High Dynamic Range in Cultural Heritage Applications

Demetris Marnerides, Vedad Hulusic, Kurt Debattista

Abstract High dynamic range (HDR) technology enables the capture, storage, transmission and display of real-world lighting at a high precision as opposed to traditional low dynamic range (LDR) imaging. One of HDR's main features is its ability to reproduce very bright and very dark areas simultaneously. Dynamic range describes the the span between these extrema in the brightness scale. HDR research investigates the generation, capturing, processing, transmission, storage and reproduction of HDR content. Cultural heritage represents our legacy that must be passed on to future generations. As it is increasingly threatened with deterioration, destruction and disappearance, its documentation, conservation and presentation is of high importance. Given the real-world dynamic range and the limitations of conventional capture and display technology, HDR imaging represents an invaluable tool for accurate documentation, virtual reconstruction and visualisation of cultural heritage. HDR is used by academics, museums, and media to visualise the appearance of sites in various periods in time. Physically-based 3D virtual reconstructions are used for studying existing or ruined cultural heritage environments. This in turn enables archaeologists to interpret the past and deduce new historical knowledge. In this chapter we present the HDR pipeline, along with its use for cultural heritage preservation, recreation and presentation.

Demetris Marnerides
University of Warwick, Warwick Manufacturing Group, UK
e-mail: Demetris.Marnerides@warwick.ac.uk

Vedad Hulusic
Bournemouth University, Faculty of Science and Technology, UK
e-mail: vhulusic@bournemouth.ac.uk

Kurt Debattista
University of Warwick, Warwick Manufacturing Group, UK
e-mail: k.debattista@warwick.ac.uk

1 Introduction

As defined by the UNESCO, cultural heritage (CH) is “the legacy of physical artefacts and intangible attributes of a group or society that are inherited from past generations, maintained in the present and bestowed for the benefit of future generations” [54]. Typically, CH is divided into tangible and intangible heritage. The former consists of historic objects, buildings, places, monuments and other artefacts with a significant value for preservation. Their significance can be related to archaeology, architecture, design, science or technology of a culture. Intangible heritage consists of practices, knowledge, customs, oral traditions, skills, oral epics and similar cultural identities. This work mainly deals with tangible CH, mainly considering 3D objects and sites.

The protection and preservation of CH is of immense importance and demonstrates the recognition of the necessity of the past and its related artefacts and stories [52]. While the main way of conservation, preservation and display of tangible heritage is still physical, allowing observing, exploring, studying and appreciating the artefacts, it is exceedingly expensive to maintain and safeguard these invaluable items from being damaged by tourist, the light when displaying them, and other human factors, such as wars and negligence, or natural disasters. Another limitation of physical exhibition is the accessibility limitation as the original artefact can be exhibited at only one physical location. One solution to these problems could be to use technology for digital documentation and presentation of CH. The whole pipeline is summarised in Figure 1.

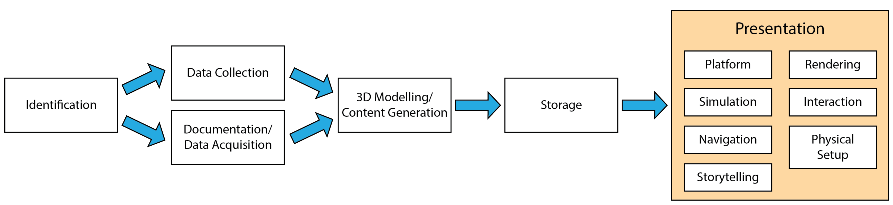


Fig. 1 3D virtual reconstruction pipeline.

After the identification of the artefact or the site based on research on its cultural and historic significance, the first step would be to either collect all relevant data in case of non-existing objects or documentation of existing artefacts or sites. Based on the collected and/or acquired data, the digital representation of the 3D object can be generated. 3D digital representation of the artefact or the site then has to be stored in appropriate format or number of them, suitable for further usage in the last stage of the pipeline. Finally, the object is presented through interactive or non-interactive media, considering several additional elements as displayed in the last stage of the pipeline in Fig. 1.

