

# REAL-TIME NOISE-AWARE TONE-MAPPING AND ITS USE IN LUMINANCE RETARGETING

Gabriel Eilertsen\*<sup>§</sup>      Rafał K. Mantiuk<sup>†§</sup>      Jonas Unger\*<sup>§</sup>

\* Media and Information Technology, Linköping University, Sweden

† Computer Laboratory, University of Cambridge, UK

§ IRYSTEC, Canada

## ABSTRACT

With the aid of tone-mapping operators, high dynamic range images can be mapped for reproduction on standard displays. However, for large restrictions in terms of display dynamic range and peak luminance, limitations of the human visual system have significant impact on the visual appearance. In this paper, we use components from the *real-time noise-aware tone-mapping* to complement an existing method for perceptual matching of image appearance under different luminance levels. The refined luminance retargeting method improves subjective quality on a display with large limitations in dynamic range, as suggested by our subjective evaluation.

**Index Terms**— High dynamic range (HDR) video, tone-mapping, retargeting, low luminance display

## 1. INTRODUCTION

High dynamic range (HDR) video brings significant benefits in viewing experience as compared to standard images and video. Using tone-mapping operators (TMOs), HDR offers many possibilities for adapting the image content to mimic the experience of viewing the original scene, to achieve artistic looks, or to adapt to display and viewing environment. While TMOs can prepare an HDR input for a certain display, a related problem is how to match image appearance between widely different display setups, taking into account the properties and limitations of the human visual system (HVS).

In this paper we address the problem of compensating images for reproduction on displays exhibiting a low effective dynamic range. This could for example be due to low peak luminance for reducing power consumption, or when viewing under high ambient lighting. We first give an overview of the *real-time noise-aware tone-mapping* (RNTM) method [1], which provides a complete real-time tone-mapping pipeline that can accommodate many of the needs in display of HDR video. We then present an algorithm that uses the content-aware tone-curves from RNTM to complement the *luminance retargeting* (LRT) method from [2]. The result is a modified low luminance display compensation that uses the image content to make better use of the available dynamic range. To validate our algorithm, we perform a subjective evaluation, in

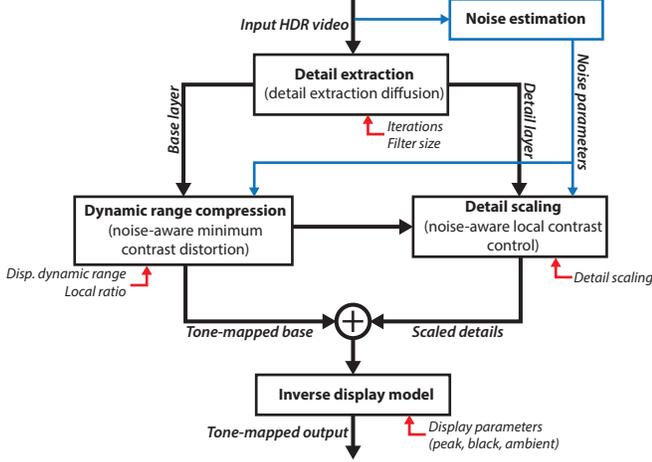
the context of a retargeting scenario for tablet displays. The results suggest that a better reproduction on a limited display is achieved using the modified LRT method proposed.

## 2. BACKGROUND

Images and video can come from a wide variety of sources, and need to be tone-mapped for display in different circumstances and for different purposes. With the development of HDR video [3, 4, 5], a number of TMOs have been proposed that take the temporal aspects into account. However, many are computationally expensive, lack in detail reproduction, or introduce artifacts such as amplification of noise, flickering and ghosting, see [6]. Recent TMOs, see *e.g.* [7, 8, 1], try to overcome these limitations. The RNTM algorithm, [1], which we consider in this paper, attempts to preserve details without visible artifacts and makes use of locally adaptive tone-curves that minimizes contrast distortions. Also, by taking the image noise into account, it can avoid increasing its visibility due to scaling of darker image areas, or due to detail enhancement.

