images are compared for different TMOs: ours and four others (see Section 2.2):

—Ward's operator [Ward 1994], a nonparametric visual performance-based linear operator.

—Larson et al.'s [Larson et al. 1997], an nonparametric histogram adjustment operator using a TVI function. We use a restriction of this paper to the histogram adjustment method based on human contrast sensitivity.

—Pattanaik et al.'s [Pattanaik et al. 1998], which provides a nonparametric computational model for adaptation and spatial vision. We use a restriction of this paper to luminance images.

—Reinhard et al.'s [Reinhard et al. 2002], an heuristic model based on the *zone system* photographic technique. We use the local *automatic dodging and burning* version of this TMO with the default values given in the paper for the *key*, *scaling,* and *threshold* parameters.

Subjective evaluation experiments [Ledda et al. 2005] indicate that [Larson et al. 1997] and [Reinhard et al. 2002] provide good results for image comparisons. These algorithms are restricted, in our evaluation, to luminance image processing.

Following [Larson et al. 1997] and [Pattanaik et al. 1998], we model the glare effects because of light sources in the scene using Spencer et al's algorithm [Spencer et al. 1995]. This preprocessing step concerns all five TMOs.

These algorithms are well known in the TMO literature and are chosen for this evaluation step as an illustration of the variety of TMO design, as well as for their respective strong point for target detection: Ward's operator is the simplest TMO to use a detection threshold; Pattanaik et al.'s includes a local adaptation estimation following Peli's contrast image representation; Larson et al. and Reinhard

et al. both lead to good visual appearance indexes in Ledda et al.'s [Ledda et al. 2005] TMO evaluation. Larson et al. [Larson et al. 1997] is a global TMO and includes a TVI model, while [Reinhard et al. 2002] is based on a photographic technique.

The experiments are carried out in a dark room dedicated to psychovisual experiments. The room is painted black and has no window so that the illumination and the experimental conditions can be reproduced at any time. We use road images (see Figs. 3 and 5) in which a Landolt ring is inserted with an uniform square background of 1° width. The ring together with the background are called "the test" in the following.

As our TMO needs as an input some photometric properties of the LDR display device (a LCD screen), it is characterized on the basis of a CIE metrological protocol [CIE 1996] using a Minolta spectroradiometer CS 1000. The HDR images are displayed with a digital light processing (DLP) video projector (see Figure 7), and the DLP display device was characterized following the same metrological protocol, in several configurations (Contrast C, with range from 0 to 255, and Brightness B, with range from -64 to 64) depending on the road scene. For each display configuration, the *AVtoL* function is measured. The LCD used in this experiment has a dynamic range of 320 : 1 and a maximum luminance of 168 cd.m^{-2}. The DLP has a dynamic range around 1000 : 1 and, depending on the tuning configuration (C and B) and on the projection surface, a minimum luminance, which can be as low as 0.25 cd.m^{-2} and a maximum luminance, which can be as high as 1392 cd.m^{-2}. The angular size of the pixels in the experimental situation was 0.013° for both DLP and LCD displays.

The Landolt ring test greatly simplifies the driving visual task, but still leads to a fundamental aspect of this task and is often used as a reference task in road visibility and road lighting studies. This test is, however, complex and the detection of the gap depends on several parameters. Some of them are set in our experiment: the time during which the gap is shown is 100 ms.; the background luminance L_b is equal to the adaptation luminance computed with [Moon and Spencer 1945], assuming the gaze is centered on the ring; and the size of the gap e is 8 min of angle (see Figure 6). Two parameters may vary:

—the contrast (Weber fraction $C = \Delta L/L$) between the background luminance L_b and the ring luminance L_R. It is positive and takes eight different values from under the perception threshold to easily detected;

—the gap can take four different positions (see Figure 6).

The images into which the test is inserted are chosen to stand for typical driving scene with specific visibility conditions. They all have the same resolution 1024×768 pixels. Three typical driving conditions are tested. Each one addresses a specific problem that all TMOs have to face with road images:

—Luminance dynamic: a daytime driving scene with fog and visible rear foglights. To display this image, a uniform luminance is projected around the image to adapt the observer's peripheral vision and limit eyestrain. The DLP configuration is $C = 128$ and $B = 0$. The projection surface is 49×37 cm (see Figure 3, DayFog image).