1.1 High-Fidelity in CH

For many years, computers and computer generated images have been used in archaeology for presentation, understanding and interpretation of ancient cultures. Computers are an integral part of every stage of the pipeline. However, the increase in hardware and software performance and decreased costs, do not inevitably lead to more accurate results. Therefore, in order to have truthful and reliable representations and virtual reconstructions, every step has to be undertaken carefully, including precise data acquisition, content generation, simulation of physical evidence from the site and presentation of the final artefact. One of the components that is often neglected or not given enough attention is the dynamic range of luminance in the original, real scene where the object or site is captured, and the virtual scene with a reconstructed artefact that is being displayed. Although, in some scenarios it might not make a significant effect on the representation, in others can lead to major shortcomings and eventually misguide the interpretation and understanding of the nature, purpose and appearance of the object or site.

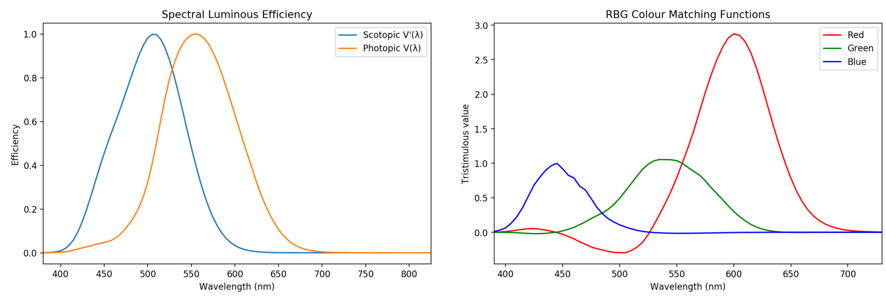
2 High-Dynamic-Range

HDR imaging aims to capture, process and display the entire distribution of visible light of real or simulated scenes. Imaging has a long history of progress; it is suspected that the Shroud of Turin from the 13$^{\text{th}}$ century might be the first recorded photograph [2]. The 20$^{\text{th}}$ century saw an acceleration in the development of photography and imaging techniques, with advanced cameras, display monitors and the transitions to colour images and video.

2.1 Light and Luminance

Light is electromagnetic energy propagating in space as a wave and as discrete particles, photons. These waves/particles are associated with specific frequencies of oscillation, ν that correspond to specific wavelengths λ bound together with a constant, the speed of light in vacuum, c, such that $\lambda\nu = c$. A subset of all possible frequencies/wavelengths forms the visible part of the light spectrum from "violet" $\lambda \approx 400nm$ to "red" $\lambda \approx 700nm$. Each photon is associated with an energy $E = h\nu$, where h is the Planck constant. While imaging in general deals with the whole spectrum of frequencies, some disciplines work with specific parts of the spectrum, for example astronomy uses radio imaging and certain areas in healthcare use X-ray imaging. HDR Imaging focuses on imaging of the visible spectrum.

To measure light distributions, two categories of quantities are considered, radiometric and photometric. Radiometric quantities measure light distributions with a direct physical correspondence to the energy of the light, for example, radiant energy,



(a) (b)

Fig. 2 (a) The photopic, $V(\lambda)$, and scotopic, $V'(\lambda)$, luminous efficiency functions (b) The CIE RGB tristimulous colour matching functions.

Q_e, radiant flux, $\Phi_e = \partial_t Q_e$, and spectral flux,$\Phi_{e,\lambda} = \partial_\lambda \Phi_e$, which is radiant flux per wavelength. The subscript "e" is to denote these quantities as *energetic*.

An important radiometric quantity, radiance, L_e, measures the radiant flux, Φ_e emitted from a surface, A, through a solid angle, ω, along a specific direction, θ, relative to the surface normal and is defined as:

$$L_e = \frac{1}{\cos\theta} \frac{\partial^2 \Phi_e}{\partial A \partial \omega}. \tag{1}$$

Photometric quantities measure light power by taking into account the HVS and its response to different frequencies. These quantities are weighted versions of the radiometric quantities, for example, luminous energy, Q_v, luminous flux, $\Phi_v = \partial_t Q_v$, and spectral luminous flux (or power), $\Phi_{v,\lambda} = \partial_\lambda \Phi_v$, which is spectral flux per wavelength. Figure 2(a) shows the spectral luminous efficiency functions that are used to weigh the radiometric quantities. These functions correspond to the sensitivity of the HVS to particular frequencies in bright (photopic) and dark (scotopic) conditions.