While tone-mapping focuses on mapping linear, or scene-referred, units to a display-referred format, another important problem is to map between the display-referred formats. This could, for example, be in order to match the appearance when viewing an image on a display with low peak luminance. For this purpose, tone-mapping operators can be utilized after applying a forward display model. However, while RNTM takes into account the display and the viewing conditions, it does not account for the HVS. The LRT method in [2] uses a comprehensive model of the effects of the HVS over different luminances in order to match the visual appearance between displays, taking into account loss of acuity and the shifts in hue and color saturation. However, given a certain retargeting scenario, the tone-curve used for matching is static and does not adapt to the image content being displayed.

We propose to complement the static tone-curves in LRT with the content adaption from RNTM, creating a novel content-aware luminance retargeting algorithm which better distributes luminances over the available dynamic range of the target display. The aim is to achieve the subjectively most preferred retargeting on a very limited dynamic range, and not necessarily the closest perceptual match.



**Fig. 1.** Diagram of the RNTM pipeline. From the incoming HDR stream, detail and base layers are extracted. The base is compressed using local tone-curves, while details are scaled taking the noise characteristic into account.

### 3. CONTENT-AWARE LUMINANCE RETARGETING

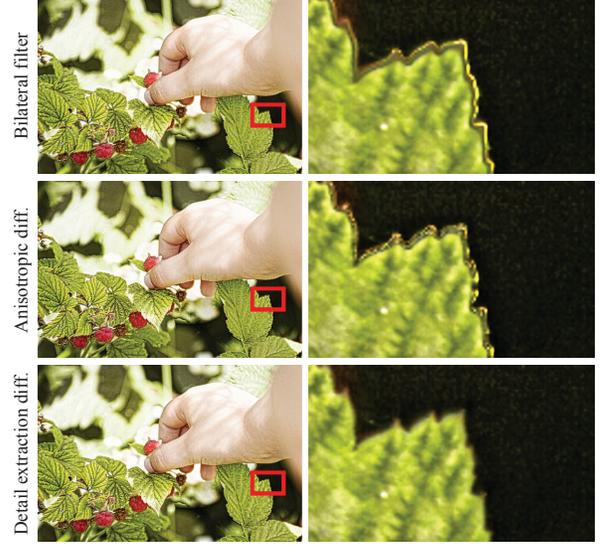
#### 3.1. Real-time noise-aware tone-mapping

The RNTM pipeline, Fig. 1, uses three central concepts based on novel techniques: 1) detail extraction with an edge-preserving filter especially designed for use in tone-mapping, 2) dynamic range compression by minimizing contrast distortions, and 3) noise-aware control over tone compression and details. All computations run in real-time, and are performed in the log domain to better account for the non-linear response of the HVS. For efficiency, only the luminance is used in the pipeline, and colors are added back at the end.

**Detail extraction:** In tone-mapping, the bilateral filter [9, 10] is often used for preserving details in the dynamic range compression. However, this and most other edge-preserving filters are not always well-suited for the purpose of tone-mapping – a problem which has not been fully recognized earlier. Comparing to *e.g.* noise reduction, a tone-mapping filter needs to work at a larger scale, and it is used to extract a detail layer. The detail layer easily reveals artifacts, such as gradient reversals near edges due to biased filtering.

Many of the 0-order filters suffer from reconstruction bias at smooth edges, resulting in banding artifacts in the detail layer (see Fig. 2). To overcome this problem an isotropic filter is used, which due to the uniform kernel cannot be biased. The filter is derived from a unified formulation, [11], of anisotropic diffusion [12] and bilateral filtering, and performs an iterative diffusion process with a gaussian kernel,  $g_{\sigma_j}$ :

$$I^{j+1}(\mathbf{p}) = (1 - w_r(\|\nabla I^j(\mathbf{p})\|)) I^j(\mathbf{p}) + w_r(\|\nabla I^j(\mathbf{p})\|) (g_{\sigma_k} * I^j)(\mathbf{p}) \quad (1)$$