—Quantization and high L_{dMin}: a nighttime urban driving scene without light source in the visual field. The DLP configuration is $C = 0$ and $B = -64$. The projection surface is 76×57 cm (see Figure 5, LN150W image).

—Luminance dynamic and quantization: a nighttime country driving scene with light sources in the visual field. The DLP configuration is $C = 128$ and $B = 0$. The projection surface is 76×57 cm (see Figure 5, NightDrive image).

Table III. Characteristics of the Subjects for DayFog, NightDrive, and LN150W Images.

	DayFog	**NightDrive**	**LN150W**
Women	2	2	1
Men	8	8	9
Total	10	10	10
$25 \leq age \leq 35$	7	7	3
$35 \leq age \leq 55$	3	3	2
$age \geq 55$	0	0	5
Total	10	10	10

4.3 Experimental Protocol

Thirty subjects were asked to perform a "Landolt ring" task, with the reference image and with the five tone-mapped images displayed on the LCD device. Three situations are tested, DayFog, NightDrive, and LN150W, with ten subjects each. For every test, we displayed an image with a full ring during 1 s, then, for 0.1 s, the gap is displayed at one of the four possible positions. The observer indicates on a gamepad the position of the gap he/she has perceived. We ask him/her to give an answer whether or not he/she has seen the gap. The answer is recorded and the next test is displayed. We showed 200 tests for the reference scene and for each of the TMOs (25 tests for each of the 8 contrasts). For a given observer, the 6 series of 200 tests were performed on different days, in order to avoid biases resulting from visual fatigue. The total number of ring presentations is $3 \times 10 \times 6 \times 200 = 36\,000$. Table III gives the characteristics of the subjects for the three images.

5. RESULTS

5.1 Uncertainty Estimation

For each of the three images, we consider an average observer whose visual performance is computed by averaging the performances of the ten observers. We compare his performances measured with the reference (HDR image) and those measured with the image processed by the TMOs.

For a better comparison, we estimate the margins of error around the reference values of the rate of correct answers, for each contrast value. For a contrast c, the tests are the realizations of a random variable, with expectation p_c, the performance of the average observer. Let us denote X_c a random variable that follows a Bernouilli distribution. X_c can take two values: 1 if the answer is correct (with probability p_c) or 0 if it is wrong (with probability $q_c = 1 - p_c$). Let us denote Y_c, a random variable corresponding to the number of correct answers for contrast c. In our case, $Y_c \in [0, 250]$, because we have 25 answers for each of the ten observers, so that Y_c follows a binomial distribution with parameters $n = 250$ and p_c. We know (binomial distribution properties) that $E(Y_c) = np_c$ and $V(Y_c) = np_c(1 - p_c)$. We consider $F_c^n = Y_c/n$ the rate of correct answers. $E(F_c^n) = p_c$ and $V(F_c^n) = p_c(1 - p_c)/n$. The Bienaymé–Tchebychev inequality gives for F_c^n:

$$P\left(\left|E\left(F_c^n\right) - F_c^n\right| \geq \varepsilon_c\right) \leq \frac{V\left(F_c^n\right)}{\varepsilon_c^2} \tag{16}$$

If we denote $a = \frac{V(F_c^n)}{\varepsilon_c^2}$, then $\varepsilon_c = \sqrt{\frac{p_c \times (1-p_c)}{na}}$. By fixing a, we can compute ε_c with:

$$\begin{aligned} &P\left(\left|F_c^n - p_c\right| \geq \varepsilon_c\right) \leq a \\ &\Leftrightarrow P\left[\left(F_c^n - \varepsilon_c\right) \leq p_c \leq \left(F_c^n + \varepsilon_c\right)\right] \geq (1 - a) \end{aligned} \tag{17}$$

