The HDR-video pipeline

- FROM CAPTURE AND IMAGE RECONSTRUCTION TO COMPRESSION AND TONE MAPPING

J. Unger†1 F. Banterle[2] G. Eilertsen[1] R. K. Mantiuk[3]

[1]Media and Information Technology, Linköping University, Sweden
[2]Visual Computing Laboratory at ISTI-CNR, Italy
[3] Computer Laboratory, University of Cambridge, UK

Abstract

High dynamic range (HDR) video technology has gone through remarkable developments over the last few years; HDR-video cameras are being commercialized, new algorithms for color grading and tone mapping specifically designed for HDR-video have recently been proposed, and the first open source compression algorithms for HDR-video are becoming available. HDR-video represents a paradigm shift in imaging and computer graphics, which has and will continue to generate a range of both new research challenges and applications. This intermediate-level *tutorial will give an in-depth overview of the full HDR-video pipeline present several examples of state-of-the-art algorithms and technology in HDR-video capture, tone mapping, compression and specific applications in computer graphics.*

Presentation slides and more information can be found at the tutorial web-page:
http://vcg.isti.cnr.it/Publications/2016/UBEM16a/

Categories and Subject Descriptors (according to ACM CCS): I.3.3 [Computer Graphics]: Display algorithms—I.4.1 [Image Processing and computer vision]: Digitization and Image Capture—I.4.2 [Image Processing and computer vision]: Compression (Coding—I.4.3 [Image Processing and computer vision]: Enhancement—

1. Introduction

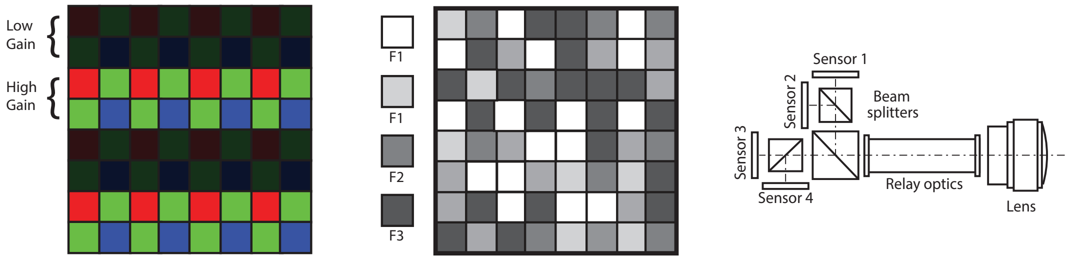
High dynamic range (HDR) video is a rapidly emerging technology, which will offer unprecedented improvements in viewing experiences for both high end cinemas as well as consumer level products. It has and will continue to generate new applications also in computer vision and graphics such as new methods for scene capture and image based lighting. Driven by the demands for extended visual fidelity and artistic freedom, significant research efforts have been spent developing new HDR technologies which are now starting to mature. On the camera and sensor side, we have seen the development of both research prototype cameras [TKTS11, KGBU13] exhibiting a dynamic range of up to 20 - 24 *f-stops*, and professional HDR-camera systems such as the Arri Alexa XT and the Red Epic Dragon with an extended dynamic range of up to 14 - 16.5 *f-stops*. On the processing and production side, major visual effects (VFX) studios are meeting this ongoing trend by developing fully HDR-enabled production pipelines putting a completely new toolset including methods and algorithms for rendering, color grading, and tone mapping in hands of the artists. Also on the display side, HDR technology is in strong focus. Manufacturers, e.g. Sim2, have moved towards extending the dynamic range using high contrast local dimming techniques and Dolby Vision X-tended Dynamic Range PRO has been announced.

HDR-video has over the last few years gone through remarkable developments, and is now starting to mature as a technology. We foresee that HDR-video will generate both a range of new applications as well as many new research challenges in imaging, graphics, and display technology.