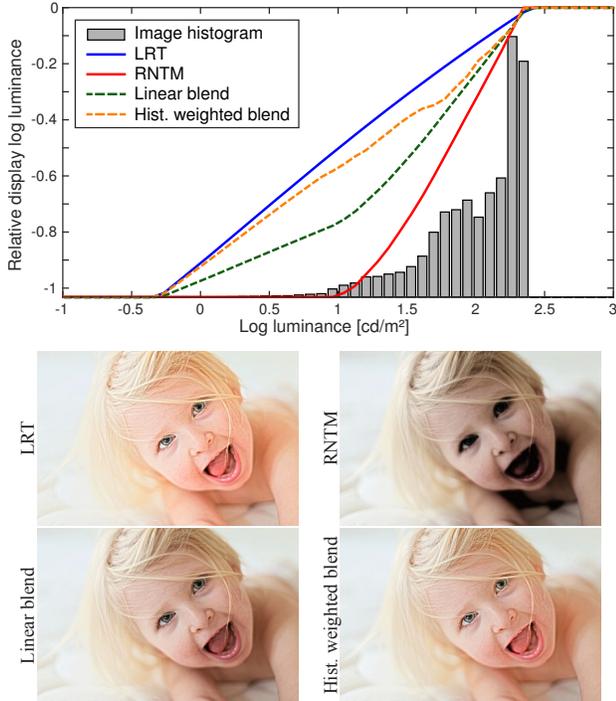
**Fig. 2.** Detail enhancement, using different edge-preserving filters. The image details have been extracted and scaled with a factor 5 before adding them back to the original image.

In each iteration  $j$ , the flow at point  $\mathbf{p}$  is weighted by the edge-stop function  $w_r$ , preventing filtering over large gradients  $\|\nabla I(\mathbf{p})\|$ . Instead of adapting the kernel to edges, this leads to less filtering (but avoids bias). In the context of tone-mapping, this is preferred over risking the artifacts associated with *e.g.* anisotropic diffusion and the bilateral filter, as shown in Fig. 2. Also, this behavior yields a simple implementation, enabling real-time processing of high resolution HDR video.

**Tone reproduction:** Having separated details from the HDR input, the base layer undergoes a compression of the dynamic range, mapping it to the luminance range of the display. Traditionally, this mapping, or tone-curve, is S-shaped to preserve contrast in middle tones at the cost of higher distortions at low and high tones. For input contrast  $G$  and the corresponding contrast  $\tilde{G}$  after tone-mapping, the expected value of these contrast distortions  $(G - \tilde{G})^2$  can be expressed using the distribution of contrast values  $p(G|l)$ , integrating over contrasts  $G$  and log-luminances  $l$ :

$$E[\|G - \tilde{G}\|_2^2] = \int p(l) \int (G - \tilde{G})^2 p(G|l) dG dl \quad (2)$$

Parameterizing the tone-curve as a piece-wise linear function with segments  $k$ , the expected distortions can be minimized as an  $L^2$  optimization problem, subject to that the tone-mapped luminances are within the dynamic range of the display and that the tone-curve is non-decreasing. The problem can be solved analytically for optimal tone-curve slopes  $s_k$ . The complete method is fast, and can be computed using the image histogram for estimating the distribution  $p(l_k)$ . An exam-



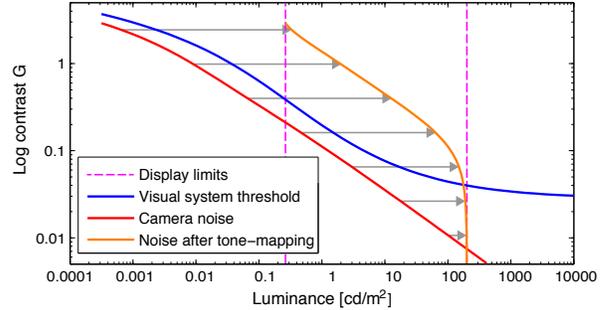
**Fig. 3. (Top):** comparing the tone-curves from LRT and RNTM, and different blendings of the two. **(Bottom):** re-targeting from 250 to 2.5  $cd/m^2$  and a dynamic range of 1.0 log units, using tone-curves in **(top)**.

ple of an RNTM tone-curve is shown in Fig. 3 (top), illustrating how the slope adapts to the image content.

