DOI: 10.1111/cgf.12687
Eurographics Symposium on Rendering 2015
J. Lehtinen and D. Nowrouzezahrai
(Guest Editors)

Volume 34 (2015), Number 4

Modeling Luminance Perception at Absolute Threshold

Petr Kellnhofer[1] Tobias Ritschel[1,2] Karol Myszkowski[1] Elmar Eisemann[3] Hans-Peter Seidel[1]

[1]MPI Informatik [2]Saarland University [3]Delft University of Technology

Figure 1: *We simulate the change of image appearance between photopic conditions* (left) *and appearance in scotopic conditions close to the absolute threshold* (right)*, where consistent vision fades into temporally varying (not reproducible in print) noise.*

Abstract

When human luminance perception operates close to its absolute threshold, i. e., the lowest perceivable absolute values, appearance changes substantially compared to common photopic or scotopic vision. In particular, most observers report perceiving temporally-varying noise. Two reasons are physiologically plausible; quantum noise (due to the low absolute number of photons) and spontaneous photochemical reactions. Previously, static noise with a normal distribution and no account for absolute values was combined with blue hue shift and blur to simulate scotopic appearance on a photopic display for movies and interactive applications (e.g., games). We present a computational model to reproduce the specific distribution and dynamics of "scotopic noise" for specific absolute values. It automatically introduces a perceptually-calibrated amount of noise for a specific luminance level and supports animated imagery. Our simulation runs in milliseconds at HD resolution using graphics hardware and favorably compares to simpler alternatives in a perceptual experiment.

Categories and Subject Descriptors (according to ACM CCS): I.3.3 [Computer Graphics]: Picture/Image Generation—Viewing algorithms

1. Introduction

The human visual system (HVS) adapts to absolute luminance through several orders of magnitude; we can perceive a bright daylight scene as well as a moonless night. Appearance drastically changes for different absolute levels: at night (scotopic) color and acuity are reduced and a shift towards blue tones is perceived, when compared to the same scene in daylight (photopic) conditions.

In visual arts, cinematography or interactive applications (e.g., games), the change of appearance is often simulated to convey the illusion of a certain adaptation level despite being in a different display condition. A skillful painter is able to depict a scene shown on a photopic canvas as if it actually was scotopic. The same holds for movies, where the so-called "Day-for-night" effect is used since the early days of cinema. For computer applications, techniques like tone mapping can convey a scotopic impression. In all cases, it is important to point out that adaptation effects are qualitatively reproduced

and might differ in quantity: night scenes are blurred only enough to become noticeable and not as much as a strict HVS simulation would require, which would lead to an unpleasant viewing experience.

Computer graphics has now routinely modeled the shift from photopic over in-between mesopic to scotopic conditions [FPSG96, PFFG98, DD00, PTYG00, TSF02, KP04, KO11, WM14] but the scotopic regime (vision close to its absolute threshold, e. g., a moonless night), has received little attention. Remarkably, the absolute threshold is close to the physical limitations of light itself; most dark-adapted subjects reliably detect flashes of light resulting in as little as 5 to 10 photons total on the retina during an integration time of 100 ms [HSP42]. Appearance under such conditions is substantially different from all other conditions. While scotopic vision can still rely on the retina as a classic noise-free sensor described by scalar ray optics, for close to absolute threshold, receptor noise due to the particle nature of light becomes apparent and requires accounting for quantum statistics.

In this paper, we complement day-for-night tone mapping to account for the effects encountered close to the absolute threshold. After explaining the background in physics, neuroscience, and human vision (Sec. 2), as well as reviewing the state of the art in modeling human scotopic perception in computer graphics (Sec. 3), we propose a neurophysiologically-motivated model of rod receptor noise which adds temporal variations to the image, as expected to be experienced in scotopic conditions close to the absolute threshold (Sec. 4). We then present the related computational aspects, involving a photon-accurate retinal image representation, an efficient rod noise generation drawn from image content-dependent distributions, and a temporal rod-signal integration (refSec-OurApproach). Using graphics hardware, our model requires 18 ms for an HD image and we compare our results to different alternatives in a perceptual evaluation (Sec. 6).

2. Background

Human vision is based on translating light into nerve signals. Light can be modeled as rays, waves, individual particles, or as their quantum statistics. In this article, different from a commonly taken viewpoint in computer graphics and vision, we choose the quantum-statistics point of view. Here, "light" for a space-time interval is not a single value anymore, but modeled as a distribution of probabilities to observe a certain number of quanta.

Light enters the human eye through the pupil, and is subject to different scattering and absorption events, before it reaches the retina, which is covered by receptors converting incoming light to electric signals. Two types of receptors, cones and rods, exist, whose performance varies drastically with absolute luminance [Shl37, KWK09]. We have to distinguish *photopic* (10^8 to 3 cd / m^2), *mesopic*, (3 to 0.1 cd / m^2), *scotopic* (0.1 to 10^{-6} cd / m^2) vision, and scotopic vision *close to absolute threshold* (less than 10^{-3} cd / m^2).

In photopic conditions, *cones* are active. They have an uneven distribution with a strong peak in the fovea and contribute to color vision. They are inactive at night [Pal99], and we will assume they do not contribute to the scotopic effects modeled in this work. We also do not consider the mesopic range, in which both rods and cones are active.

In scotopic night vision, only *rods* are active. They have a different response to light of different wavelengths [Wal45] and do not contribute to color vision. Their peak density is lower than for cones, but their distribution over the retina is more uniform and shows a slower falloff with the retinal eccentricity. Modelling these differences is the foundation of many day-for-night tone mappers [DD00, PTYG00, TSF02, KP04, KO11, WM14].

Rods convert light into nerve signals using a *photochemical* cascade. Each rod contains rhodopsin, that is isomerized by exposure to light, resulting in a small change of potential to become a nerve signal [Alp71]. In each following step of the cascade, non-linear functions amplify the signal, while at the same time suppressing noise. The temporal aspects of photo-transduction are the cause of afterimages [RE12].