2. Tutorial outline and schedule

This *intermediate-level* tutorial will go through the full HDR video pipeline and give an in-depth overview of methods

† jonas.unger@liu.se



(a) varying gain settings (b) spatially varying filters (c) multiple sensors with ND-filters

Figure 1: illustrates two approaches for capturing HDR-video. (a) shows a sensor with a Bayer pattern filter array running two different ISO/gain settings simultaneously for different pairs of pixel rows, see e.g. [HKU14], (b) a sensor with a spatially varying natural density (ND) filter mask with random filter positions (the different filter densities are denoted F0, F1, F2, and F3) , see e.g. [NB03], and (c) an illustration of a multi-sensor HDR camera using a beam-splitter arrangement and different ND-filters for each sensor, see e.g. [KGBU13]

and algorithms from HDR-video capture and processing to tone mapping and specific applications in computer graphics. The goal is to show that, from a research perspective and many practical applications, much of the technology is readily available to implement and use. The tutorial will go through both the theory and practical tools required to do so. The tutorial covers an introduction to the field and five different sub-topics:

1. Introduction
2. HDR-video capture
3. Visual models in tone mapping
4. Tone mapping for HDR-video
   Break 15 min
5. HDR-video compression
6. Applications of HDR-video in graphics
7. Future work and outlook

The next sections give an introduction to each of the topics covered in the tutorial.

2.1. Introduction

This part will start with the relevant background in HDR imaging in general, introduce the various subtopics related to HDR-video, and define what we mean by *dynamic range*. In the introduction, focus will be put on showing examples of state-of-the-art applications from imaging and graphics using HDR-video as an integral component.

2.2. HDR-video capture

A number of HDR-video camera systems have been presented over the last few years, both research prototypes and commercial systems. Current state-of-the-art methods for capture and algorithms for image reconstruction will be divided into three broad classes: multi-exposure capture [KUWS03, UGOJ04, HLL*10] (often referred to as exposure bracketing), multi-sensor imaging systems [TKTS11, KGBU13, FGE*14], and single sensor cameras where the exposure settings varies over the sensor using e.g. filter masks, [NB03, YMIN10, SBS*12, AAG*14] or spatially varying per pixel gain settings over the sensor [HKU14, HKU15, SHG*16]. Figure 1 illustrates HDR capture systems using: (a) two different per pixel gain settings varying spatially over the sensor, (b) spatially varying natural density filters, and (c) multiple sensor imaging the scene through a common optical system and a beam-splitter arrangement. The three classes, approaches, for HDR video capture are explained in detail with examples of real, up-to-date prototype implementations. Specific attention is given to how sensor noise can be modeled, and how it is used for statistical image reconstruction [GAW*10, HDF10, KGB*14]. This part of the tutorial also gives an overview of commercial cameras for HDR or extended dynamic range capture that are available off-the-shelf such as the Arri Alexa XT and the Red Epic Dragon.

2.3. Visual models in tone mapping

Tone-mapping most often tries to achieve certain perceptual goals, such as reproducing image appearance on a display that has different brightness and contrast than the original scene. Such perceptual goals are quantitatively formulated with the help of visual models. In this part of the tutorial we will review a few examples of such models and analyze how they can be integrated into a formulation of a tone-mapping problem.

First, we introduce a display model, which let us link the digital units driving a display with the physical quantities of light, which are emitted by the display and perceived. Then, we will analyze the three main intents of tone-mapping: best subjective quality, visual simulator and scene reproduction. This will be followed by overview of the four main approaches to tone-mapping and how they rely on visual models. Finally, we will discuss a few most frequently used visual models on the examples of perception-aligned tone mapping operators.

The written notes covering major part of the tutorial can be found in [MMS15]. More information on tone-mapping can be found in Chapters 7 and 8 of [RHD*10] and in Chapters 6 and 7 of [DCMM16].

2.4. Tone mapping for HDR-video

Tone mapping for still images is by now a well researched problem with a wide range of algorithms producing high quality results for a range of different purposes. However, it is not until recently robust tone mapping operators (TMOs) for video have started to appear. One reason for this is that HDR-video footage has not been available. Now, along with the advent of high quality HDR-video cameras (see Section 2.2), researchers have started to adress this challenge.