In order to further maintain local image contrasts, the tone-curves are computed over image regions corresponding to 5 visual degrees. Per-pixel tone-curves are then derived from interpolation between the local tone-curves. Flickering due to rapidly changing image statistics between frames is prevented by filtering the nodes of the tone-curves over time. Filtering of the limited number of nodes is efficient compared to filtering individual pixels, and since the detail extraction filter utilized does not introduce visible inconsistencies over time, the complete tone-mapping chain is temporally stable.

**Noise-awareness:** In the tone-mapping process, contrast is manipulated and low luminance levels are in many cases increased. This may lead to amplified noise in the final image, as it is pushed above the HVS visibility threshold, see Fig. 4. To prevent this from happening, the RNTM introduces a *noise-awareness* concept in the tone-mapping process, effectively concealing the noise in the darker parts of the image.

The operator has two mechanisms controlling noise visibility. Firstly, the saliency of tone levels ( $p(l)$  in Eq. 2) is lowered for those tones that are affected by noise. As noise is the highest for the darkest tones, these tones have their contrast and brightness reduced when a tone-curve is computed. Secondly, the amplification of the detail layer is constrained so



**Fig. 4.** Although noise is invisible in the original image, when adapting to the limits of the display this may be pushed above the visibility threshold of the visual system.

that the noise cannot be boosted above the visibility threshold. Those two mechanisms combined effectively conceal (rather than remove) the noise in an HDR sequence. The technique can be used in combination with regular video denoising.

### 3.2. Luminance retargeting

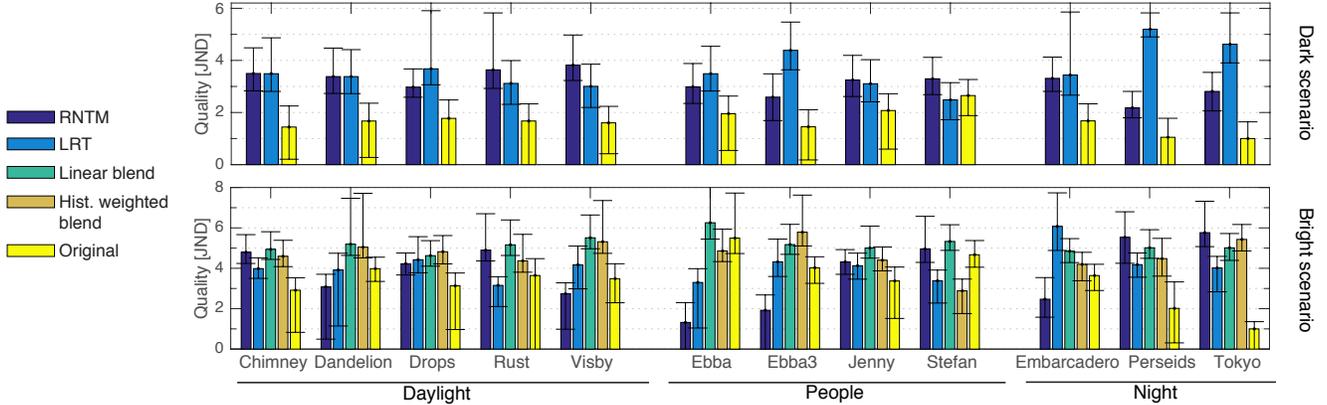
In the global contrast retargeting step of the LRT method, tone-curves are derived using a similar concept as the tone-curves in RNTM (Section 3.1), but optimized for minimal difference in perceived contrast. Also, the tone-curves do not account for image content, and thus stay static for a certain retargeting scenario irrespective of input. However, considering widely different image content it may in many situations be beneficial to adapt the tone-curve based on the image [13]. In the case of LRT this could be accomplished in the optimization of the tone-curves, utilizing the image histogram as a saliency term. However, this results in a non-trivial problem without a simple solution. Instead, we look at the content-adaptive tone-curves from RNTM which can be found in a very fast procedure. We use these in combination with the LRT tone-curves, since simply replacing them would defeat the purpose of minimizing perceived contrast differences.

