High Dynamic Range Video in Postproduction

Barbara Pröglhöf

Diplomarbeit Jahrgang 2011/2012
Masterstudiengang Digitale Medientechnologien, FH St. Pölten

barbara.proeglhoef@gmail.com

1 Introduction

Capturing High Dynamic Range images is a popular way of depicting the real world as closely to what can be seen with the human eye as possible and can't be processed by common camera sensors. While even amateur photographers can easily handle the workflow for still pictures, the technology of HDR video is still a field that is growing and has yet to be fully explored. In this paper, it is investigated how far the field HDR has progressed and it takes a special view on postproduction techniques that can be used to work with high dynamic range footage. In order to do so, an overview is given regarding the most important aspects of HDR imagery, that are crucial in the understanding of any workflows that one might use when capturing and editing such material. It is explained how HDR-video can be captured, including both actual hardware solutions and ways to overcome the limitations of LDR hardware by combining multiple components of already known equipment.

An important piece of this work is the analysis of various tone mapping operators, which make it possible to view high contrast images and video on standard low dynamic range devices. It is studied which operators produce a pleasing result and which methods and extensions have to be added to apply photographic operators to moving imagery. Furthermore, software solutions for postproduction of high dynamic range videos are searched for, compared and tested regarding possibilities and optical quality.

The remainder of this paper is organized as follows: Section 2 reviews current technologies for capturing HDR-content. Section 3 discusses the various approaches in HDR-Tonemapping, while section 4 will describe current available software and its appliance on test footage.

2 Capturing HDR Content

As stated before, capturing HDR-photographs is a well-known procedure even to amateur photographers. The most commonly used method to capture an image with enhanced dynamic range – exposure bracketing – is to take a sequence of images of the same visual motif. Each image is captured with a differing exposure, as the resulting pictures can later be combined to form one image with a high dynamic range in a process called 'exposure fusion'.

Exposure can be varied by three means: One option is the use of sundry neutral density (ND) filters between exposures. This way it can be assured that motion blur and depth of field are the same in each taken picture, which can boost the overall quality of the resulting HDR-Image. A disadvantage of this method is that changing the filters takes a lot of time and also moves the camera which can create artifacts on the final output. Another option to differ exposure is to change the f-stop value. While it is quicker than using ND-filters and doesn't require additional hardware, it results in varying depth of field in each of the images, posing a problem in the later applied exposure fusion, leading to artifacts. The third option is to vary shutter speed. This provides a constant depth of field throughout the image sequence.

After capturing the sequence, all the images are combined by exposure fusion. Assuming that each pixel in a picture was correctly illuminated in at least one of the exposures, the actual scene value can be calculated using a weighting function (Debevec & Malik, 1997).

Capturing HDR-Video content with the method described above poses some obvious problems: To film moving objects, it would be necessary to have a camera that can capture multiple exposures for each frame; Also, using varying shutter speed to gain those different images would result in artifacts due to a varying amount of motion blur in each exposure.

In order to capture high dynamic range video content, two basic principles have been developed: temporal & spatial exposure differentiation. Temporal exposure differentiation implies a camera that captures multiple exposures for each frame. The advantage here is that the full resolution of a camera can be used; but in order to have multiple exposures a very high frame rate is necessary, which leads to very short exposure times for each frame. Filming dark scenes with temporal differentiation can therefore lead to noise artifacts due to increased amplifying of the signal (Myszkowski et al., 2008, p. 17). Spatial exposure difference, on the other hand, captures multiple exposures of the same image by either splitting the incoming light and redirecting it to various lenses (or sensor chips), or by applying a pattern of filters onto a chip, which then captures i.e. four different exposure values within a block of 2 x 2 pixels. Those pixels can later be merged to one HDR-data pixel. An obvious drawback here is that the resolution of the final image gets reduced by 3/4 of the original (ibid., p. 19). An example for cameras using a beam splitter is the 'AMP HDR' camera, which splits the light beam into three parts that are redirected by mirrors onto three different sensor chips (Tocci et al., 2011).

3 HDR-Video Tonemapping Techniques

Due to technological limitations HDR content cannot be displayed on conventional display devices such as computer monitors. Even though there are HDR displays that are capable of doing so, it is still necessary to find ways to work with HDR content on common devices. Methods that were developed to accomplish this are called tonemapping. The goal of tonemapping is to compress dynamic range (so an image can be displayed on a normal monitor) while keeping the details that were attained through HDR capturing and exposure fusion.

Tonemapping algorithms, also called tonemapping operators (TMO), can be divided in two categories: local and global TMOs. The main difference between local and global tonemapping is that global tonemapping only takes features into account that can be taken from the picture as a whole, like the average luminance in a picture. Using these features a global tonemapping operator applies a linear filter to decrease dynamic range. To calculate the luminance value of a certain pixel, a local TMO also looks at neighboring pixels which makes them harder to implement and slower in calculation, but give better results compared to global operators (Reinhard, 2005).

Even though different local tonemapping operators use varying techniques to achieve good results, for most of them a basic functionality can be described: To compress the dynamic range of a given picture, it is divided in two or more layers. In most cases, two layers are used; One for details and one base layer with the global illumination values. While the detail layer stays untouched, the base layer gets compressed to be then combined with the details again. This way it can be assured that detail is not lost while the dynamic range is compressed enough to be displayed on a common device.

The factor that directly correlates to the resulting output quality is the extraction of the base layer (Impoco, Marsi & Ramponi, 2005).