Radiance as defined in Equation 1 fully describes the light field distribution. However, a photometric quantity is preferred over a radiometric one in the case of imaging, since it is more representative of what the result looks like to the human observer. Luminance, L_v, is the photometric quantity that corresponds to radiance, weighing it according to a response function, typically the photopic luminosity function, $V(\lambda)$ from Figure 2(a) and formally defined by the Commission internationale de l'éclairage (CIE) as:

$$L_v = k_m \int_0^\infty L_{e,\lambda} V(\lambda)\, \mathrm{d}\lambda \tag{2}$$

where $L_{e,\lambda} = \partial_\lambda L_e$ is the spectral radiance and $k_m = 683.002 \mathrm{lm\,W^{-1}}$ is the luminous efficacy of a 555 nm (ideal) source. Luminous efficacy measures how well a light source produces visible light. Luminance is measured in candelas per square meter

($\mathrm{cd\,m^{-2}}$) In HDR imaging, the main quantity of interest used is luminance and is denoted as L for ease of notation.

2.2 Colour

The HVS has two types of receptors, *cones*, which are sensitive to relatively high luminance levels (photopic vision: 10^{-2}– $10^{8}\mathrm{cd\,m^{-2}}$), and *rods*, which are sensitive to lower luminance levels (scotopic vision: 10^{-6} – 10 $\mathrm{cd\,m^{-2}}$) [3]. While rods are more abundant (10 – 20 times more), they are of only one type and thus can only produce images of one colour tone, interpreted as levels of grey by the human brain. This is why colour cannot be perceived in dark scenarios.

Cones are of three different types, each responding to different parts of the visible spectrum (with overlap). The combination of these three different signals is interpreted in the brain and perceived as colour. The CIE has defined the RGB colour-matching functions, $r(\bar{\lambda})$, $g(\bar{\lambda})$, $b(\bar{\lambda})$, based on the colour matching experiments of Wright and Guild, as shown in Figure 2(b).

From these functions, given a spectral power distribution, $S_{e,\lambda}(\lambda)$, the CIE RGB *colour space* can be calculated as follows:

$$(R, G, B) = \int_{380}^{830} I(\lambda)\,(r(\lambda), g(\lambda), b(\lambda))\,\mathrm{d}\lambda \qquad (3)$$

Another important colour space is the CIE XYZ colour space which was defined using a different set of colour-matching curves such that the Y component matches the luminance of the pixel. Many other colour spaces, e.g. Adobe-RGB and sRGB [42], are based on or derived from the CIE XYZ colour space, which plays a central role in colour space transformations. The use of colour spaces is multifold, since they not only need to describe the light information, but they also, in many cases, need to correspond to other properties, for example display capabilities, hence the abundance of colour space standards. Most colour spaces can be mapped back to CIE XYZ and this is the main way that the luminance of an image can be computed.

2.3 Contrast and Dynamic Range

When describing a scene, be it real world, virtual or in an image, global or local differences in luminance can be expressed using contrast ratios. While these can be defined in multiple ways [3], the most commonly used definitions are the direct ratio, C_r, Weber contrast, C_W, and Michelson contrast, C_M:

$$C_r = \frac{L_{\max}}{L_{\min}},, \quad C_W = \frac{L_{\max} - L_{\min}}{L_{\min}}, \quad C_M = \frac{L_{\max} - L_{\min}}{L_{\max} + L_{\min}}. \qquad (4)$$