A number of different tone-curve combinations were tested, both in the intensity and gradient domains. Among these, two methods were chosen for evaluation, which were confirmed to produce most preferred results in an initial test:

**Linear blending:** Given tone-curves  $v_{lrt}(l)$  and  $v_{rntm}(l)$ , that maps from log luminance  $l$  to display value  $v_{lrt}$  or  $v_{rntm}$ , a simple linear blend is performed of the tone-curve nodes using a blend value  $\alpha \in [0, 1]$ :

$$v(l) = (1 - \alpha)v_{lrt}(l) + \alpha v_{rntm}(l) \quad (3)$$

**Histogram weighted blending:** The content-aware capabilities of the RNTM tone-curve can also be reflected in the blending with the LRT tone-curve, so that the blending uses more of the LRT version where the histogram bin probabilities are low. This is achieved by blending with the normalized image histogram,  $\alpha = \bar{p}(l)$ , in Equation 3.



**Fig. 5.** Results of the pairwise comparison experiment. **(Top):** dark scenario with no ambient lighting and  $5 \text{ cd/m}^2$  display peak luminance. **(Bottom):** bright scenario with  $1000 \text{ lux}$  ambient lighting and a peak luminance of  $35 \text{ cd/m}^2$ .

#### 4. EVALUATION

In order to evaluate which tone-curve works best we performed a subjective pairwise comparisons experiment, which compared to *e.g.* ranking or rating experiments tend to reveal subtle differences better [14]. We used a Samsung Galaxy Tab S2 viewed under two conditions: a completely dark environment with a display peak luminance of  $5 \text{ cd/m}^2$ , and an approximately  $1000 \text{ lux}$  ambient light environment using a display peak luminance of  $35 \text{ cd/m}^2$ . Each scenario used 12 input images, capturing a variety of scenes. The experiment included 10 participants, who in each trial were shown two images and asked to choose the one they preferred.

For the dark environment, the shape of the RNTM tone-curve was close to a linear ramp due to the low black level, *i.e.* no content-adaptation is needed. The mixed tone-curves were therefore excluded, and the experiment was performed only in order to confirm that LRT was preferred over the original images, both with original and RNTM tone-curves.

In the tone-curve combinations (Equation 3), the RNTM tone-curve was evaluated globally and without noise-awareness. The linear blend was set to the mean of the curves,  $\alpha = 0.5$ .

##### 4.1. Results

The results of the pairwise-comparison experiment are shown in Figure 5. These are scaled in JND units under Thurstone Case V assumptions, where 1 JND corresponds to 75% discrimination threshold. The error bars denote 95% confidence intervals computed by bootstrapping. The results are reported per scene rather than averaged, because JND is a relative measure with different absolute values per scene. The scenes have been sorted in to daylight scenes, scenes with people and night scenes. As expected, looking at the plots, the results vary a lot with both scene and with experimental environment.

From the results we make the following observations: 1) for all images at least one of the LRT versions is preferred over, or on par with, the original image; 2) the scenes containing people show results where the original image performs on

par with the retargeted images; 3) excluding people scenes, overall the retargeted images clearly perform better than the original images; 4) in the dark environment, both LRT with the original and RNTM tone-curve clearly outperform the original version. From the observations, it seems the most difficult cases are the ones with people, where often the original image performs equally to retargeted images. One reason could be that for faces a softer look may be more appealing, so that loss of details due to low peak luminance is actually advantageous. Comparing the results, it is clear that overall the retargeted images perform better than their original counterparts. However, it is difficult to tell which tone-curve used performs the best, since it varies between scenes. Instead, inspecting the consistency of the results, we see that the linear blend is among the top 3 preferred in all images, and it is always preferred over the original image. Its benefits are also confirmed by visual inspection – while the separate tone-curves in some situations can give too extreme results, the linear interpolation always improves on these, see Fig. 3.