Not all photons hitting a receptor are actually transduced into an electrical signal (false negative) because it might happen that the photon does not hit a rhodopsin molecule. The ratio of transduction (ca. 0.06–0.2, [HSP42]) is called *quantum efficiency*. At the same time, it happens that rhodopsin is transduced in the absence of light (false positive) [Bar56]. These aspects will be detailed in our treatment of near-absolute-threshold light levels in Sec. 4.

Finally, other entoptic phenomena, which are not directly caused by light in the common way, such as floaters, phosphenes, visual snow, the blue-field entoptic effect [RP80], or afterimages [RE12], can occur under specific conditions but are not related to scotopic vision and will not be modelled in this work.

3. Previous work

In this section, we discuss tone-mapping solutions for night scenes. In particular, we focus on the role of perceived noise in scotopic vision and, in this context, we overview other sources of noise in images such as sensor noise, film grain, and synthetically generated noise, which under certain conditions can improve perceived realism and quality. Finally, we discuss photon-accurate eye modeling as is required near absolute threshold, which is central for this work.

Tone mapping: Night scene depiction A key goal of tone mapping operators (TMO) is to reproduce scotopic scene appearance on a photopic display [RWD*10] by simulating a blue-shift and the loss of color vision, visual acuity, contrast and brightness characteristic for night vision [FPSG96, PFFG98, KP04, KO11, WM14]. Typically, such

simulations cover higher levels of scotopic luminance (0.001–0.1 cd/m^2) including the transition to mesopic conditions, while luminance levels near absolute thresholds are not specifically addressed. Furthermore, the time-course of adaptation [DD00, PTYG00], the bleaching phenomenon [GAMS05], or stereovision in darkness [KRV*14] were modeled in computer graphics.

Nightly impressions have been convincingly reproduced in painting [Liv02], digital arts, computer games, and feature films without referring to any rigorous simulation of scotopic vision. Empirical solutions inspired by "day-for-night" shooting have been proposed by Thompson et al. [TSF02]. The success of empirical techniques indicates that rigorous simulations of scotopic vision not always lead to a subjectively optimal night-like look, especially in photopic conditions. Consequently, our strategy is to apply psychophysical data when available, and otherwise refer to empirical techniques, including the case when such data does not generalize to images presented on photopic displays.

Most importantly, Thompson et al. [TSF02] observed that adding noise to day-for-night tone mapping can improve the scotopic impression. They add static, zero-mean, uncorrelated Gaussian noise with a fixed standard deviation to each pixel, to achieve subjectively optimal visual results. Still, it is not clear how to apply their approach for animated content where calibration in absolute luminance is crucial such as close to absolute thresholds. An example of a video showing a transition (Fig. 1, left to right) from photopic over mesopic conditions down to scotopic conditions near the absolute threshold illustrates the two remaining main challenges: First, a transition from a noise-free image over subtle noise to a state in which only grey noise is expected to remain. To this extent, we introduce a calibration by absolute luminance not available from previous work. Second, changing image content, e. g., a simple camera pan, will require the noise to change. A simple overlay would result in a "shower door effect" [KP11]. Instead, we model accurate change dynamics, principled by physiological data to feature additive and multiplicative components (Sec. 4).

Image noise Noisy images are often undesirable in image synthesis and imaging applications, where denoising techniques are common. However, noise can be explicitly added to enhance perceived image sharpness [JF00]. Fairchild and Johnson [FJ05] have hypothesized that noise as a repetitive pattern tends to be suppressed by the HVS, which might subjectively enhance image saliency. In general, procedural noise is often used in image synthesis to enhance the visual richness of rendered images [LLC*10].

Sensors in digital cameras are prone to different temporal and spatial sources of noise [Jan01]. In particular, temporal photon and dark current shot noise show similarities to the nature of noise in rods (refer to Sec. 4) and are also modeled via Poisson distributions. Readout noise could be considered as an analog of retinal circuitry processing beyond rods and cones, which we ignore in this work.

The exposure and development of silver-halide crystals dispersed in the emulsion of analog films results in forming tiny blobs of metallic silver or dye clouds, which creates the familiar film grain effect whose density tends to follow a Gaussian distribution [Alt77]. Film grain as a form of noise might be manipulated or even intentionally added for technical or artistic reasons in movie post-production [Sey11] and can be acquired through film-stock scanning or synthesized following its noise-power spectrum [SS07, GLC13]. Stereoscopic processing of film grain was recently described by Templin et al. [TDMS14]. Similarly, in digital photography, the suppression of a synthetic look and masking of digital sensor noise are desirable [KMAK08]. Simulating the complete photographic process, including important characteristics of photographic materials, such as its response to radiant energy, spectral sensitivity, emulsion resolution, and graininess [GM97] can render results very realistic. While some analogies to our work are apparent, we refer to photon-accurate simulation, extremely low light levels are not supported in [GM97].

Photon-accurate eye modeling Deering [Dee05] models individual cones and accounts for photon events to generate per-cone photon counts that have been photoisomerized by unbleached photopigments. Hereby, cone-level rendering can be simulated for any physical display configuration and photopic image content. Our goals are similar but we focus on rods in scotopic vision near absolute threshold conditions, which has a strong impact but can be ignored in photopic conditions. We also share the limitation that retinal circuitry beyond cones and rods is ignored.

4. Modeling rod noise close to absolute thresholds

While the source of scotopic noise is well understood in physiology, it has not yet been considered in computer graphics, where noise is added in an ad-hoc way. This section introduces the reader with a computer-graphics background to the physiology of luminance perception at levels close to the absolute thresholds.