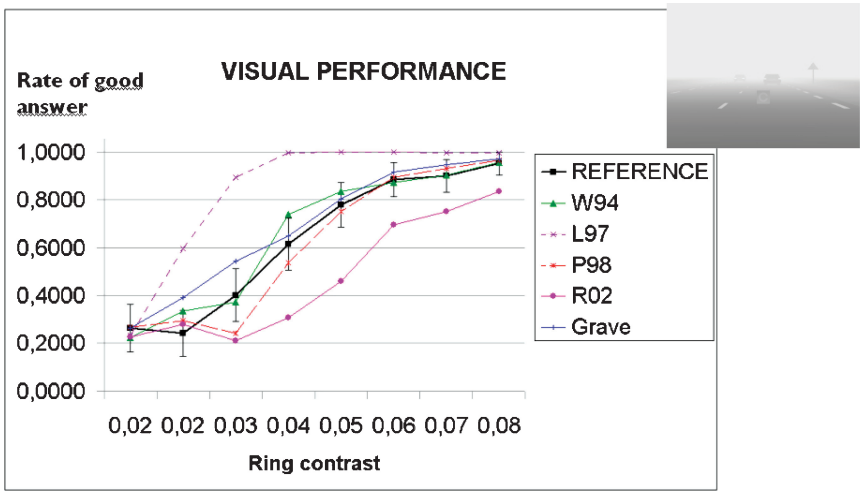

Fig. 8. Visual performances measured with the DayFog image, ploted against the luminance contrast (i.e., the Weber fraction). The reference image is compared to the image processed by five TMOs, including ours.

In the following, we set $a = 0.1$, which is the same as looking for a margin of error ($\pm\varepsilon_c$ around the value of F_c^n), which gather together 90% of the values that F_c^n can take. We can then estimate the values of $\varepsilon_c = \sqrt{\frac{p_c(1-p_c)}{0.1n}}$.

5.2 Results with the DayFog Image

Figure 8 shows the visual performance (the uncertainty estimations are only displayed for the reference situation, for easy reading) measured with the DayFog reference image and the five TMOs. The 50% detection threshold obtained with the reference image is for a contrast $c \in [0.03; 0.04]$.

The detection thresholds obtained with W94, P98, and our TMO are within this interval. For contrasts just below the perception threshold, we overestimate and P98 underestimates the detection rate. For contrasts beyond the threshold, W94 overestimates the detection rate. Whereas the rendering of contrast detection is rather satisfactory with these three operators, the two others do not meet the objective as far as this daytime road scene is concerned. L97 overestimates the detection rate, while R02 underestimates it.

5.3 Results with the LN150W Image

Figure 9 shows the visual performance (the uncertainty estimations are only displayed for the reference situation, for easy reading) measured with the LN150W reference image and the five TMOs. The 50% detection threshold obtained with the reference image is for a contrast around $c = 0.08$.

Two operators successfully render the visual performance: ours with a threshold slightly lower than 0.08 and P98, with a threshold slightly higher than 0.08. L97 enhances the contrast so that contrasts, which should not be detected, are seen. W94 and R02 also overestimate the contrast, with a threshold that is close to the minimum displayed contrast of 0.036.

5.4 Results with the NightDrive Image

Figure 10 shows the visual performance (the uncertainty estimations are only displayed for the reference situation, for easy reading) measured with the NightDrive reference image and the five TMOs. The 50% detection threshold obtained with the reference image is for a contrast $c \in [0.18; 0.28]$.

