Survey and Evaluation of Tone Mapping Operators for HDR Video

Supplementary Material

Gabriel Eilertsen[1]* Jonas Unger[1] Robert Wanat[2] Rafał Mantiuk[2]

[1]Linköping University, Sweden [2]Bangor University, United Kingdom

Tone mapping operators

The tone mapping operators (TMOs) used in the evaluation are listed, and shortly described, in Table 1. All operators in the evaluation include a model for processing of temporal variations in the HDR input data. Recognizing that different TMOs may have different aims, we organize them into two different classes based on their intent; **Visual System Simulators (VSS)**, which simulates the properties and limitations of the human visual system, and **Best Subjective Quality (BSQ)** operators which try to achieve the subjectively preferred image. The table also divides the TMOs into *global* and *local* operators, where the global performs the same calculations for all pixels in the image, while the local differs from pixel to pixel, based on *e.g.* processing of neighboring pixels.

HDR video sequences

In the evaluation five different HDR video sequences were used, containing a range of different phenomena:

- **Park**: Walk-through scene in a park, with the sun flickering through the trees, creating rapid temporal changes in illumination.
- **Window**: Person moving in front of a window, occluding the sun and creating a large temporal change in illumination when stepping to the side.
- **Corridor**: People walking towards camera in a corridor, generating high contrast edges moving in the image. The sequence also contains a significant amount of noise in the darker parts of the image.
- **Entrance**: Walk-through in a corridor from light to dark environment, making a slow transition in illuminance.
- **Lab**: Static lab scene with a light bulb switched on and off, isolating only a rapid temporal change in illumination.

Example frames from the sequences, together with tone mappings, are shown in Figure 1.

Light adaptation models

In common for all the video tone mapping operators in our study is that they include some model for changes in input luminance over time. Figure 2 and 3 visualize the temporal response of the different TMOs for the same input data. The plots show how the relative output luminance for each operator varies over time in the lab scene. The luminances are averaged over small areas for two different locations in the image; one far away from the light bulb, where the change in intensity when switching it on and off is low, and one close to the light bulb where the dynamic range over time is large. The luminances have, for presentation purposes, been shifted along the y-axis so that the mean value within the considered region is the same for input and tone mapping when the light is switched off. In this way the differences in how the operators adapt to a change in illumination is easy to see.

*e-mail:gabriel.eilertsen@liu.se

Operator	Publication	Processing	Intent	Description
Model of visual adaptation	[4]	Global	VSS	Use of data from psychophysical experiments to simulate adaptation over time, and effects such as color appearance and visual acuity. Visual response model is based on measurements of threshold visibility as in [11].
Time-dependent visual adaptation	[9]	Global	VSS	Based on published psychophysical measurements [5]. Static responses are modeled separately for cones and rods, and complemented with exponential smoothing filters to simulate adaptation in the temporal domain. A simple appearance model is also included.
Local model of eye adaptation	[6]	Local	VSS	Temporal adaptation model operating on a local level using bilateral filtering, where experimental data is used to simulate the adaptation.
Per-pixel virtual exposures	[1]	Local	BSQ	Bilateral filter applied both spatially for local processing, and separately in time domain for temporal coherence.
Model of human cones	[10]	Global	VSS	Dynamic system modeling the cones in the human visual system over time. A quantitative model of primate cones is utilized, based on actual retina measurements.
Gradient domain tone mapping	[7]	Local	BSQ	Extension for temporal coherence of the gradient domain TMO in [3]. Motion vectors are included when solving the Poisson equation on the gradient field, to ensure temporal stability in the output.
Display adaptive tone mapping	[8]	Global	BSQ	Display adaptive tone mapping, where the goal is to preserve the contrasts within the input (HDR) as close as possible given the characteristic of an output display. Temporal variations are handled through a filtering procedure.
Retina model	[2]	Local	VSS	Biological retina model where the time domain is used in a spatio-temporal filtering for local adaptation levels. The spatio-temporal filtering, simulating the cellular interactions, yields an output with whitened spectra and temporally smoothed for improved temporal stability and for noise reduction.

Table 1: The list of reviewed video tone mapping operators. Processing refers to either global processing that is identical for all the pixels within a frame or local processing that may vary spatially. Intent refers to the main goal of the operator, and corresponds to our classification.