Absolute threshold On average, 60 % of flashes with 510 nm wavelength, a duration of 1 ms, emitting 54–148 photons in total towards a retinal area covering 500 receptors located off the fovea (which has no rods) will be detected by dark-adapted subjects [HSP42]. The fraction of photons that actually reach the retina is then only 10%. The key result of this experiment is that, close to absolute threshold, answers can only be given with certain probabilities, not with absolute certainty. In consequence, photon counts are related to detection likelihoods via *receiver operating curves* (ROCs).

Quantization noise The seminal work of Hecht [HSP42] has shown that the quantization of light into photons has

actually a practical perceivable consequence. In conditions close to the absolute threshold, photon count is important as, for rare discrete random events, noise is to be expected. Such noise can be modeled by a Poisson distribution, which estimates the probability density function P of observing k events given the expected number of such events μ:

$$P(k,\mu) = \frac{\exp(-\mu)\mu^k}{k!}.$$

The probability of observing Θ or more events is P's tail distribution (complementary cumulative distribution function)

$$F(\Theta,\mu) = \sum_{k>\Theta}^{\infty} P(k,\mu).$$

The probability of seeing a flash of N photons per unit time (integration time) at the cornea is

$$F_{\text{Quant}} = F(\Theta, qN), \tag{1}$$

where Θ is the minimal number of photons that can be perceived and q is the quantum efficiency. Such noise is qualitatively multiplicative (i. e., its magnitude depends on the image), as it depends on the actual number N of photons at the cornea. Hecht et al. [HSP42] have fitted the parameters of their ROC measurements against such a model and found $\Theta \approx 6$ and $q \approx 0.06$. Note, that for $N = 0$ the probability is zero, which cannot explain seeing noise in the absence of light. Furthermore, a quantum efficiency of $q \approx 0.06$ is judged to be too low with respect to other physiological data [FSR05]. Consequently, the model needs to be extended.

Photon-like noise An alternative source of noise has been identified in spontaneous photo-transduction [Bar56, BLY79, AF77]. Once in two thousand years, a rhodopsin molecule is isomerized without any reason, leading to false-positive responses, which becomes important when explaining the perception of noise in the absence of all light. There are 60,000,000 rods [JSN92] and given 2,000,000,000 rhodopsin molecules in each [YMB79], results in 0.032 events per second and rod. While this is little compared to the excitation rates above absolute threshold (high N), it is perceivable close to absolute threshold where N and the rate of such spontaneous photo-transductions become similar. The probability of seeing a flash due to such photon-like chemical events is

$$F_{\text{Dark}} = F(\Theta, D) \tag{2}$$

where D is a dark-noise constant, which characterizes the rate of spontaneous photo-transductions. This noise is qualitatively additive (it does not depend on the actual number N of photons but on a constant D) and could explain perceived noise in the practical absence of light. When fitting behavioral data [TPV*82] to such a model, one can find $\Theta \approx 40$ and $D \approx 50$. The best fit however, is produced by a model that accounts for both quantum and dark noise.

Combined noise Lillywhite [Lil81] has shown physiological evidence that photo-transduction near absolute threshold is in fact not a Poisson process. A Poisson process assumes that events are statistically independent. This does not hold as bleaching causes a non-linear response of the photoreceptors to consequent photons [GAMS05]. A better model can be obtained if the noise is assumed to be a combination of both quantization and photon-like noise. The probability distribution of observing exactly k photons is given by

$$P_{\text{All}}(k,N) = P(k, qN + D)F(\Theta, \alpha k), \tag{3}$$

where α is a *constant of growth*. Fitting to behavioral data yields $\alpha = 0.5$, $q = 0.2$, $D = 19$, and $\Theta = 15$ [FSR05], which is in good agreement with all physiological evidence [TPV*82].

5. Computational model

Overview The input to our system is a sequence of HDR images [RWD*10] storing "retinal" radiance, i. e., after including the eye's complete optical transfer. The simulated noise is added to an LDR image, produced from the HDR image by a day-for-night tone mapping of choice, leading to changes in chroma, saturation and acuity. Decoupling noise and tone mapping allows us to maintain full control over the appearance.A modular design also leads to easy and efficient integration into existing systems. For all results in this paper, we used the tone mapping by Thompson et al. [TSF02]. The output is an LDR image sequence to be displayed on an LDR display at photopic levels, which is perceived as similar to a scotopic experience close to absolute threshold.

Fig. 2 summarizes the computational pipeline from photons emitted by an HDR image to the triggered rod responses.First, the HDR input is converted into photon counts per unit time and area (Sec. 5.1). Next, the according retinal response is simulated (Sec. 5.2). As light perception depends on an integration period, which is particularly long in night vision (Bloch's law [Blo85]), we also consider eye motion (Sec. 5.3).

5.1. Photon counts

Close to absolute thresholds, the actual number of photons is important. Hence, we need to convert the image or the frames of an image sequence into photon counts per time and receptor. We follow the derivation of Deering [Dee05, Sec. 6].For our case, we account for the increase in pupil size [WY12] and the spectral shift in the HVS sensitivity when we derive photometric quantities. Hereby, we can estimate the number of photons that reach the retina through the pupil from a given screen pixel under standard viewing conditions. Since the data we use in Sec. 5.2 are acquired for the wavelength λ = 507 nm [HSP42], we assume the whole multi-spectral radiant power of the image to be concentrated at this wavelength. Given the relatively small range of visible frequencies, the effect on photon count and, thus, noise appearance is minor. Especially, since spectral radiance is not always available, it