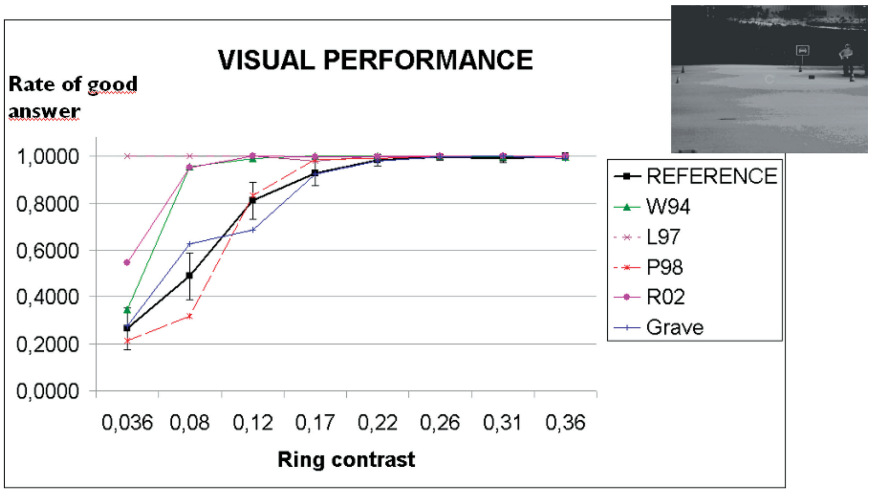

Fig. 9. Visual performances measured with the LN150W image. The reference image is compared to the image processed by five TMOs, including ours.

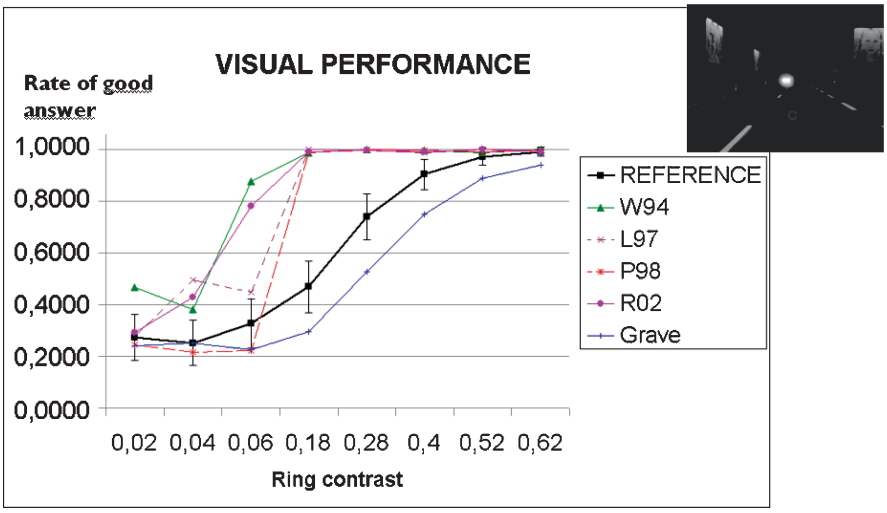

Fig. 10. Visual performances measured with the Nightdrive image. The reference image is compared to the image processed by five TMOs, including ours.

Our TMO gives a detection threshold of $c = 0.28$, which means that we underestimate the detection rate. However, the performance curve obtained with our operator has a shape close to the reference curve's shape. W94 and R02 enhance the detection rate, since their threshold is close to 0.04. P98 and L97 also overestimate the detection rate, but with thresholds closer to the reference, around 0.10. None of the tested TMOs fit the reference curve well, which is consistent with the fact that this image includes both the luminance dynamic and the quantization issue, which is, of course, harder to cope with for a TMO than a single issue.

Table IV. Comparison of Five TMOs (for a Detection Task and for Three Typical Road Images

	W94	L97	P98	R02	Grave
DayFog	G	B	G	B	G
LN150W	B	B	G	B	G
NightDrive	B	B	B	B	F

[a]Good, good; F, Fair; B, Bad.

6. DISCUSSION

Table IV sums up the performance tests for the three selected road images. These results suggest that our algorithm suits its purpose, except for the NightDrive situation, where the other TMOs we have tested did not give better results. This confirms that there was indeed a need for such a specific algorithm.

Two algorithms seem to stand out of the group of TMOs, P98 and ours, concerning the performance rendering for all three driving scenes. However, we cannot make general conclusions about TMOs quality out of this limited experiment, for at least three reasons:

1. We only tested three road images, chosen in order to emphasize three specific TMO issues, which are commonly addressed in road-images rendering (light sources, quantization, low luminance levels). This cannot be generalized without caution to other visual environments.
2. The algorithms chosen for this comparison are well known, but we cannot reject the idea that others may have given better results, even for target detection. For instance, Reinhard et al. [Reinhard et al. 2002] include parameters which were set to their default value. An optimization of these parameters may have lead to better results. However, road vision applications would benefit more from a nonparametric TMO, if its has to be used by "naive" users (in the computer graphics sense).
3. Our quality index is based on a visual performance task. It says nothing about the visual appearance of the tone-mapped images, which is the main purpose of most TMOs design and evaluation techniques.