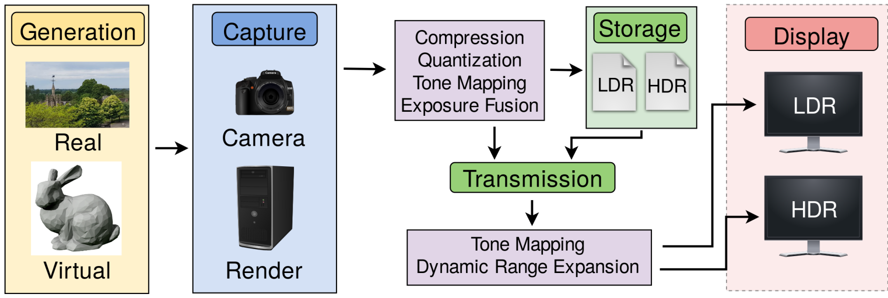


Fig. 3 The HDR pipeline, from generation to display. The capture, storage and display stages also have their own individual processes which change the generated information, for example Camera Response Functions (CRFs) and gamma encoding.

While contrast measures the relative intensity between the maximum and minimum luminance values, *dynamic range* can have slightly different meaning depending on the context. For example, in a real world scene with a (practically) continuous luminance distribution, the dynamic range of the scene is synonymous with the contrast ratio. It describes the extent of the luminance range. However, when the same scene is captured by a camera sensor, due to the imperfections and noise in the capturing process, it is more useful to describe the dynamic range of the captured content using the Peak Signal to Noise Ratio (PSNR):

$$\mathrm{PSNR}_{\mathrm{capture}} = 20 \log_{10} \left(\frac{L_{\max}}{\sigma_{\mathrm{noise}}} \right). \tag{5}$$

where σ_{noise} is the standard deviation of the noise and $L_{\max}$ is the maximum captured luminance.

The dynamic range of a digitally stored image can be thought of as a measure of the global contrast ratio, in combination with the light information/entropy of the contents of the image image, or alternatively, the minimum number of "grey" levels needed to describe that image. The more information in the scene, the more bits are required to encode it digitally.

3 High Dynamic Range Pipeline

There are multiple stages involved in the lifetime of an image. The first stage is the generation of the image content and its capture, followed by the storage or transmission of the captured scenes. The stored (or transmitted content) is then processed for display. Figure 3 shows an overview of this process. The rest of this section discusses these stages in more detail.

3.1 Capture

The two main types of content originate from real world scenes that are captured using camera sensors, or from virtual scenes captured by rendering, for example Path Tracing. While rendering can in theory be performed with arbitrary dynamic range and precision, capturing real world scenes with a camera is limited by the dynamic range and sensitivity of the camera sensor.

Camera sensor cells can only receive a certain amount of light before becoming saturated. This, combined with the acceptable noise level of the sensor signal, defines the maximum dynamic range the sensor can capture and can be described in terms of the PSNR value from Equation 5. If the sensor is exposed to incoming radiation, eventually all pixels will become saturated. The *exposure time* is adjusted to minimise the saturated cells. However, with low exposure times, dark parts of the scene might not get enough exposure time and thus become indistinguishable from sensor noise and are interpreted as black pixels. This sensor limitation means that it is in many cases impossible to capture the whole dynamic range of a scene with a single exposure.

The most popular way of capturing HDR content using current sensor technology is to take multiple exposures with varying exposure times and then combine them to form a single image [44], a process referred to as *exposure bracketing* followed by *exposure fusion*. Given N different exposures, I, with exposure times T, then the resulting image $\hat{I}$ can be computed using:

$$\hat{I} = \frac{1}{\sum_{i=1}^{N} w(I_i)} \sum_{i=1}^{N} \frac{1}{T_i} w(I_i) I_i \tag{6}$$

where w is a weighing function, providing higher weights for well-exposed mid-range values and less weight for noisy black-level and saturated pixel values. Equation 6 assumes that the camera responses are linear, which in most cases are not, and the non-linear Camera Response Function (CRF), f_{CRF}, must be taken into account, such that:

$$\hat{I} = \frac{1}{\sum_{i=1}^{N} w(I_i)} \sum_{i=1}^{N} \frac{1}{T_i} w(I_i) f_{\mathrm{CRF}}^{-1}(I_i). \tag{7}$$

Another assumption is the spatial correspondence of the pixels between exposures, which is often not true, since the camera might not be very stable during or between captures, or the scene might change due to motion. Misalignment due to camera shaking can be alleviated by aligning the images before merging [60]. Motion in the scene causes ghosting artefacts, which can be addressed using multiple methods [53], for example patch matching, weighted filtering of intensity transfer functions, or weighted non-parametric models.