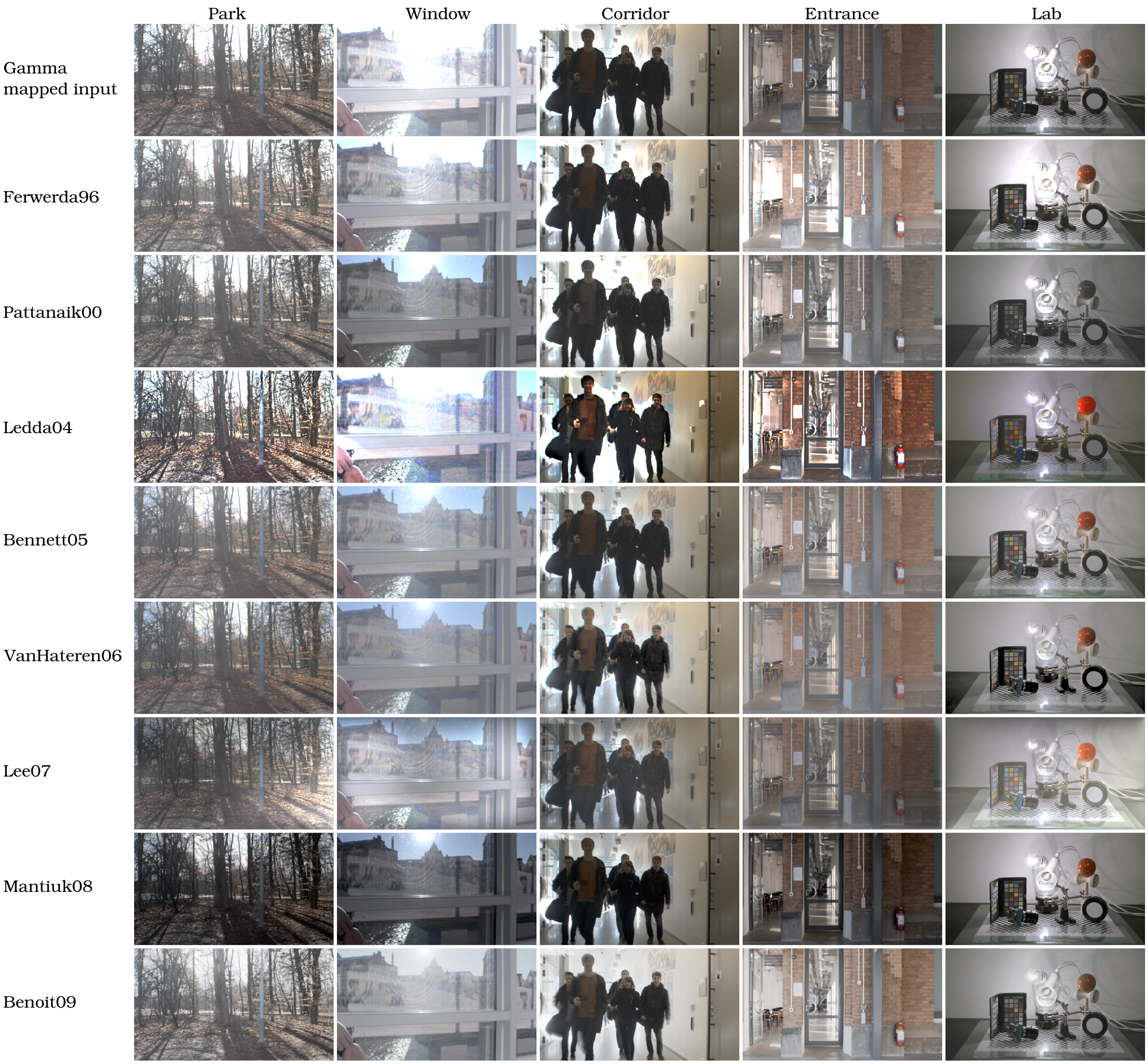

Figure 1: Example frames from the different HDR videos, and TMOs applied. The top row shows the input sequences with a gamma mapping ($\gamma = 2.2$).