#### 5. CONCLUSION AND FUTURE WORK

This paper presented a tone mapping algorithm combining content-aware tone-curves from RNTM together with the static tone-curves of LRT. The evaluation shows that the approach has promising results, and suggests that a combination has advantages. For future work we will extend the approach with other features from RNTM, such as local tone-curves, noise-awareness and detail preservation, and evaluate it in the context of retargeting using HDR video.

#### 6. ACKNOWLEDGEMENTS

This project was funded by the Swedish Foundation for Strategic Research (SSF) through grant IIS11-0081, Linköping University Center for Industrial Information Technology (CENIIT), the Swedish Research Council through the Linnaeus Environment CADICS, and COST Action IC1005.

## 7. REFERENCES

- [1] G. Eilertsen, R. K. Mantiuk, and J. Unger, “Real-time noise-aware tone mapping,” *ACM Trans. Graph.*, vol. 34, no. 6, pp. 198:1–198:15, Oct. 2015.
- [2] R. Wanat and R. K. Mantiuk, “Simulating and compensating changes in appearance between day and night vision,” *ACM Trans. Graph.*, vol. 33, no. 4, pp. 147:1–147:12, July 2014.
- [3] M. D. Tocci, C. Kiser, N. Tocci, and P. Sen, “A versatile hdr video production system,” *ACM Trans. Graphics*, vol. 30, no. 4, pp. 41:1–41:10, 2011.
- [4] J. Kronander, S. Gustavson, G. Bonnet, A. Ynnerman, and J. Unger, “A unified framework for multi-sensor HDR video reconstruction,” *Signal Processing : Image Communications*, vol. 29, no. 2, pp. 203 – 215, 2014.
- [5] J. Froehlich, S. Grandinetti, B. Eberhardt, S. Walter, A. Schilling, and H. Brendel, “Creating Cinematic Wide Gamut HDR-Video for the Evaluation of Tone Mapping Operators and HDR-Displays,” in *Proc. SPIE 9023, Digital Photography X*, 2014, pp. 90230X–90230X–10.
- [6] G. Eilertsen, R. Wanat, R. K. Mantiuk, and J. Unger, “Evaluation of Tone Mapping Operators for HDR-Video,” *Computer Graphics Forum*, vol. 32, no. 7, pp. 275–284, 2013.
- [7] R. Boitard, R. Cozot, D. Thoreau, and K. Bouatouch, “Zonal brightness coherency for video tone mapping,” *Signal Processing: Image Communication*, vol. 29, no. 2, pp. 229–246, 2014.
- [8] T. O. Aydin, N. Stefanoski, S. Croci, M. Gross, and A. Smolic, “Temporally coherent local tone mapping of HDR video,” *ACM Trans. Graphics*, vol. 33, no. 6, pp. 1–13, 2014.
- [9] V. Aurich and J. Weule, “Non-linear gaussian filters performing edge preserving diffusion.,” in *DAGM-Symposium*, Gerhard Sagerer, Stefan Posch, and Franz Kummert, Eds. 1995, Informatik Aktuell, pp. 538–545, Springer.
- [10] C. Tomasi and R. Manduchi, “Bilateral filtering for gray and color images,” in *Proc. International Conference on Computer Vision 6*, 1998, pp. 839–846.
- [11] F. Durand and J. Dorsey, “Fast bilateral filtering for the display of high-dynamic-range images,” *ACM Trans. Graphics*, vol. 21, no. 3, pp. 257–266, 2002.
- [12] P. Perona and J. Malik, “Scale-space and edge detection using anisotropic diffusion,” *IEEE Trans. Pattern Analysis Machine Intelligence*, vol. 12, no. 7, pp. 629–639, 1990.
- [13] G. Ward Larson, H. Rushmeier, and C. Piatko, “A visibility matching tone reproduction operator for high dynamic range scenes,” *IEEE Trans. Visualization and Computer Graphics*, vol. 3, no. 4, pp. 291–306, 1997.
- [14] R. K. Mantiuk, A. Tomaszewska, and R. Mantiuk, “Comparison of four subjective methods for image quality assessment,” *Computer Graphics Forum*, vol. 31, no. 8, 2012.