3.2 Storage

Captured or generated HDR images are usually stored in a digital form using floating point numbers. They are stored as linear RGB (not gamma corrected) but the luminance values are usually relative and do not correspond to the real world luminance values. One of the first formats introduced for storing HDR content was the Ward RGBE format [19] (.hdr or .rad extension) which was initially used to store radiance values from rendered scenes. It uses 8 bits per channel, but introduces a fourth channel E for the exponent of the RGB values (which are assumed to be close, hence they share an exponent), effectively using 32 bits instead of 24 for each pixel. The format can also use the XYZE encoding. The compression used is run-length encoding, which uses a count-value strategy to represent repeated consecutive values [44].

A more recent HDR image format, the OpenEXR (.exr extension - EXtended Range), developed by Industrial Light & Magic (ILM), uses 16 bit half-float precision for each of the three channels but also supports 32 and 24 bit per channel encodings. This results in larger file sizes than the .hdr format but can represent content much more accurately. The EXR format uses the PIZ lossless wavelet compression [44]. The JPEG-HDR encoding uses a tone mapped version of the original, along with a grey-scale ratio image to lossy compress the HDR image.

3.3 Display

The final stage of the HDR pipeline is the display of the captured or stored content. For the display of any type of content, be it LDR or HDR, there needs to be some adaptation, particularly of the dynamic range and contrast, such that an accurate and perceptually preferred viewing experience is achieved.

LDR Displays

LDR Content is linearised (if it was gamma encoded) and its colour space is adapted according to the display specifications. Its range is matched to that of the display, so there is no adaptation in terms of contrast or dynamic range. A gamma correction curve is applied, again depending on the specifications. Most displays are created with a specific colour space as a specification and vice versa. *HDR Content* needs to be altered to fit the dynamic range of the display. This requires for information to be removed from the HDR image. There are multiple ways to remove this information, each one resulting in a different viewing experience. These methods are called Tone Mapping Operators (TMOs) and are discussed in more detail in Section 4.

HDR Displays

LDR Content must be mapped to match the dynamic range of the screen via some EO. These are introduced and discussed in more detail in Section 5. *HDR Content* needs to have its dynamic range adapted to match that of the display, depending on the level of mismatch. This case is essentially the same as that of HDR content being displayed on an LDR display (just one with a wider range), since no display can match the (theoretical) real world lighting that HDR content can encode.

4 Tone Mapping

Tone mapping is the process of reducing the dynamic range of an image for display. Dynamic range reduction is performed by compressing, clipping or quantising the range or most commonly via a combination of the three. Multiple TMOs have been introduced, each one focusing on different aspects of the problem, but in general, these operators can be categorised as either global or local [3]. Global operators perform an operation on all pixels in the same fashion, whereas local operators adjust the operation given local context of the specific pixel neighbourhood. Local operators may depend on local properties, for example average contrast, or frequency aspects, for example if an edge exists (detail) or not.

There are multiple properties to consider when tone mapping an image, for example how global and local contrast is preserved, the average image brightness, as well as the tonal contrast between different colours. In the case of 8-bit precision, which is the most commonly used LDR format, the operation is defined as follows:

$$I_{\mathrm{TM}} = f_{\mathrm{TM}}(I_{\mathrm{HDR}}), \text{ where } f_{\mathrm{TM}} : \mathbb{R}^+ \rightarrow [0, 255] \tag{8}$$

where I_{HDR} is the HDR content, I_{TM} is the resulting tone mapped image and f_{TM} is the TMO. In most algorithms the image luminance, L_{HDR}, is first computed by taking the Y channel from the CIE XYZ colour space. The tone mapped single channel luminance, L_{LDR}, of the result is then computed, which is used to scale all three RGB input channels:

$$(R, G, B)_{\mathrm{TM}} = \frac{L_{\mathrm{LDR}}}{L^s_{\mathrm{HDR}}} (R, G, B)^s_{\mathrm{HDR}} \tag{9}$$

where $s \in (0, 1]$ is a factor affecting colour saturation. For display, or storage of the result, a gamma encoding is applied and the tone mapped image is quantised and clipped at the two ends (0, 255).