Concerning the second point, it is obvious that our experiment cannot lead to the conclusion that our TMO is better than any others. However, considering the growing literature on TMOs for the past 20 years, an exhaustive comparison is not possible and we are not aware of a unique and broadly used TMO evaluation procedure in the computer graphic literature. Our point in this comparison was to assess our TMO to well-known algorithms, which are either based on psychophysical data, or with good evaluation indexes in experimental psychology experiments. The comparison shows, at least, that a TMO design devoted to a specific visual task may give good results for this task.

Concerning third point, our TMO was designed for a detection task, which calls for the visual performance of observers, so we evaluated it with a performance test. However, at a more general level, two classes of quality indexes can be used for perceptual evaluation: performance indexes (as presented above), based on error rates or on reaction times, and appearance indexes, which are purely subjective evaluations. Appearance indexes are widely used for TMO evaluation and are also important in road visibility applications considering the immersion feeling in driving simulators. We have thus completed the performance evaluation of our TMO with an appearance evaluation.

Our operator is compared to the same other TMOs, with an experimental protocol close to [Ledda et al. 2005]. 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. For a given observer, the reference image remains the same during the entire test. On the LCD device, two images processed by two different TMOs are displayed successively and the subjects are asked to select the one that they

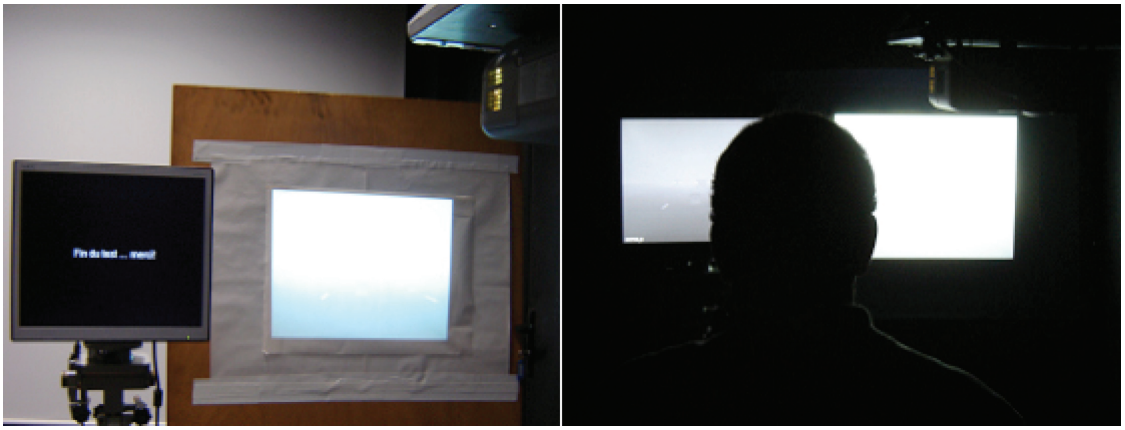

Fig. 11. The visual appearance rendering evaluation experiment.

Table V. Visual Appearance Tests: Score of the Five TMOs

	W94	L97	P98	R02	Grave	max.
DayFog	9	1.3	6.1	3.3	10.3	12.0
NightDrive	5.9	8.9	0.8	2.9	11.5	12.0
LN150W	9.4	0.8	3.1	6.8	9.9	12.0

feel is "closest" to the reference image. The observers may go back and forth from one image to another as many times as needed to make their choice. The images and the observers are the same as in the performance tests. The same five TMOs are compared, which results in ten different TMO pairs of images. Each pair is displayed three times, which means that the maximum number of times a TMO can be chosen is 12. Table V shows the average score of each operator, over the ten subjects.