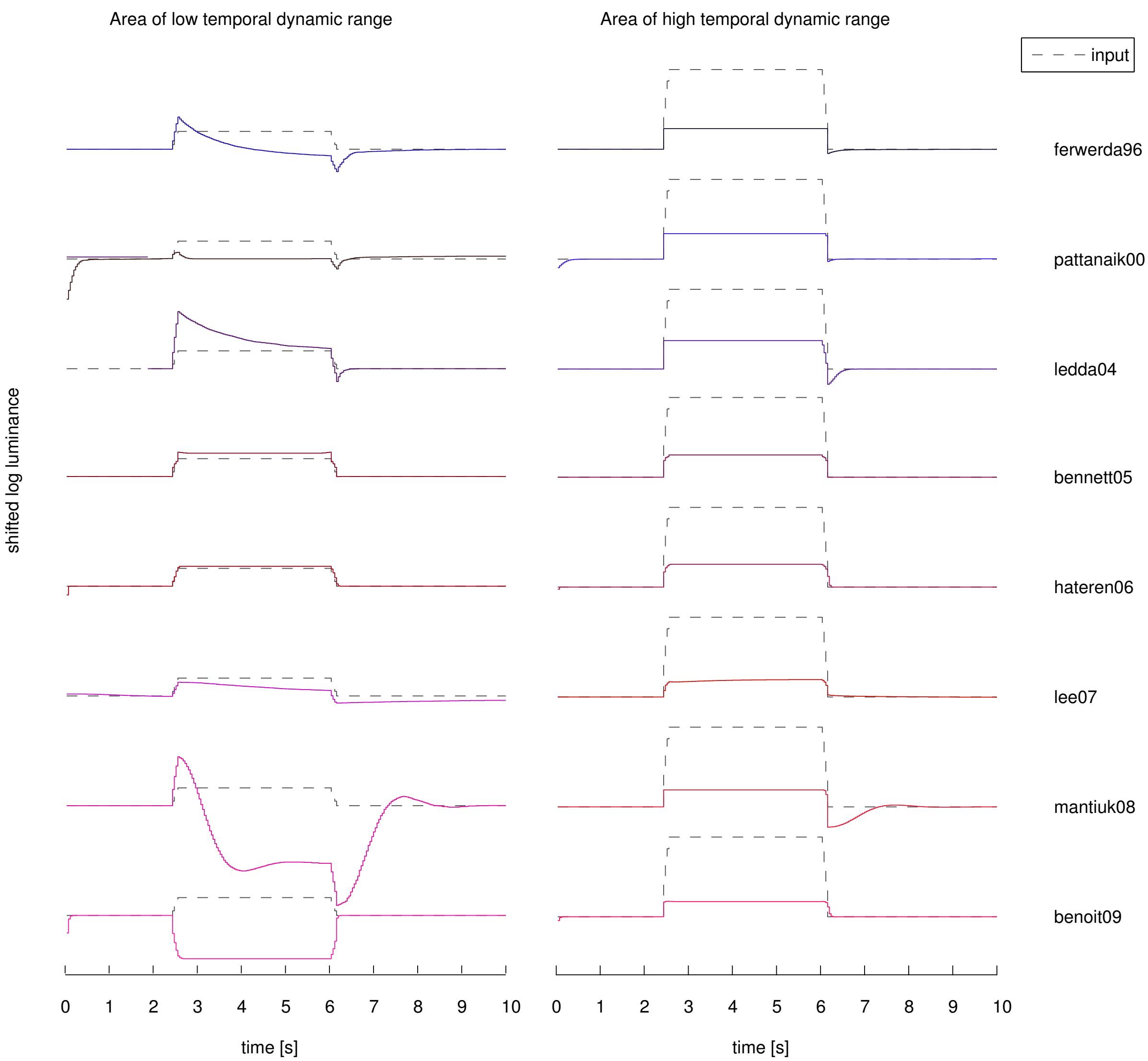

Figure 2: Visualization of the light adaptation process over time for the TMOs used in the evaluation. The left column shows how the luminance in a region of low temporal dynamic range changes over time, while the right is for a region of high temporal dynamic range. The low range region is represented by the background wall at distance from the light in the lab scene, and the high range region by the lamp base right next to the light bulb (see Figure 3). The curves have been shifted along the y-axis so that the luminance when the light is switched off is the same for input and tone mappings, to simplify comparisons.