$$I_{\mathrm{LDR}} = \lfloor \mathrm{clamp}(I_{\mathrm{TM}}, 0, 255)^{\gamma} \rfloor \tag{10}$$

The simplest way to tone map is to use a linear function. This is a global operation and is equivalent to taking a single exposure from the HDR image, as a camera would from a real scene (assuming a linear CRF):

$$L_{\mathrm{TM}} = eL_{\mathrm{HDR}} \tag{11}$$

where the factor e adjusts the exposure level. When the number of well exposed (non clipped) pixels is maximised, then e is called automatic exposure. Debattista et al. [8] propose the selection of e such that the information loss due to the clipping and quantisation of the HDR histogram is minimised, in which case e is termed the optimal exposure.

Drago et al. [14] introduce a global adaptive logarithmic function which adjusts luminance in logarithmic space with the logarithm base adapted according to the luminance of the given pixel. Logarithmic spaces better correspond to the HVS sensitivity to brightness. The display adaptive operator by Mantiuk et al. [33], takes into account the specific characteristics of the target display (e.g. LCD, OLED or e-paper) such that visible contrast distortions are minimised. An that takes into account the HVS is introduced by Reinhard et al. [41], inspired by photoreceptor physiology. The resulting LDR image is computed by modelling the adaptation of cones to different levels of luminance.

Urbano et al. [57] evaluate TMOs when applied for small screen devices, for example mobile phones. The psychophysical experiments performed show that the choice of TMO depends on the display type. They considered Liquid Crystal (LCD), Cathode Ray Tube (CRT) and Seven-Segment Displays (SSD). The different characteristics of the display type required different adjustment to the TMOs in order to get optimal performance.

HDR video is more difficult to tone map since there is an additional time correlation in effect. Tone mapping video sequences using a simple frame by frame approach frequently exhibit artefacts, for example brightness flickering [6]. Boitard et al. [5] propose a post-processing step that also addresses object incoherency and brightness incoherency besides temporal flickering. Eilertsen et al. [16] provide a comprehensive comparative review of video TMOs.

Tone mapping is a forward problem, whose solution varies depending on the constraints of the user, for example HVS response or display properties. Reconstructing HDR signals from LDR is the inverse version of this problem is discussed in the next section.

5 Dynamic Range Expansion

Expansion Operators (EOs), also known as inverse or reverse tone mapping operators, attempt to generate HDR content from LDR content. EOs can generally be expressed as:

$$\tilde{I}_{\mathrm{HDR}} = f_{\mathrm{EO}}(I_{\mathrm{LDR}}), \text{ where } f_{\mathrm{EO}} : [0, 255] \rightarrow \mathbb{R}^{+} \tag{12}$$

where $\tilde{I}_{\mathrm{HDR}}$ is the expanded HDR content and f_{EO} is the EO. In this context, dynamic range expansion could be considered an ill-posed problem since there is no unique solution. The missing information that needs to be recovered could have been one of many valid forms, for example a single exposure image with over-exposed sky could have had clouds or clear blue skies.

A variety of methods have emerged that attempt to tackle the issue. Most non-learning based EOs follow some, or all of five steps [3]. First, the LDR image is linearised, via the removal of gamma correction and the application of inverse CRFs. Then the range of well-exposed, non-saturated areas is expanded via some function, locally or globally, depending on the algorithm. The expansion is computed on the linearised single-channel luminance, which is ultimately used to scale all three channels. At this point, as an additional step, some methods attempt to reconstruct badly-exposed areas as well. The final two steps attempt to reduce artefacts due to quantisation or compression, and apply colour-corrective algorithms that might be necessary due to partial saturation of one of the three RGB channels.

Global methods use a straightforward function to expand the content equally across all pixels. One of the first of such methods was the technique presented by Landis [30] that expands content for displaying on an HDR display based on power functions. Akyüz et al. [1] proposed a method that uses a linear transformation with a gamma correction to expand single exposure.