Our operator is the first chosen with the NightDrive and DayFog images and, it is the first, almost equal with Ward's operator, in the LN150W image. With the restriction (1) and (2) above, these results tend to show that our operator keeps a fair image fidelity in terms of visual appearance, despite the fact that it is built on visual performance models. One may find that these results are unexpected, after Ledda et al.'s experiments [Ledda et al. 2005]. In our opinion, this kind of subjective evaluation is sensitive to the choice of the HDR images (see [Yoshida et al. 2006] for an exhaustive comparision over 25 images), so that we recommend to restrict the conclusions to road images. In addition, we cannot reject the possibility that the observers, having already been participating to a detection task with the same images, may have answered with a bias toward detection criteria.

A drawback of our TMO is that we noticed halo artifacts (contrast inversion) in the tone-mapped images, as well as with other local operators like Pattanaik et al's. Because of the control over the linear factors k_{ad}^{l} for the adaptation image transformation, halos were hardly perceived in the LDR images of road scenes processed by our operator; however, halos were clearly visible in the images processed by Pattanaik et al., which may explain the bad results this TMO obtained with the appearance evaluation (it has to be noted that this drawback does not disturb the visual performances as long as the visual task is not too close to a light source). Even if the contrast inversion is hardly visible in the displayed images, this is to our opinion the main problem we have to address in order to improve our TMO. However, going back to a global operator is not a valuable solution as we need to keep the visual performance fidelity.

7. CONCLUSION AND FUTURE WORK

We proposed a new tone-mapping operator, designed for a detection task, for road visibility, and road lighting applications. Our main objective was to preserve the observer's visual performance despite the luminance dynamic range compression, the possible high level of minimum displayable luminance, and the image quantization. To assess our algorithm and to check that it fulfills these conditions, we carried out a psychophysical experiment in order to compare visual performance measured with reference road images (HDR images), and with tone-mapped images. The comparison images were made of the reference images tone mapped and displayed on a LDR display device. We performed these tests with three typical road images: daytime and nighttime with and without direct light sources. Over the three scenes, considering a mean of ten observers for each, our operator seems to render the visual performance correctly, which is not always the case with four other TMOs from the computer graphic literature that we have tested. To complete the TMO assessment, we carried out appearance-rendering comparisons with the same images and observers. Our operator came first for the three driving scenes. This means, if we merge the results of both tests, that our TMO fulfills its objectives: it allows to display road images preserving the contrast perception and the general visual appearance.

We have showed that for a detection task, in roadlike visual conditions, our operator is effective. However, the comparison situation is too restricted to allow an extrapolation of these results to more general situations and visual tasks. Our experimental results only suggest that our TMO is well suited to the detection task for road images, which is consistent with the fact that the vision models used in the TMO directly address this visual task. However, in our opinion, the main conclusion in terms of design is that we propose a practical example of a TMO design driven by the user's needs (in our case, road engineering).

The next steps of this work should be at least in three directions. Even if color is not a critical parameter for a detection task, the ecological validity of our operator would need to extend the performance tests to color images and to motion detection, as well [Kelly 1985]. The second aspect is the integration of our algorithm into the visual loop of a driving simulator. This includes a real-time implementation of the algorithm and also a re-configuration of some aspects of the design of usual visual loops (including a photometrical calibration of the display systems) in order to cope with a photometrical description of the 3D databases [Brémond and Gallée 2002; Dupuis and Grezlikowski 2006]. Finally, we must admit that even if our reference images allowed to show significant differences between the TMOs, which were compared, they were not full HDR images, because of our limited DLP image display system. A promising future work would be to use a full HDR display device [Seetzen et al. 2004], as was done by [Ledda et al. 2004], to assess the quality of our TMO with more realistic reference situations.

ACKNOWLEDGMENTS

We would like to thank the 30 participants of the experimental study. We would also like to thank E. Dumont, J.-P. Farrugia, B. Péroche and F. Viénot for their help in the progress of this work, and the anonymous reviewers for many suggestions which improved the manuscript.