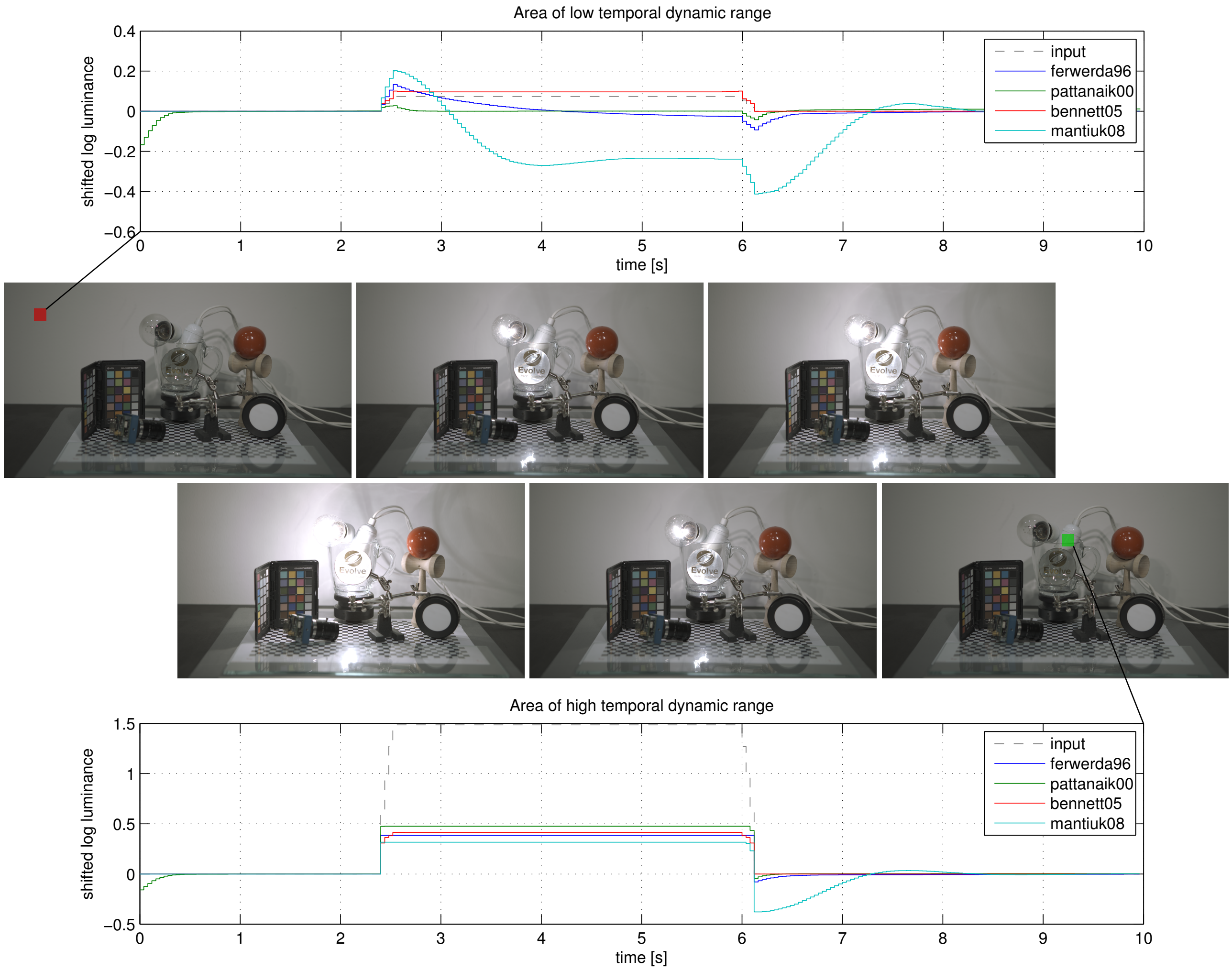

Figure 3: Visualization of the light adaptation process over time for the TMOs used in the evaluation. The upper figure shows how the luminance in a region of low temporal dynamic range changes over time, while the bottom figure is for a region of high temporal dynamic range. The low range region is represented by the background wall at distance from the light in the lab scene, and the high range region by the lamp base right next to the light bulb. The curves have been shifted along the y-axis so that the luminance when the light is switched off is the same for input and tone mappings, to simplify comparisons.

Parameter settings

Parameters available for tuning the tone mappings are shown in Table 2, where throughout the table $L = L(x, y)$ is used to refer to the luminance of the image. The table also shows the default values where available, and the settings used for the different sequences in the evaluation.