Applying a global or a local TMO designed for static images often lead to flickering artifacts (global or local). Moreover, simple statistics filtering [RJH04] or computing global statistics [KUWS03] can produce other issues such as temporal object incoherency [BCTB14a] or overall dark appearance. This part of the tutorial will give an overview of the challenges related to tone mapping for HDR-video, and describe in detail what additional requirements the transition from static images to video puts on the filtering processes involved. We will described the current state-of-the-art in the field including methods which try to reduce or eliminate artifacts from strong local contrast variations in the temporal domain [KRTT12, BCTB14b]. We will also review recent improvements in video TMOs [ASC*14, EMU15] which have brought local enhancements from local TMOs for static images [RHD*10, BADC11] into the video domain without introducing ghosting artifacts [LK07] or over-enhancement of the sensor noise. Finally, we will show how to properly evaluate such algorithms using psyschophysical experiments from small to large displays [EWMU13, MBDC14].

2.5. HDR-video compression

Content creators and content distributors have started to switch from standard dynamic range (SDR) to HDR videos. This is driven by the rapidly emerging video and film cameras achieving up to 18 *f-stops* (e.g. see Section 2.2), new tone mapping algorithms, and the appearance on the market of affordable HDR televisions (e.g. the Vizio R-series, http://www.vizio.com/r-series). Content distributors are already offering HDR-ready films, tv-series, etc. For example, Amazon has recently announced that HDR-ready content soon will be available on Amazon Prime TV. However, broadcasting or streaming over the Internet of high resolution HDR-footage requires very high bandwidth. A highly important challenge, which is starting to be addressed in the research community, is therefore HDR-video compression.

After the seminal works by Mantiuk et al. [MKMS04, MEMS06], researchers have started to investigate in efficient compression methods [LK08, MT10, MMM*11, BAPN14, BMP15]. Most of these algorithms are drawback compatible. An HDR-video is typically tone mapped and encoded using a MPEG stream, and its reverse TMO and/or residuals are usually stored in the MPEG stream metadata and/or in another stream. This part will be covered in the tutorial highlighting the key concepts and ideas for achieving modern high quality HDR-video compression. The final part of the tutorial will be focused on the recent trends in HDR-video compression which are defining and standardizing tone mapping curves or electro-optic transfer functions (e.g. MPI's PUs encoding or Dolby's PQ encoding) for having few perceptual distortions and high compression rates [BMP15]. Recently, the MPEG committee, http://mpeg.chiariglione.org/, has started the standardization process of HDR-video into its next standard. We will also give an overview of the recently released open source HDR video codec LUMA HDRv, http://lumahdrv.org.

2.6. Applications of HDR-video in graphics

HDR-video has, in addition to film and video recording, also found other applications in computer graphics. This part of the tutorial will give an overview of how HDR-video is and can be used in rendering applications. In particular, this section will focus on how HDR-video has been used to extend traditional image based lighting (IBL) [Deb98] to include lighting effects in the temporal, [UKL*13b], and spatial, [UKL*13a], domains. Figure 2 shows a comparison between (a) traditional image based lighting using a single panoramic HDR image to represent the lighting in the scene, and (b) the level of visual fidelity achievable using HDR-video to recover both the geometry and the spatially varying lighting conditions in the scene. The advantages and challenges involved in HDR-video based image based lighting will be explained in detail and illustrated with several practical examples. This section will also give an overview of the underlying theory, and how the rendering equation, [Kaj86], can be formulated in order to take into account the different parts of the light transport when rendering virtual objects into real scenes.

2.7. Outlook and future work

The conclusion of the tutorial is that HDR-video is starting to mature as a technology with many current and potential applications. There are, however, a number of important and from a research perspective interesting challenges that still remains to be solved. This part of the tutorial will present an outlook and summarize what those challenges are. Specific topics that will be covered include: temporally robust filtering for base detail layer separation for HDR-video, generation of tone curves (local and global), and video tone mapping. Another important aspect covered in this section will

(a) A single HDR panorama

(b) A scene model recovered using HDR-video

Figure 2: shows a comparison between (a) an IBL rendering where the lighting is captured at a single point in space, and (b) a VPS rendering from [UKL*13a], where the lighting environment is captured as a combination of geometry and HDR lighting information. Both renderings are of high quality, but it is evident that the spatial variations in the scene illumination captured in (b) are missing in the IBL rendering in (a).

be strategies for efficient HDR-video compression with high fidelity image reconstruction.