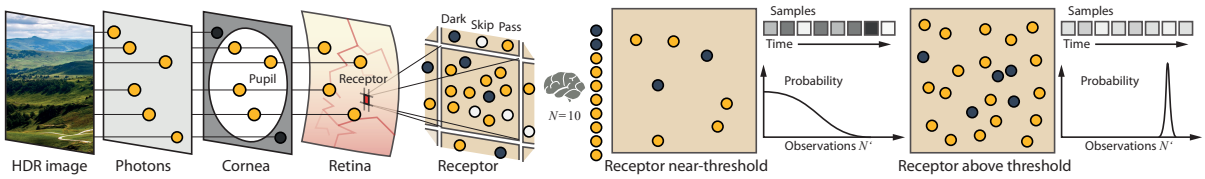

Figure 2: *Luminance perception at absolute thresholds* (Left to right). *Starting from the input HDR image, we compute the number of photons reaching the retina per unit time. At a single receptor, a photon either contributes to luminance perception* (yellow) *or is skipped (quantum efficiency)* (white). *Additionally, dark photon-like events* (blue) *contribute to the perceived sum, here ten. Near absolute threshold, the probability distribution is wide* (curve). *Luminance samples drawn over time according to this distribution are unstable and vary* (grey squares). *Well above absolute threshold, the probability distribution is peaky. Samples drawn from this distribution are stable and very similar to the expected value.*

is a reasonable approximation and represents the peak of the scotopic luminous efficiency $V'(\lambda)$ [Wan95], which could otherwise be used to compute wavelength-dependent luminous flux.

We then derive the approximate number of rod cells covered by a pixel's projection on the retina. To this extent, we assume a density of 100,000 rods / mm^2. We chose this value as a representative average density, since the rod acuity peak has a density of 150,000 rods / mm^2 at the eccentricity of 1.5 mm from the fovea center [JSN92, MKN84]. Combining the number of rod cells and photon count, we derive the per-receptor estimate. The full derivation is detailed in the supplemental material, leading to a single coefficient ϕ to convert scotopic luminance L into photon count per retinal region N and 100 ms integration time. Typically, $\phi \approx 1.2 \cdot 10^5$ for $L = 10^{-3}$ cd / m^2, hence $N = \phi L \approx 120$ photons.

Discussion The goal of this work is not a strict simulation of a complete perceptual pipeline for night vision, although our spatio-temporal model of handling photons at rods could potentially serve as input for higher-order processes. Using a photon-unit scale is a means to offer control over the day-for-night processing.

5.2. Simulation

The simulation is performed independently for all pixels in time steps matching the rendering framerate. Relying on our analytical model, this choice is both practical and performance-efficient. In the following, we will discuss the simulation outcome for a single receptor.

Eq. 3 is used to sample the number N' of photons perceived, depending on the number N of photons. It is not a Poisson process (contrary to simple shot-noise in Eq. 1, or dark noise-model in Eq. 2) and analytically drawing samples is not straightforward. As the evaluation is needed for all pixels per frame, an efficient procedure is required. To make sampling tractable, we use an inversion via a lookup table.

First, the values $P_{\mathrm{All}}(k,N)$ of Eq. 3 are tabulated for all values 0 to k and all values 0 to N. From this table, a complementary cumulative sum $F_{\mathrm{All}}(k,N) = \sum_{i=1}^{N} P_{\mathrm{All}}(k,i)$ is created numerically. Note, that each row in Eq. 3 is already a PDF and its integral is 1. The inverse of each row, $F_{\mathrm{All}}^{-1}(\xi,N) = \min\{k | F_{\mathrm{All}}(k,N) > \xi\}$, is stored as a look-up table.

The lookup table is constructed offline, but we also provide it in the supplemental material. To convert the physical photon count N into a photo-transduced photon count N', a random number $\xi \in [0,1]$ is generated and used to look up $N' = F_{\mathrm{All}}^{-1}(\xi,N)$ in constant time.

As the values $\alpha = 0.5$, $q = 0.2$, $D = 19$ and $\Theta = 15$ were derived for stimuli that covered 500 receptors, and a duration of 100 ms [HSP42, Lil81], the number N' computed above is valid for 500 receptors and 100 ms. However, we would need to apply a conversion to a single receptor, but computing the response of every individual receptor is computationally costly. Further, we actually should consider ρ receptors covered by a pixel ($\rho \approx 5$ for the display in our experiment). To accelerate the computation, we assume that the probability for an observation is uniform in a spatial neighborhood of 500 receptors in a time window of 100 ms and that observation events are independent between different receptors. Under these conditions, the probability that a pixel observes M events is given by the binomial distribution

$$P_{\mathrm{Final}}(M,N') = \binom{M}{N'} (\rho/500)^M (1-\rho/500)^{N'-M}.$$

Again, a single sample M' is drawn from this distribution using $M' = F_{\mathrm{Final}}^{-1}(\xi,N')$ using the inversion method and a lookup table for each N'.

Finally, the number of transduced photons for this pixel M' is converted back to a displayable value. At this point, we have to account for the factor ϕ that relates luminance L' and photon counts, as well as for the quantum efficiency that reduced the photon count due to the eye optics: $L' = \phi^{-1} \cdot q^{-1} \cdot (\rho/500)^{-1} \cdot M'$. In order to preserve chroma, the tone-mapped RGB values are first converted to YCrCb, the noise is applied to the luminance Y and the resulting Y'CrCb is

converted back to RGB. The noise is determined by the ratio of the photon count M' and the expected photon count given by the HDR luminance L. It is applied to Y as a combination of gain and bias, where the gain represents the multiplicative noise from the light quantization and depends on the size of qN, and the bias represents the additive photon-like noise and depends on the dark-noise constant D, as well as a noise baseline K

$$Y' = \frac{M'}{(qN+D)\rho/500}\left[\left(\frac{qN}{qN+D}\right)Y + \left(1-\frac{qN}{qN+D}\right)K\right].$$

Applying the noise to a toned-mapped image provides fine appearance control, as we can choose the noise intensity K in totally black regions, where no evidence about the absolute scale is available. $K = 0.06$ was used in our results. Photopic and mesopic conditions are practically noise-free and seamlessly covered by our simulation because $L \approx L'$, as dark noise can be neglected when $N \gg D$ and the standard deviation of quantum noise is small for $N \gg 0$. Because both, Poisson and binomial distributions, converge to the normal distribution for sufficiently large samples, we keep only up to 1000 values in the lookup tables and samples from larger distributions are drawn using the Box-Muller method.