Operator	Param.	Description	Default	Park	Wind.	Corr.	Entr.	Lab
Model of visual adapt.	$L_{d,max}$	Maximum screen luminance.	100	100	1000	1000	1000	100
	f	Frame rate.	25	25	25	25	25	25
Time-dependent visual adapt.	G_{cone}	Cones adaptation level.	1° foveal weight [12]	$geo-mean(L)$	$gm(L)$	$gm(L)$	$gm(L)$	$gm(L)$
	G_{rod}	Rods adaptation level.			$gm(L)$	$gm(L)$	$gm(L)$	$gm(L)$
	f	Frame rate.	-	25	25	25	25	25
Local model of eye adaptation	α	Power constant used in deriving half saturation parameter for local sigmoid mapping from the adaptation level.	0.69	0.69	0.69	0.69	0.69	0.69
	β_{cone}	Multiplicative constant used in deriving half saturation parameter for cone response.	2.0	2.0	2.0	2.0	2.0	2.0
	β_{rod}	Multiplicative constant used in deriving half saturation parameter for rod response.	5.83	5.83	5.83	5.83	5.83	5.83
	n	Power constant for local sigmoid mapping.	0.7-2.0	3	3	2.8	3	2.2
	σ_f	Standard deviation of Gaussian filter for spatial component of bilateral filter.	-	0.2	0.4	0.4	0.2	0.4
	g	Standard deviation scale of Gaussian filter for intensity component of bilateral filter.	0.15	0.15	0.15	0.15	0.15	0.15
	σ_{disp}	Display adaptation level.	30-50	30	30	30	30	30
	τ_{cone}	Cone dark adaptation time constant.	600	5	5	5	5	5
	τ_{rod}	Rod dark adaptation time constant.	2400	5	5	5	5	5
	ρ_{cone}	Cone light adaptation time constant.	20	5	5	5	5	5
	ρ_{rod}	Rod light adaptation time constant.	20	5	5	5	5	5
	f	Frame rate.	-	25	25	25	25	25
Per-pixel virtual exposures	ψ_1	Base layer attenuation profile parameter.	40	10	10	10	10	50
	ψ_2	Details layer attenuation profile parameter.	700	700	700	700	700	700
	σ_s	Spatial standard deviation of bilateral filter.	-	2	2	2	2	2
	σ_r	Intensity domain standard deviation of bilateral filter.	-	0.05	0.05	0.05	0.05	0.1
	σ_t	Standard deviation of temporal bilateral filter.	-	3	3	3	3	3
	σ_{tr}	Intensity domain standard deviation of temporal bilateral filter.	-	0.05	0.05	0.05	0.05	0.1
	σ_c	Standard deviation of Gaussian blur of chroma.	-	0.1	0.1	0.1	0.1	0.1
Model of human cones	p_a	Pupil area.	10	10	10	10	10	10
	k_β	Gain.	$1.6 \cdot 10^{-4}$	$1.6 \cdot 10^{-4}$	$1.6 \cdot 10^{-4}$	$1.6 \cdot 10^{-4}$	$1.6 \cdot 10^{-4}$	$1.6 \cdot 10^{-4}$
	a_C	Calcium feedback constant.	$9 \cdot 10^{-2}$	$9 \cdot 10^{-2}$	$9 \cdot 10^{-2}$	$9 \cdot 10^{-2}$	$9 \cdot 10^{-2}$	$9 \cdot 10^{-2}$
	c_β	Residual activity.	$2.8 \cdot 10^{-3}$	$2.8 \cdot 10^{-3}$	$2.8 \cdot 10^{-3}$	$2.8 \cdot 10^{-3}$	$2.8 \cdot 10^{-3}$	$2.8 \cdot 10^{-3}$