REFERENCES

Adrian, W. 1989. Visibility of targets: model for calculation. *Lighting Research and Technology 21*, 181–188.

Brémond, R. and Gallée, G. 2002. Image quality for driving simulation experiments. In *Proc. of Driving Simulation Conference 2002*. INRETS-Renault.

Burt, P. J. and Adelson, E. H. 1983. The laplacian pyramid as a compact image code. *IEEE Transactions on Communications 31,* 4 (Ap.), 532–540.

Campbell, F. and Robson, J. 1968. Application of fourier analysis to the visibility of gratings. *Journal of Physiology 197*, 551–566.

CIE. 1981. *An Analytic Model for Describing the Influence of Lighting Parameters upon Visual Performance*. Report 19/2.1. Commission Internationale de l'Eclairage.

CIE. 1988. *Roadsigns*. Report 74. Commission Internationale de l'Eclairage.

CIE. 1992. *Contrast and Visibility*. Report 95. Commission Internationale de l'Eclairage.

CIE. 1996. *The Relationship between Digital and Colorimetric Data for Computer-Controlled CRT Displays*. Report 122. Commission Internationale de l'Eclairage.

CIE. 1999. *Visual Adaptation to Complex Luminance Distribution*. Report 135/5. Commission Internationale de l'Eclairage.

DRAGO, F., MARTENS, W. L., MYSZKOWSKI, K., AND SEIDEL, H.-P. 2003. Perceptual evaluation of tone mapping operators. In *ACM SIGGRAPH 2003 Sketches & Applications*. ACM Press, New York. 1–1.

DUPUIS, M. AND GREZLIKOWSKI, H. 2006. Opendrive, an open standard for the description of roads for driving simulations. In *Proceedings of the Driving Simulation Europe*. INRETS-Renault, 25–36.

FERWERDA, J. A. 2001. Elements of early vision for computer graphics. In *IEEE computer graphics and applications*. IEEE, Los Alamitos, CA. 21–33.

FERWERDA, J. A., PATTANAIK, S. N., SHIRLEY, P., AND GREENBERG, D. P. 1996. A model of visual adaptation for realistic image synthesis. In *Proceedings of ACM SIGGRAPH '96*. ACM, New York. 249–258.

GINSBURG, A. P. 1986. Chapter 34: spatial filtering and visual form perception. In *Handbook of Perception and Visual Performance*, K. R. Boff, L. Kaufmann, and J. P. Thomas, Eds. Wiley, New York. 1–41.

GRAVE, J. AND BRÉMOND, R. 2005. Designing a tone mapping algorithm for road visibility experiments. In *APGV '05: Proceedings of the 2nd Symposium on Applied Perception in Graphics and Visualization*. ACM Press, New York. 168.

HILLS, B. L. 1980. Vision, visibility, and perception in driving. *Perception 9*, 183–216.

HOC, J.-M. 2001. Towards ecological validity of research in cognitive ergonomics. *Theoretical Issues in Ergonomic Science 2,* 3, 278–288.

ISHIDA, T. AND IRIYAMA, K. 2003. Estimating light adaptation levels for visual environments with complicated luminance distribution. In *Proceedings of the 25th Session of CIE*. D, vol. 1. CIE. 10–13.

KELLY, D. H. 1985. Visual processing of moving stimuli. In *Journal of the Optical Society of America*. 2, vol. 2. 216–225.

KUANG, J., YAMAGUCHI, H., JOHNSON, G., AND FAIRCHILD, M. 2004. Testing hdr image rendering algorithms. In *Proceedings of IS and T/SID 12th Color Imaging Conference: Color Science and Engineering Systems, Technologies, Applications*. SID, 315–320.

LARSON, G. W., RUSHMEIER, H., AND PIATKO, C. 1997. A visibility matching tone reproduction operator for high dynamic range scenes. *IEEE Transactions on Visualization and Compter Graphics*. 291–306.

LEDDA, P., SANTOS, L. P., AND CHALMERS, A. 2004. A local model of eye adaptation for high dynamic range images. In *AFRIGRAPH '04: Proceedings of the 3rd International Conference on Computer Graphics, Virtual Reality, Visualisation and Interaction in Africa*. ACM Press, New York. 151–160.