5.3. Temporal integration

The temporal resolution at scotopic conditions is low [Wan95, Fig. 7.23] and even lower close to absolute threshold [USB08, Fig. 1] (unfortunately, in the range 10^{-6}–10^{-5} cd/m^2 we were not able to find the relevant data for human vision). Consequently, noise undergoes filtering of temporal frequencies above 10 Hz. A simple solution would store the last 100 ms and average them. Instead, we use a closed-form solution routinely applied in graphics [KP11]: To simulate the current frame, the old frame is blended with the new one weighted by $\alpha = \exp(-2\pi \cdot f_c/f_s)$, and $1-\alpha$ [Smi97], where $f_c = 0.5$ Hz is the cutoff frequency and f_s is the frame rate. The cutoff of $f_c = 0.5$ was tuned by manual inspection to achieve a result most similar to averaging multiple frames.

As the noise process occurs on the retina, the noise pattern is expected to move with the eye over the image. In the absence of eye tracking, we make the assumption that the eye follows the optical flow for given pixel [KP11] during the integration-time period. We warp the noise pixel-by-pixel along this flow, including a depth test if depth is available.

Pixel-by-pixel warping is used, as repeated warping of an image at a high framer-rate would quickly blur out all high spatial frequencies that are important for our noise. However, this warping, as well as disocclusion (when depth is used), results in holes.

To address this issue, we *jitter* the target position of every pixel by a small random offset of one pixel. Doing so reduces the regular structure, which would become apparent in a smooth flow field. Further, we *fill* the remaining holes with new noise values, but, as they did not undergo temporal integration, a careful choice is needed, otherwise, their brightness statistics would differ from the warped pixels. One approach would be to draw multiple samples over time and average them over the integration period. A more efficient solution is to directly change the distribution from which these hole-filling samples are drawn to match the mean and standard deviation of the temporally-integrated distribution. Such a distribution can be obtained by properly scaling the standard simulation time of 100 ms and the number of photons. Intuitively, larger time leads to higher mean values in the simulation and therefore lower relative noise. As our integration procedure is an exponential smoothing filter and our cumulative distribution behaves like a box filter, we can find a proper scaling for the simulation time by looking for a box filter length such that the corresponding exponential and uniform distribution mean values and standard deviations are equal. The resulting scaling factor is $(1+\alpha)/(1-\alpha)$.

6. Results

Acquiring a reference noise for comparison is impossible; it exists solely as a neural representation, for which no imaging technology is available. This section complements the quantitative fit to physiological data, which we have provided so far, with performance evaluations, a qualitative assessment in form of actual images, and a perceptual experiment.

Performance Our implementation computes an HD noise frame from an input image in 18 ms on a Nvidia Geforce 660 GTX. Most time is spent on the image-warping with respect to the estimated eye movement (9 ms). Producing samples from the distribution is fast when using the lookup tables (1.8 ms).

Images Fig. 1, Fig. 5 and Fig. 3 show typical results, compared to other noise models at different scales of absolute luminance. Gaussian noise follows the recommendation by Thompson et al. [TSF02]; adding Gaussian noise with a small standard deviation, independent of the luminance in the image. We see that the noise does not adapt to the image content and it is unclear how it should be scaled in respect to adaptation. The quantum and dark noise show implementations of Eq. 1 and Eq. 2 respectively. The quantum noise reacts to image content but lacks noise in dark areas. The dark noise has an inverted behavior. Only our model combines all properties into a consistent omnipresent noise that reacts to luminance. Note, that Fig. 1 and Fig. 3 span a range of adaptation levels for didactic purposes, while images only have one dominant adaptation level in practice. The supplemental video shows animated versions of these two figures.

Perceptual experiment The key questions using our noise model concern the nocturnal mood impression (realism), viewing comfort, and the observer's overall preference. To this extent, we performed a perceptual experiment, which

Figure 3: *Gaussian [TSF02], quantum (Eq. 1), dark noise (Eq. 2) and our model* (top to bottom)*, applied to an image that contains different absolute scales of luminance* (horizontal)*, including a condition close to absolute threshold. Insets show the power distribution of the noise* (black) *and a Gaussian reference* (grey) *at four specific pixels together with their photon count N.*

was precedented by a calibration phase where dark-adapted subjects could possibly experience scotopic noise themselves. Afterwards, subjects were shown videos with day-for-night tone mapping applied, with either (1) white additive noise as suggested by Thompson et al. [TSF02] or (2) our full model of noise distribution temporarily simulated either as (1) a static noise frame (a single white noise pattern or our full model with a constant random seed), (2) a dynamically changing phenomena. Subjects were shown pairs of videos and asked, which one: (1) depicts the scene more realistically? (2) is more comfortable to look at? (3) is preferred in general?

To explore the dynamic aspects of noise and its interaction with the image content the video sequences exhibit camera motion. In total four different short movies (10 s) were used; two computer-generated (ARCHITECTURE, TUNNEL) and two captured (COUNTRYSIDE, CAR). The animation in CAR was produced by horizontally panning across the teaser in Thompson et al. [TSF02]. Three of the videos (ARCHITECTURE, COUNTRYSIDE, TUNNEL) contain temporal changes of absolute luminance (see supplemental video).

Ten subjects took part in the experiment comprising a *calibration* and a *query* phase. In the calibration phase, they were instructed to adapt for 10 minutes in a dark room. It

was done merely to let them investigate the appearance of several presented objects and at the same time experience the scotopic noise for themselves. While more time is typically required to achieve full scotopic adaptation [FPSG96,PTYG00], we found scotopic noise to become apparent already after ten minutes in our setting.Longer adaptation is expected to produce an even stronger effect, but results in fatigue.The query phase was performed in photopic conditions under controlled dim office illumination. Subjects were shown all of the $4\times2\times2$ combinations of the above stimuli and noise variants in a random order. Stimuli involving changes in distribution and temporal behavior simultaneously were skipped to reduce fatigue. Videos were looped and after three repetitions (30 s), subjects were asked to answer the three questions stated above. The used display was a Dell U2412M with a resolution of 1920×1200 pixels and stimuli were played at 30 Hz. They were shown next to each other at a resolution of 800×800 pixels in front of a dark background, at a distance of 60 cm, at which a pixel covers a visual angle of 1.45 arcmin. Subjects were adapted to dim, photopic office-lighting conditions.