	τ_R	Temporal filter size, photo-transduction cascade.	$3.4 \cdot 10^{-3}$	$3.4 \cdot 10^{-3}$	$3.4 \cdot 10^{-3}$	$3.4 \cdot 10^{-3}$	$3.4 \cdot 10^{-3}$	$3.4 \cdot 10^{-3}$
	τ_E	Temporal filter size, photo-transduction cascade.	$8.7 \cdot 10^{-3}$	$8.7 \cdot 10^{-3}$	$8.7 \cdot 10^{-3}$	$8.7 \cdot 10^{-3}$	$8.7 \cdot 10^{-3}$	$8.7 \cdot 10^{-3}$
	τ_C	Temporal filter size, calcium feedback loop.	$3.0 \cdot 10^{-3}$	$3.0 \cdot 10^{-3}$	$3.0 \cdot 10^{-3}$	$3.0 \cdot 10^{-3}$	$3.0 \cdot 10^{-3}$	$3.0 \cdot 10^{-3}$
	f	Frame rate.	25	25	25	25	25	25
Gradient domain tone mapping	α_{sc}	Gradient magnitude scaling level. Here determined as $\alpha \equiv \alpha_{sc}$ times the average gradient magnitude.	0.1	0.8	0.8	0.8	0.8	0.2
	β	Gradient compression exponent.	0.8-0.9	0.95	0.95	0.95	0.95	0.95
	λ	Temporal significance.	-	0.01	0.01	0.01	0.01	0.01
	s	Saturation exponent.	-	1	1	1	1	1
Display adaptive tone mapping	γ	Display gamma.	2.2	2.2	2.2	2.2	2.2	2.2
	L_{max}	Peak luminance of display.	100	100	100	100	100	100
	L_{black}	Black level; luminance of a black display pixel.	1	1	1	1	1	1
	k	Reflectivity of display screen.	0.01	0.01	0.01	0.01	0.01	0.01
	E_{amb}	Ambient illumination where image is to be displayed, in lux.	50	50	50	50	50	50
	c	Color saturation factor.	1	1	1	1	1	1
	e	Contrast enhancement.	1	1	1	1	1	1
Retina model	h	Histogram clipping value.	0	0	0	0	0	0
	s	Color saturation factor.	3	3	3	3	3	3
	β_h	Retina horizontal cells gain.	40	40	40	40	40	40
	$R_{p,0}$	Local adaptation photo-receptors.	197	193	193	193	193	182
	$R_{g,0}$	Local adaptation ganglion cells.	190	186	186	186	186	175

Table 2: Operator parameters, and settings used for the different sequences.

References

[1] Eric P. Bennett and Leonard McMillan. Video enhancement using per-pixel virtual exposures. *ACM Trans. Graph.*, 24(3):845–852, July 2005.

[2] Alexandre Benoit, David Alleysson, Jeanny Herault, and Patrick Callet. *Spatio-temporal Tone Mapping Operator Based on a Retina Model*, pages 12–22. Springer-Verlag, Berlin, Heidelberg, 2009.

[3] Raanan Fattal, Dani Lischinski, and Michael Werman. Gradient domain high dynamic range compression. In *Proceedings of the 29th annual conference on Computer graphics and interactive techniques*, SIGGRAPH '02, pages 249–256, New York, NY, USA, 2002. ACM.

[4] James A. Ferwerda, Sumanta N. Pattanaik, Peter Shirley, and Donald P. Greenberg. A model of visual adaptation for realistic image synthesis. In *Proc. of SIGGRAPH*, pages 249–258, New York, NY, USA, 1996. ACM.

[5] R. W. G. Hunt. *The Reproduction of Colour.* Fountain Press, 1995.

[6] Patrick Ledda, Luis Paulo Santos, and Alan Chalmers. A local model of eye adaptation for high dynamic range images. In *Proceedings of the 3rd international conference on Computer graphics, virtual reality, visualisation and interaction in Africa*, AFRIGRAPH '04, pages 151–160, New York, NY, USA, 2004. ACM.

[7] Chul Lee and Chang-Su Kim. Gradient domain tone mapping of high dynamic range videos. In *ICIP (3)'07*, pages 461–464, 2007.

[8] RafałMantiuk, Scott Daly, and Louis Kerofsky. Display adaptive tone mapping. In *ACM SIGGRAPH 2008 papers*, SIGGRAPH '08, pages 68:1–68:10, New York, NY, USA, 2008. ACM.

[9] Sumanta N. Pattanaik, Jack Tumblin, Hector Yee, and Donald P. Greenberg. Time-dependent visual adaptation for fast realistic image display. In *Proceedings of the 27th annual conference on Computer graphics and interactive techniques*, SIGGRAPH '00, pages 47–54, New York, NY, USA, 2000. ACM Press/Addison-Wesley Publishing Co.

[10] J. H. Van Hateren. Encoding of high dynamic range video with a model of human cones. *ACM Trans. Graph.*, 25:1380–1399, October 2006.

[11] Greg Ward. A contrast-based scalefactor for luminance display. In Paul S. Heckbert, editor, *Graphics gems IV*, pages 415–421. Academic Press Professional, Inc., San Diego, CA, USA, 1994.

[12] Gregory Ward Larson, Holly Rushmeier, and Christine Piatko. A visibility matching tone reproduction operator for high dynamic range scenes. In *ACM SIGGRAPH 97 Visual Proceedings: The art and interdisciplinary programs of SIGGRAPH '97*, SIGGRAPH '97, pages 155–, New York, NY, USA, 1997. ACM.