LEDDA, P., CHALMERS, A., TROSCIANKO, T., AND SEETZEN, H. 2005. Evaluation of tone mapping operators using a high dynamic range display. In *Proceedings of ACM SIGGRAPH '05*. 249–258.

LI, Y., SHARAN, L., AND ADELSON, E. H. 2005. Compressing and companding high dynamic range images with subband architectures. *ACM Transactions on Graphics 24,* 3, 836–844.

MOON, P. AND SPENCER, D. 1945. The visual effect of non-uniform surrounds. *Journal of the Optical Society of America 35,* 3, 233–248.

MOVE. 2005. *Performance based model for mesopic photometry*. Helsinki University of Technology.

PATTANAIK, S. N., FERWERDA, J. A., FAIRCHILD, M. D., AND GREENBERG, D. P. 1998. A multiscale model of adaptation and spatial vision for realistic image display. *Computer Graphics 32,* Annual Conference Series. 287–298.

PATTANAIK, S. N., TUMBLIN, J., YEE, H., AND GREENBERG, D. P. 2000. Time-dependent visual adaptation for fast realistic image display. In *In Proceedings of ACM SIGGRAPH 2000*. ACM, New York. 47–54.

PELI, E. 1990. Contrast in complex images. *Journal of the Optical Society of America 7,* 1 (Oct.), 2032–2040.

REINHARD, E., STARK, M., SHIRLEY, P., AND FERWERDA, J. 2002. Photographic tone reproduction for digital images. In *Proceedings of ACM SIGGRAPH 2002*. 267–276.

REINHARD, E., WARD, G., PATTANAIK, S., AND DEBEVEC, P. 2005. *High dynamic range imaging: acquisition, display, and image based lighting*. Morgan Kaufmann.

SEETZEN, H., HEIDRICH, W., STUERZLINGER, W., WARD, G., WHITEHEAD, L., TRENTACOSTE, M., GHOSH, A., AND VOROZCOVS, A. 2004. High dynamic range display systems. In *Proceedings SIGGRAPH 2004*. 3, vol. 23. 760–768.

SHAPLEY, R. 1991. Chapter 12: Neural mechanisms of contrast sensitivity. In *Spatial Vision*, D. M. Regan, Ed., MacMillon, New York. 290–306.

SMITH, K., KRAWCZYK, G., MYSZKOWSKI, K., AND SEIDEL, H.-P. 2006. Beyond tone mapping: Enhanced depiction of tone mapped hdr images. In *Computer Graphics Forum, Proceedings of Eurographics 2006*. 3, vol. 25. Eurographics, 427–438.

SPENCER, G., SHIRLEY, P., ZIMMERMAN, K., AND GREENBERG, D. 1995. Physically-based glare effects for computer generated images. *Proceedings ACM SIGGRAPH '95*. 325–334.

VIÉNOT, F., BOUST, C., COSTA, E. D., BRÉMOND, R., AND DUMONT, E. 2002. Psychometric assessment of the look and feel of digital images. In *Driving Simulation Conference*. INRETS-Renault.

WARD, G. 1994. A contrast-based scalefactor for luminance display. *Graphics Gems IV*, 415–421.

YOSHIDA, A., BLANZ, V., MYSZKOWSKI, K., AND SEIDEL, H.-P. 2005. Perceptual evaluation of tone mapping operators with real-world scenes. In *Human Vision and Electronic Imaging X, IS& T/SPIE's 17th Annual Symposium on Electronic Imaging*. SPIE Proceedings Series, vol. 5666. San Jose, CA. 192–203.

YOSHIDA, A., MANTIUK, R., MYSZKOWSKI, K., AND SEIDEL, H.-P. 2006. Analysis of reproducing real-world appearance on displays of varying dynamic range. In *Computer Graphics Forum, Proceedings of Eurographics 2006*. 3, vol. 25. Eurographics, 415–426.

Received July 2006; revised Sept. 2006/Jan. 2007; accepted April 2007