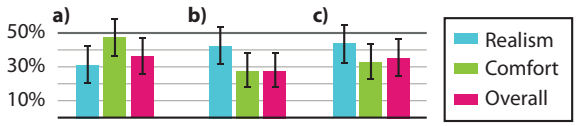

Figure 4: *Study statistics for three different comparisons* a–c) *and three qualities* (colors)*. Each bar's height denotes preference compared to a reference. Notches denote 95 % confidence intervals (CI). Comparison with a CI not intersecting the 50 %-null hypothesis line are statistically significant.*

First, we compare our full approach to a variant using static instead of dynamic noise (Fig. 4, a). It is significantly preferred (all significant effects reported are, $p < .05$ binomial test) overall ($63.7\,\%, \mathrm{CI}\,[11.5, 10.5]\,\%$) and, in particular, in terms of realism ($68.8\,\%, \mathrm{CI}\,[11.3, 9.9]\,\%$), while no significant effect on comfort was present ($52.5\,\%, \mathrm{CI}\,[11.5, 11.3]\,\%$). The finding indicates that adding dynamic noise is useful because static noise is perceived unnatural, in particular for dynamic scene.

Second, we compare our full approach to dynamic white noise (Fig. 4, b). Again it is found to be significantly better overall ($72.5\,\%, \mathrm{CI}\,[11.1, 9.4]\,\%$) and in terms of comfort ($72.5\,\%, \mathrm{CI}\,[11.1, 9.4]\,\%$), while the improvement in realism is not significant ($57.5\,\%, \mathrm{CI}\,[11.6, 11.0]\,\%$). The finding indicates that adding dynamics alone is not preferred over adding it to the appropriate noise distribution. Probably, uncorrelated white noise fluctuations are perceived unnatural as they not adjust to the image content.

The comparison of a static variant of our noise and static white noise (Fig. 4, c), leads to a significant preference of the static variant of our approach in terms of comfort ($67.5\,\%, \mathrm{CI}\,[11.4, 10.1]\,\%$) and preference ($65.0\,\%, \mathrm{CI}\,[11.5, 10.3]\,\%$) with no significant effect on realism ($56.3\,\%, \mathrm{CI}\,[11.6, 11.1]\,\%$). The finding indicates that besides all dynamics, our choice of physiologically-principled noise is an important factor; only a mix of additive and multiplicative noise, as well as adaption to the actual luminance seems to appear plausible.

In summary, the experiments indicate, that previous work adds to the nocturne mood in static images, but might be incomplete for animated imagery. Further, the noise dynamics as a function of scene content is not trivial and the type of noise distribution leads to perceivable differences. Still, extensive noise can reduce viewing comfort and, ultimately, if an artist decides to use noise to depict scotopic conditions, a tradeoff is possible.

7. Conclusion

We derived a physiologically-motivated model of noise perception close to the absolute luminance threshold. The model is practical and can be computed efficiently. Our specialized warping maintains noise details and leads to temporal coherence and might be useful in other contexts or other forms of temporally-coherent high-frequency noise. The experimental evaluation shows that our simulated noise is always overall preferred, and more comfortable to watch than previous solutions, which are based on white noise. Our dynamic-noise solution can be potentially less comfortable than its static counterpart, but it consistently improves realism. The artistic intent should be the key factor when choosing between apparent realism and viewing comfort. Our model does not yet account for any higher-level effects. We assume that the day-for-night tone mapping (chroma change, acuity loss) is independent of and happens before the receptor noise. While this is physiologically plausible, future work could account for higher-level processes in order to reproduce all major scotopic phenomena, including also the Purkinje shift, and the scotopic (temporal) contrast, as a consequence of the noisy retinal signal processing itself.

References

[AF77] ASHMORE F., FALK G.: Dark noise in retinal bipolar cells and stability of rhodopsin in rods. *Nature 270* (1977), 69–71. 4

[Alp71] ALPERN M.: Rhodopsin kinetics in the human eye. *J Phys. 217*, 2 (1971), 447–471. 2

[Alt77] ALTMAN J.: The sensitometry of black and white materials. In *The Theory of the Photographic Process* (1977). 3

[Bar56] BARLOW H. B.: Retinal noise and absolute threshold. *J. Opt. Soc. Am. 46*, 8 (1956), 634–639. 2, 4

[Blo85] BLOCH A. M.: Experience sur la vision. *C.r. Séanc. Soc. Biol. 37* (1885), 493–495. 4

[BLY79] BAYLOR D., LAMB T., YAU K.-W.: Responses of retinal rods to single photons. *J Phys. 288*, 1 (1979), 613–34. 4

[DD00] DURAND F., DORSEY J.: Interactive tone mapping. In *Proc. EGWR* (2000), pp. 219–230. 2, 3

[Dee05] DEERING M.: A photon accurate model of the human eye. *ACM Trans. Graph. (Proc. SIGGRAPH) 24*, 3 (2005), 649–58. 3, 4

[FJ05] FAIRCHILD M. D., JOHNSON G. M.: On the salience of novel stimuli: Adaptation and image noise. In *IS&T/SID 13th Color Imaging Conference* (2005), pp. 333–338. 3

[FPSG96] FERWERDA J. A., PATTANAIK S., SHIRLEY P., GREENBERG D.: A model of visual adaptation for realistic image synthesis. In *Proc. SIGGRAPH* (1996), pp. 249–58. 2, 8