While there are a variety of tonemapping operators available for still imagery, these cannot be used for the purpose of HDR video without any alteration. The main problem is that algorithms for photographs are implemented to work in a way to get the most out of a single picture. If one would apply an algorithm like that to video footage, it would result in artifacts, mostly flickering, as each frame of the video would be compressed independently to its predecessor or successor.

Algorithms that address this problem use different approaches to deal with flickering (and other) artifacts. Chul Lee and Chang-Su Kim extended Fattals work (Durand & Dorsey 2002) – which is based on a gradient map that is derived of an image and then compressed – to be suitable for video by calculating motion vectors for each pixel. If the difference in luminance for each pixel between two frames exceeds a certain threshold, a weighting function is used to minimize it (Lee & Kim 2007, p. 2). In his work "Adaptive Reduction of the Dynamics of HDR Video Sequences" Impoco uses an approach based on the formula that was described earlier in this section: The HDR footage is divided in two layers using a blur filter, and the luminance layer of a given frame is not only compressed, but also compared to the luminance layer of its antecessor, equalizing the difference (Impoco, Marsi & Ramponi, 2005). Another approach on HDR video tonemapping is Jinnos tonemapping based on gamma blending: In this case, two gamma curves are combined and applied to the illumination layer – one enhancing detail in low luminance regions, the other in high luminance regions – and a threshold again defines how much the illumination between two consecutive can vary without being equalized by a weighting function (Jinno, Mouri & Okuda, 2010).

4 HDR Video Postproduction Software

A main goal of the presented work was to research which software solutions are currently available to work with high dynamic range video footage, and to compare the output that can be produced regarding the visual quality. Tonemapping, as described in section 3, is one important step in postproduction because current workflows are still adapted to a final output in low dynamic range. There is a variety of software solutions available that offer video tonemapping, but they differ greatly in respect of their functional scope. The following section will give a quick overview of the most important findings.

Pfstools, a set of commandline tools, implements eight popular tonemapping algorithms for still images and gives the user extensive possibilities to c ontrol the final output. Three TMOs – developed by Sumanta Pattanaik (Pattanaik 2000), D Rafal Mantiuk (Mantiuk 2008) and Erik Reinhard (Reinhard 2002) – were adapted for using them on video footage, but tests done with an HDR timelapse video show that even though the quality of the tone -mapping itself is pleasing, flickering is still an issue. Another available option is to use plugins for compositing packages – in this case AfteEffects – which are easier to use as they have a GUI, but offer a lot less functionality when compared to pfstools. 'GingerHDR' has multiple modes that can be chosen from depending on the video footage, but, as seen with pfstools, flickering artifacts cannot be avoided without further processing. 'EFX HDRI Comp', another plugin for After Effects, is very limited in its functionality as it is not able to work with actual HDR-files and cannot combine more than two different exposures. Flickering is not an issue here, but the overall visual impression lacks quality.

Using the compositing package Nuke, a simple tonemapping operator was built with the standard nodes available in the software. The main architecture of the TMO is based on the algorithm developed by Durand (Durand & Dorsey 2002) and uses fast bilateral filtering to separate the overall luminance from detail information. To avoid flickering artifacts, a de-flickering plugin was used.

When taking HDR video postproduction a step further, it becomes clear that working with actual HDR data throughout the whole pipeline will be a main aspect in future appliances. In order to accomplish this, a standardization of video compression and encoding will be necessary. GoHDR media suite is the only currently public available software that offers HDR-compression without tonemapping the footage. Only a few encodings are available, and as of now, only the included videoplayer can read the generated output files.

In summary, it can be said that there is still room for a lot of developments in the field of HDR video editing, as it is very likely that the future tendency will be away from tonemapping towards a full HDR-postproduction pipeline including HDR-displays. In order to accomplish this, standardizations have to be made and the underlying tools need to be further developed to produce pleasing results.

References

Debevec, Paul & Malik, Jitendra (1997), Recovering High Dynamic Range Radiance Maps from Photographs, in: *Siggraph* '97 – *Proceedings of the 24th annual conference on Computer graphics and interactive techniques*. ACM Press/Addison-Wesley Publishing, pp. 369–378

Durand, Frédo & Dorsey, Julie (2002), Fast bilateral filtering for the displayof high-dynamic-range images, in: *Siggraph '02 – Proceedings of the 29th annual conference on Computer graphics and interactive techniques*. ACM, pp. 257–266

Impoco, G., Marsi, S. & Ramponi, G. (2005), Adaptive Reduction of the Dynamics of Video Sequences, in: *IEEE International Conference on Image Processing 2005 (ICIP 05)*, Genova, Italy, 11–14 Sept. 2005.

Jinno, T., Mouri, K. & Okuda, M. (2010), HDR video tone mapping based on gamma blending, in: *IEEE International Conference on Image Processing*, 26–29 Sept. 2010, pp. 2521–2524

Lee, Chul & Kim, Chang Su (2007), Gradient Domain Tone Mapping of High Dynamic Range Videos, in: *IEEE International Conference on Image Processing*, 16 Sept.–19 Oct., 2007, pp. III-461–III-464

Myszkowsi, Karol, et al. (2008). *High Dynamik Range Video*. Morgan & Claypool Publishers

Reinhard, Erik (2005), Dynamic Range Reduction inspired by Photoreceptor Physiology, in: *IEEE Transactions on Vizualization and Computer Graphics* Vol. 11, Issue 1 (Jan. 2005), pp. 13–24. Available at: http://www.cs.bris.ac.uk/~reinhard/papers/tvcg2005.pdf

Tocci, M., Kiser, C., Tocci, N. & Sen, P. (2011), A Versatile HDR Video Production System, in: *ACM Transactions on Graphics (Proceedings of SIGGRAPH 2011)* vol. 30, number 4