Local methods, similarly to the local tone mapping methods discussed in Section 4, use analytical functions that depend on local image neighbourhoods. These methods also use a locally dependent expand map.

One of the earliest local ITM methods introduced by Banterle et al. [4], initially expands the range by applying the inverse of the PTR TMO [43], although any other invertible TMO could be used. A smooth low frequency expand map is generated by selecting a constellation of bright points and expanding them via density estimation (importance sampling). The resulting luminance is computed by interpolating between the expanded luminance, L_{ITM}, that was computed via the inverse TMO, and the LDR image luminance, L_{LDR}, as follows:

$$\tilde{L}_{\mathrm{HDR}} = E L_{\mathrm{ITM}} + (1 - E) L_{\mathrm{LDR}} \tag{13}$$

where E is the expand map. The role of the expand map is to to avoid quantisation errors that would arise via inverse tone mapping only and also reconstruct content in over-exposed areas of the image.

Classification based methods such as by Meylan et al. [37, 36] and Didyk et al. [12], operate on different parts of the image by classifying these parts accordingly, for example into diffuse and specular areas. The classification is performed via thresholding and filtering in the first method, or by a support vector machine classifier which is also corrected by a user in the second.

Eilertsen et al. [15], use a deep Convolutional Neural Network (CNN) to predict values for saturated areas of badly exposed content, whereas non-saturated areas are linearised by applying an inverse camera response curve. Endo et al. [17] also use a CNN model that predicts multiple exposures from a single exposure which are then used to generate an HDR image using standard merging algorithms. Marnerides et al. [35] use a multi-branch CNN architecture to combine multiple scales of the image and predict HDR images directly from the LDR inputs in an end-to-end fashion.

6 Using HDR in CH applications

HDR imaging holds an important role in CH research and is used not only for extending the luminance range during capture and display, but also for temporal light acquisition of movable light sources and participating media, and for light/sun simulations. The acquisition of HDR data, allows for more accurate documentation and recreation of tangible CH scenarios and understanding of how sites once looked and how they were used.

6.1 Documentation

One way to document CH is by using still and video cameras. This allows for capturing shape, colour and light of existing CH objects in their original or adjusted real world locations. Documenting objects or sites using traditional hardware with LDR capabilities might be insufficient. Therefore, either specialised HDR cameras or suitable iTMO should be utilised to preserve details in high-contrast scenes. This not only provides with greater accuracy and an extended luminance range [27] but also allows for capturing specific light phenomena, such as caustics or participating media which might play an important role in the life of an artefact or historical site. Caustics are formed when light rays get reflected or refracted by a curved specular surface and then hit another surface. Typical examples are those formed through a glass of water or wine, through a water surface or of a metallic surface such as wedding ring. Due to a high luminance value of the concentrated light produced by this phenomenon, traditional LDR technology might not be able to adequately capture and reproduce this effect accurately.

While HDR could be very useful in overcoming such limitations, some artefacts might occur in dynamic scenes. In addition, intrinsically they are not suitable for temporal data acquisition of light that might be important for documentation and investigation of light and colour properties of movable light sources and high-contrast environments. For example, we might want to simulate the day cycle of a historical site, or simply to capture the environment lit by a candle with some airflow around its flame causing it to move. Although the former could be achieved with a still HDR camera using a discrete sampling of a relatively slow moving light/environment, the latter could lead to ghosting or similar bending artefacts. For example, capturing dynamic caustics might result in ghosting effect when combining multiple LDR images with different exposures as shown in Figure 4.

Rendering of caustics is still a challenging task as it is typically not suitable for a real-time simulations and requires manual tuning of the parameters. Although, sig-

Fig. 4 Visible artefacts as a result of HDR bracketing used for a moving caustics [22]

nificant improvements have been made, this is still an unresolved issue [20]. Another important light phenomenon that is of significant importance in CH reconstruction and simulation is the participating media, typically in forms of a smoke, fog and dust particles [21, 34].