[FSR05] FIELD G. D., SAMPATH A. P., RIEKE F.: Retinal processing near absolute threshold: from behavior to mechanism. *Annu. Rev. Physiol. 67* (2005), 491–514. 4

[GAMS05] GUTIERREZ D., ANSON O., MUNOZ A., SERON F.: Perception-based rendering: eyes wide bleached. In *EG Short Paper* (2005), pp. 49–52. 3, 4

[GLC13] GOMILA C., LLACH J., COOPER J.: Film grain simulation method, 2013. US Patent 8,447,127. 3

[GM97] GEIGEL J., MUSGRAVE F. K.: A model for simulating the photographic development process on digital images. In *Proc. SIGGRAPH* (1997), pp. 135–142. 3

[HSP42] HECHT S., SHLAER S., PIRENNE M. H.: Energy, quanta, and vision. *J Gen Phys 25*, 6 (1942), 819–840. 2, 3, 4, 5

[Jan01] JANESICK J. R.: *Scientific charge-coupled devices*, vol. 83. SPIE press, 2001. 3

[JF00] JOHNSON G. M., FAIRCHILD M. D.: Sharpness rules. In *IS&T/SID 8th Color Imaging Conference* (2000), pp. 24–30. 3

[JSN92] JONAS J. B., SCHNEIDER U., NAUMANN G. O.: Count and density of human retinal photoreceptors. *Graefe's archive for clinical and experimental ophthalmol. 230*, 6 (1992), 505–10. 4, 5

[KMAK08] KURIHARA T., MANABE Y., AOKI N., KOBAYASHI H.: Digital image improvement by adding noise: An example by a professional photographer. In *Image Quality and System Performance V* (2008), vol. 6808 of *SPIE*, pp. 1–10. 3

[KO11] KIRK A. G., O'BRIEN J. F.: Perceptually based tone mapping for low-light conditions. *ACM Trans. Graph. (Proc. SIGGRAPH) 30*, 4 (2011), 42:1–10. 2

[KP04] KHAN S. M., PATTANAIK S. N.: Modeling blue shift in moonlit scenes by rod cone interaction. *J Vis. 4*, 8 (2004). 2

[KP11] KASS M., PESARE D.: Coherent noise for non-photorealistic rendering. *ACM Trans. Graph. (Proc. SIGGRAPH) 30*, 4 (2011), 30. 3, 6

[KRV*14] KELLNHOFER P., RITSCHEL T., VANGORP P., MYSZKOWSKI K., SEIDEL H.-P.: Stereo day-for-night: Retargeting disparity for scotopic vision. *ACM Trans. Appl. Percept. 11*, 3 (2014), 15:1–17. 3

[KWK09] KIM M. H., WEYRICH T., KAUTZ J.: Modeling human color perception under extended luminance levels. *ACM Trans. Graph. (Proc. SIGGRAPH 2009) 28*, 3 (2009), 27:1–9. 2

[Lil81] LILLYWHITE P.: Multiplicative intrinsic noise and the limits to visual performance. *Vis. Res. 21*, 2 (1981), 291–296. 4, 5

[Liv02] LIVINGSTONE M.: *Vision and art: the biology of seeing*. Harry N. Abrams, 2002. 3

[LLC*10] LAGAE A., LEFEBVRE S., COOK R., DEROSE T., DRETTAKIS G., EBERT D. S., LEWIS J. P., PERLIN K., ZWICKER M.: State of the art in procedural noise functions. In *EG 2010 - State of the Art Reports* (2010). 3

[MKN84] MARIANI A. P., KOLB H., NELSON R.: Dopamine-containing amacrine cells of Rhesus monkey retina parallel rods in spatial distribution. *Brain Res. 322*, 1 (1984), 1 – 7. 5

[Pal99] PALMER S. E.: *Vision science: Photons to phenomenology*, vol. 1. MIT press Cambridge, 1999. 2

[PFFG98] PATTANAIK S. N., FERWERDA J. A., FAIRCHILD M. D., GREENBERG D. P.: A multiscale model of adaptation and spatial vision for realistic image display. In *Proc. SIGGRAPH* (1998), pp. 287–98. 2

[PTYG00] PATTANAIK S. N., TUMBLIN J. E., YEE H., GREENBERG D. P.: Time-dependent visual adaptation for fast realistic image display. In *Proc. SIGGRAPH* (2000), pp. 47–54. 2, 3, 8

[RE12] RITSCHEL T., EISEMANN E.: A computational model of afterimages. *Comp. Graph. Forum (Proc. EG) 31*, 2 (2012), 529–534. 2

[RP80] RIVA C., PETRIG B.: Blue field entoptic phenomenon and blood velocity in the retinal capillaries. *J. Opt. Soc. Am. 70*, 10 (1980), 1234–1238. 2

[RWD*10] REINHARD E., WARD G., DEBEVEC P., PATTANAIK S., HEIDRICH W., MYSZKOWSKI K.: *High Dynamic Range Imaging*. Morgan Kaufmann Publishers, 2nd edition, 2010. 2, 4

[Sey11] SEYMOUR M.: Case study: How to make a Captain America wimp. *fxguide* (2011). 3

[Shl37] SHLAER S.: The relation between visual acuity and illumination. *J Gen Phys 21* (1937), 165–188. 2

[Smi97] SMITH S. W.: *The scientist and engineer's guide to digital signal processing*. California Technical Pub., 1997. 6

[SS07] STEPHENSON I., SAUNDERS A.: Simulating film grain using the noise-power spectrum. In *Theory and Practice of Computer Graphics* (2007), pp. 69–72. 3

[TDMS14] TEMPLIN K., DIDYK P., MYSZKOWSKI K., SEIDEL H.-P.: Perceptually-motivated stereoscopic film grain. *Comp. Graph. Forum (Proc. Pacific Graphics) 33*, 7 (2014). 3