3. Author biographies

Francesco Banterle is a post-doc researcher at the Visual Computing Laboratory at ISTI-CNR Italy. He received a PhD in Engineering from Warwick University in 2009. During his PhD he developed Inverse Tone Mapping which bridges the gap between Low Dynamic Range Imaging and High Dynamic Range (HDR) Imaging. He holds a BSc and a MSc in Computer Science from Verona University. He is the first co-author of the book "Advanced High Dynamic Range" published by AK Peters in 2011. His main research fields are HDR Imaging and Rendering.

Gabriel Eilertsen is a Ph.D. student in the department of Science and Technology at Linköping University (Sweden). He received his M.Sc. in Media Technology and Engineering from Linköping University (2010), with a master thesis work carried out at visual effects company Tippett Studio (2010, Berkeley, California). Prior to his Ph.D. studies, Gabriel worked as a research engineer on material characteristics and capturing for computer graphics. During the last years, Gabriel has focused on display and distribution of HDR-video, with a thorough survey and evaluation of state-of-the-art within video tone mapping, development of new video tone mapping techniques, and with implementation of the LUMA HDRv open source solution for encoding of HDR-videos in 2015.

Rafał K. Mantiuk is a senior lecturer at the Computer Laboratory, University of Cambridge (UK). He received his PhD from the Max-Planck-Institute for Computer Science (2006, Germany), was a postdoctoral researcher at the University of British Columbia (Canada) and a lecturer at Bangor University (UK). He has published numerous journal and conference papers presented at ACM SIGGRAPH, Eurographics, CVPR and SPIE HVEI conferences, has been awarded several patents and was recognized by the Heinz Billing Award (2006). Rafal Mantiuk investigates how the knowledge of the human visual system and perception can be incorporated within computer graphics and imaging algorithms. His recent interests focus on designing imaging algorithms that adapt to human visual performance and viewing conditions in order to deliver the best images given limited resources, such as bandwidth, computation time or display contrast.

Jonas Unger is an associate professor at the department of Science and Technology at Linköping University, where he is leading the group for Computer Graphics and Image Processing. He received his PhD in visualization and interactive techniques from Linköping University in 2009, and holds a MSc in Media Technology from Linköping University. He has previously worked as a visiting researcher at University of Southern California's Institute for Creative Technologies. Jonas Unger's research interests lie at the intersection of computer graphics and image processing, where he is currently driving projects directed towards High Dynamic Range video including capture, image reconstruction, tone mapping and compression, as well as techniques for capture, reconstruction and reproduction of real world environments for high quality image synthesis.

Acknowledgements

This project was funded by the Swedish Foundation for Strategic Research through grant IIS11-0081, Linköping University Center for Industrial Information Technology, and the Linnaeus research environment CADICS.

References

[AAG*14] AGUERREBERE C., ALMANSA A., GOUSSEAU Y., DELON J., MUSÉ P.: Single shot high dynamic range imaging using piecewise linear estimators. In *ICCP* (2014). 2

[ASC*14] AYDIN T. O., STEFANOSKI N., CROCI S., GROSS M., SMOLIC A.: Temporally coherent local tone mapping of HDR video. *ACM Trans. on Graph. 33*, 6 (2014), 1–13. 3

[BADC11] BANTERLE F., ARTUSI A., DEBATTISTA K., CHALMERS A.: *Advanced High Dynamic Range Imaging: Theory and Practice*. AK Peters (CRC Press), Natick, MA, USA, 2011. 3

[BAPN14] BANITALEBI-DEHKORDI A., AZIMI M., POURAZAD M. T., NASIOPOULOS P.: Compression of high dynamic range video using the HEVC and H.264/AVC standards. In *10th International Conference on Heterogeneous Networking for Quality, Reliability, Security and Robustness, QShine 2014, Rhodes, Greece, August 18-20, 2014* (2014), pp. 8–12. 3

[BCTB14a] BOITARD R., COZOT R., THOREAU D., BOUATOUCH K.: Survey of temporal brightness artifacts in video tone mappin. In *HDRi2014 - Second International Conference and SME Workshop on HDR imaging* (March 2014), IC-1005. 3

[BCTB14b] BOITARD R., COZOT R., THOREAU D., BOUATOUCH K.: Zonal brightness coherency for video tone mapping. vol. 29, pp. 229–246. 3

[BMP15] BOITARD R., MANTIUK R., POULI T.: Evaluation of color encodings for high dynamic range pixels. In *Proc. SPIE 9394, Human Vision and Electronic Imaging XX* (March 2015), SPIE. 3

[DCMM16] DUFAUX F., CALLET P. L., MANTIUK R. K., MRAK M.: *High Dynamic Range Video - From Acquisition to Display and Applications*. Elsevier, 2016. 3

[Deb98] DEBEVEC P.: Rendering synthetic objects into real scenes: bridging traditional and image-based graphics with global illumination and high dynamic range photography. In *SIGGRAPH* (1998), pp. 1–10. 3

[EMU15] EILERTSEN G., MANTIUK R., UNGER J.: Real-time noise-aware tone mapping. *ACM Transactions on Graphics (TOG) SIGGRAPH Asia '15* (2015). 3

[EWMU13] EILERTSEN G., WANAT R., MANTIUK R. K., UNGER J.: Evaluation of Tone Mapping Operators for HDR-Video. *Computer Graphics Forum 32*, 7 (2013), 275–284. 3

[FGE*14] FROEHLICH J., GRANDINETTI S., EBERHARDT B., WALTER S., SCHILLING A., BRENDEL H.: Creating Cinematic Wide Gamut HDR-Video for the Evaluation of Tone Mapping Operators and HDR-Displays. In *Proceedings of SPIE Electronic Imaging* (2014). 2

[GAW*10] GRANADOS M., AJDIN B., WAND M., THEOBALT C., SEIDEL H., LENSCH H.: Optimal hdr reconstruction with linear digital cameras. In *CVPR* (2010). 2

[HDF10] HASINOFF S., DURAND F., FREEMAN W.: Noise-optimal capture for high dynamic range photography. In *CVPR* (2010). 2

[HKU14] HAJISHARIF S., KRONANDER J., UNGER J.: HDR reconstruction for alternating gain (ISO) sensor readout. In *Eurographics 2014 Short Papers* (May 2014), Eric Galin M. W., (Ed.). 2

[HKU15] HAJISHARIF S., KRONANDER J., UNGER J.: Adaptive dualiso hdr-reconstruction. *EURASIP Journal on Image and Video Processing 2015*, 1 (2015), 1–13. 2

[HLL*10] HEO Y. S., LEE K. M., LEE S. U., MOON Y., CHA J.: Ghost-free high dynamic range imaging. In *ACCV* (2010), pp. 486–500. 2

[Kaj86] KAJIYA J. T.: The rendering equation. *Computer Graphics (Proceedings of SIGGRAPH '86) 20*, 4 (1986), 143–150. 3

[KGB*14] KRONANDER J., GUSTAVSON S., BONNET G., YNNERMAN A., UNGER J.: A unified framework for multi-sensor hdr video reconstruction. *Signal Processing : Image Communications 29*, 2 (2014), 203 – 215. 2

[KGBU13] KRONANDER J., GUSTAVSON S., BONNET G., UNGER J.: Unified HDR reconstruction from raw CFA data. In *Proceedings of the IEEE International Conference on Computational Photography* (2013). 1, 2

[KRTT12] KISER C., REINHARD E., TOCCI M., TOCCI N.: Real time automated tone mapping system for HDR video. In *Proceedings of the IEEE International Conference on Image Processing* (Orlando, Florida, September 2012), IEEE, IEEE, pp. 2749–2752. 3

[KUWS03] KANG S. B., UYTTENDAELE M., WINDER S., SZELISKI R.: High dynamic range video. *ACM Trans. Graph. 22* (July 2003), 319–325. 2, 3

[LK07] LEE C., KIM C.-S.: Gradient domain tone mapping of high dynamic range videos. In *ICIP07* (2007), pp. 461–464. 3