[TPV*82] TEICH M., PRUCNAL P. R., VANNUCCI G., BRETON M. E., MCGILL W. J.: Multiplication noise in the human visual system at threshold. *J. Opt. Soc. Am. 72*, 4 (1982), 419–31. 4

[TSF02] THOMPSON W. B., SHIRLEY P., FERWERDA J. A.: A spatial post-processing algorithm for images of night scenes. *J. Graph. Tools 7*, 1 (2002), 1–12. 2, 3, 4, 6, 7, 10

[USB08] UMINO Y., SOLESSIO E., BARLOW R. B.: Speed, spatial, and temporal tuning of rod and cone vision in mouse. *J Neur 28*, 1 (2008), 189–98. 6

[Wal45] WALD G.: Human vision and the spectrum. *Science 101* (1945), 653–58. 2

[Wan95] WANDELL B. A.: *Foundations of Vision*. Sinauer Associates, 1995. 5, 6

[WM14] WANAT R., MANTIUK R.: Simulating and compensating changes in appearance between day and night vision. *ACM Trans. Graph. (Proc. SIGGRAPH) 33*, 4 (2014). 2

[WY12] WATSON A. B., YELLOTT J. I.: A unified formula for light-adapted pupil size. *J Vis 12* (2012). 4

[YMB79] YAU K., MATTHEWS G., BAYLOR D.: Thermal activation of the visual transduction mechanism in retinal rods. *Nature 279* (1979), 806–7. 4

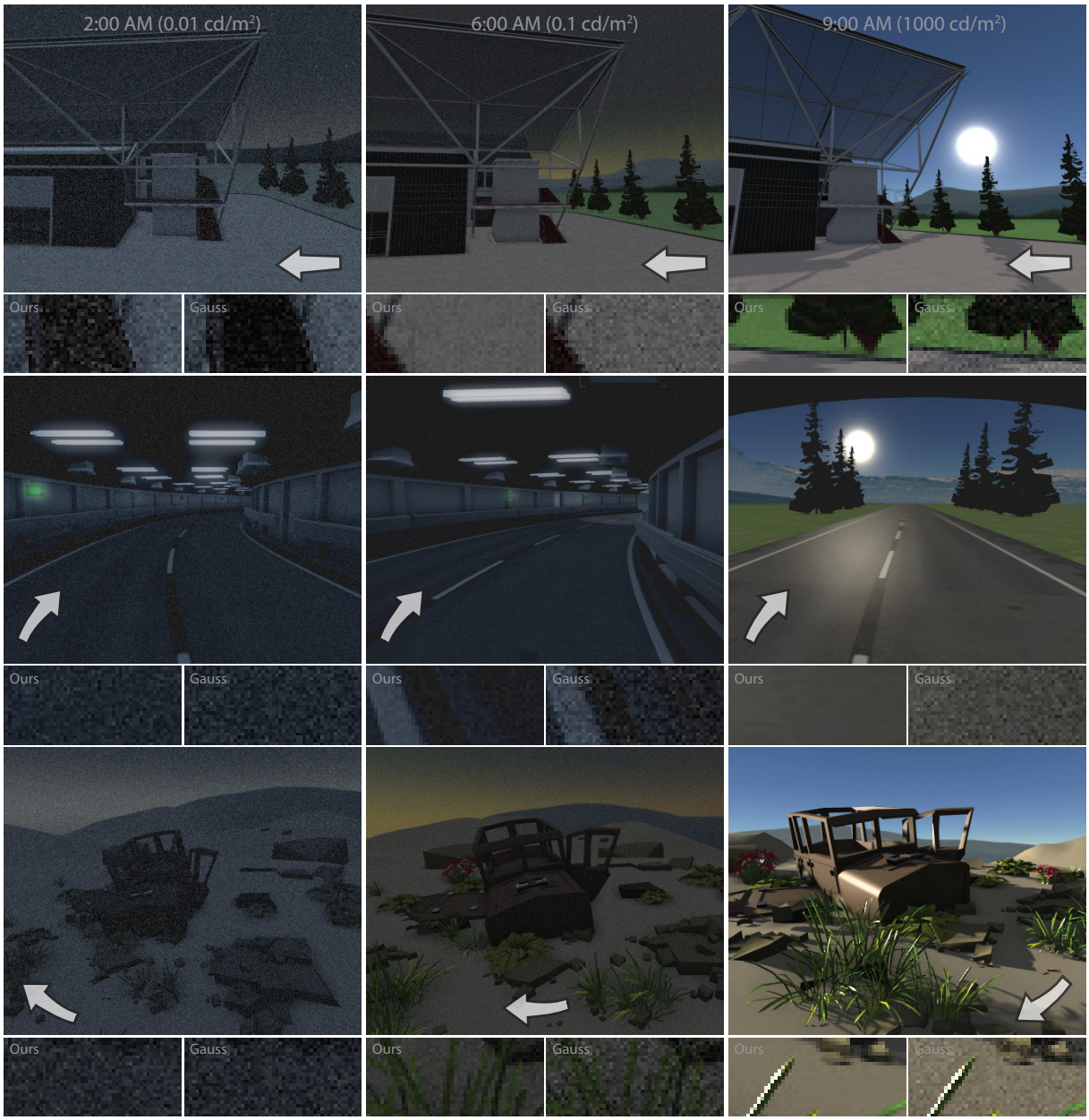

Figure 5: *Results of our full model applied to different CG animations* (Rows) *with adaptation luminance changing over time* (Columns). *The first row is a time-lapse architectural visualization from night over morning to daylight. The second row is a driving simulation in a tunnel. The last row shows a time-lapse animation of a three-dimensional animated 3D scene with a setting similar to the photography used by Thompson et al. [TSF02]. The pairs of insets (the left is ours, the right is Gaussian noise) show a magnified part of the frames above. Our approach changes over time, adapts to luminance changes, and does not suffer from the shower-door effect, which is typical of screen-space patterns, such as Gaussian noise, which only works well for a photo with a fixed luminance level.*