[LK08] LEE C., KIM C.-S.: Rate-distortion optimized compression of high dynamic range videos. In *16th European Signal Processing Conference (EUSIPCO 2008)* (2008), pp. 461–464. 3

[MBDC14] MELO M., BESSA M., DEBATTISTA K., CHALMERS A.: Evaluation of hdr video tone mapping for mobile devices. *Image Commun. 29*, 2 (Feb. 2014), 247–256. 3

[MEMS06] MANTIUK R., EFREMOV A., MYSZKOWSKI K., SEIDEL H.-P.: Backward compatible high dynamic range mpeg video compression. *ACM Trans. Graph. 25*, 3 (2006), 713–723. 3

[MKMS04] MANTIUK R., KRAWCZYK G., MYSZKOWSKI K., SEIDEL H.-P.: Perception-motivated high dynamic range video encoding. *ACM Trans. Graph. 23*, 3 (2004), 733–741. 3

[MMM*11] MAI Z., MANSOUR H., MANTIUK R., NASIOPOULOS P., WARD R., HEIDRICH W.: Optimizing a tone curve for backward-compatible high dynamic range image and video compression. *IEEE Trans. on Image Processing 20*, 6 (2011), 1558 – 1571. 3

[MMS15] MANTIUK R. K., MYSZKOWSKI K., SEIDEL H.-P.: High Dynamic Range Imaging. In *Wiley Encyclopedia of Electrical and Electronics Engineering*. Wiley, 2015, pp. 1–81. 3

[MT10] MOTRA A., THOMA H.: An adaptive logluv transform for high dynamic range video compression. In *ICIP* (2010), pp. 2061–2064. 3

[NB03] NAYAR S., BRANZOI V.: Adaptive Dynamic Range Imaging: Optical Control of Pixel Exposures over Space and Time. In *IEEE International Conference on Computer Vision (ICCV)* (Oct 2003), vol. 2, pp. 1168–1175. 2

[RHD*10] REINHARD E., HEIDRICH W., DEBEVEC P., PATTANAIK S., WARD G., MYSZKOWSKI K.: *High Dynamic Range Imaging: Acquisition, Display, and Image-Based Lighting*, 2nd edition ed. Morgan Kaufmann, 2010. 3

[RJH04] RAMSEY S., JOHNSON T., HANSEN C.: Adaptive temporal tone mapping. In *The 7th IASTED International Conference on Computer Graphics and Imaging (CGIM)* (August 2004). 3

[SBS*12] SCHOBERL M., BELZ A., SEILER J., FOESSEL S., KAUP A.: High dynamic range video by spatially non-regular optical filtering. In *Image Processing (ICIP), 2012 19th IEEE International Conference on* (2012), IEEE, pp. 2757–2760. 2

[SHG*16] SERRANO A., HEIDE F., GUTIERREZ D., WETZSTEIN G., MASIA B.: Convolutional sparse coding for high dynamic range imaging. *Computer Graphics Forum 35* (2016). 2

[TKTS11] TOCCI M. D., KISER C., TOCCI N., SEN P.: A versatile hdr video production system. *ACM Trans. Graph. 30*, 4 (July 2011), 41:1–41:10. 1, 2

[UGOJ04] UNGER J., GUSTAVSON S., OLLILA M., JOHANNESSON M.: A real time light probe. In *Eurographics, Short Papers* (2004), pp. 17–21. 2

[UKL*13a] UNGER J., KRONANDER J., LARSSON P., GUSTAVSON S., LÖW J., YNNERMAN A.: Spatially varying image based lighting using HDR-video. *Computers & graphics 37*, 7 (2013), 923–934. 3, 4

[UKL*13b] UNGER J., KRONANDER J., LARSSON P., GUSTAVSON S., YNNERMAN A.: Temporally and spatially varying image based lighting using hdr-video. In *EUSIPCO* (2013). 3

[YMIN10] YASUMA F., MITSUNAGA T., ISO D., NAYAR S. K.: Generalized Assorted Pixel Camera : Postcapture Control of Resolution , Dynamic Range , and Spectrum. *IEEE Transactions on Image Processing 19*, 9 (2010), 2241–2253. 2