Another acquisition technique that might benefit from HDR is the reflectance transform imaging (RTI). This computational photographic method captures object's shape and colour from a single point of view but varying the light position and thus highlights and shadows on the surface in each image [32]. This eventually enables the interactive relighting of the object from any direction and computational enhancements of its surface and colour features, thus better analysis, understanding and interpretation of the artefacts. Therefore, given the importance of the image details in this technique, capturing higher dynamic range for highly specular surfaces might be necessary for preserving reliable pixel values.

Relighting virtually reconstructed sites and objects using captured light information at the original site is an important aspect of virtual reconstruction of CH. In this case a 3D rendering technique termed Image Based Lighting (IBL) can be utilised [11, 9]. This technique comprises of omnidirectional real-world light capture, typically a dome or a sphere, and the projection of this information onto a scene geometry, for simulating the lighting in a virtual scene. Although the light probe can be generated at the same physical place, it cannot be captured in the original time, i.e. in the past, as for example in the reconstruction of the Parthenon by Debevec [10]. Nonetheless, these probes, typically called environment maps can be used instead of standard synthetic light sources to achieve more accurate representation of light for the given scene [24]. In the CH context, it might be necessary to reconstruct the surrounding and distant objects and, at the same time, remove them from the captured light probe. In addition, all the light sources not belonging to the target time should be switched off. One of the main limitations of this method is the fact that each probe is captured at a single point in space, at a single point in time. This does not allow for recreating spatially and temporally varying lighting conditions. The solution for this would be a creation of a 4D HDR light field, typically called Incident Light Fields (ILFs) [56, 55]. Alternatively, the sky can be modelled or simulated [39, 28, 45]. One of the most commonly used sets of generic sky models with conditions from completely clear to overcast is the one of the International Commission on Illumination (CIE) [7].

6.2 Visualisation

When it comes to visualisation of CH material, the first thing that should be considered is whether it will be a passive or active experience. The former, can be rendered offline and requires no interaction or input from the user. It can be delivered with an extremely high level of detail utilising physically-based rendering techniques. The latter assumes certain level of user interaction and, although elements could be pre-rendered, relies on real-time rendering. Physically-based rendering is beyond the scope of this chapter, but the literature suggests that it is still not possible to deliver such imagery at very high resolutions and frame rates. The purpose of the (inter)active experiences is two-fold: for the research purposes by the experts, such as the documentation of the excavations or monitoring of the degradation of the artefacts; or for the enhancement and promotion of CH through virtual museums, serious games or other types of interactive presentations [50].

One of the best media for presentation of CH sites and artefacts are virtual museums (VM). Virtual museums are digital entities complementing, enhancing or augmenting the museum experience [v-must.net]. They are intrinsically immersive and interactive way of enhancing the understanding and connection with the past and present world around us. VMs can represent a digitised exhibition, section, department or a full physical museum, or can be an independent collection of artefacts acting as a separate, virtual entity. In addition, they can be multiplatform – standalone, mobile, web-based, VR/AR; have different physical setup and accessibility - running on site or remotely; and level of engagement - single or multi-user, game like experience [59, 31, 47, 38]. They could also be categorised based on various criteria: content, interaction technology, duration, communication, level of immersion, format, scope, sustainability [58].

One of the main goals of VMs is to communicate knowledge to a wide audience, whether they are general public, scholars or professionals. This is typically achieved or enhanced by using digital storytelling [23, 40, 46]. The stories can be about the artefacts or the site, contextualising them culturally, historically, geographically and through time, providing a wider picture of how they have been created, used, preserved, and changed over periods of time, thus augmenting the experience compared to a typical museum presentation.

6.2.1 Modes of delivery

As discussed previously, dynamic range of the original scene could be of significant importance when capturing the environment. Similarly, it might be necessary to display that same range on the display. Therefore, different platform and types of display devices will be discussed.

The most conventional platform for CH application is desktop, either as a standalone or web-based system. In either case, the display hardware is a monitor that might support touch-screen interaction with the application. Although HDR seems to be widely present in current technology, from TVs, computer monitors and mobile phones, there is still a lot of debate in the scientific community what “true” HDR is. The first HDR display is considered to by Brightside’s DR 37P HDR, Figure 5 (left).

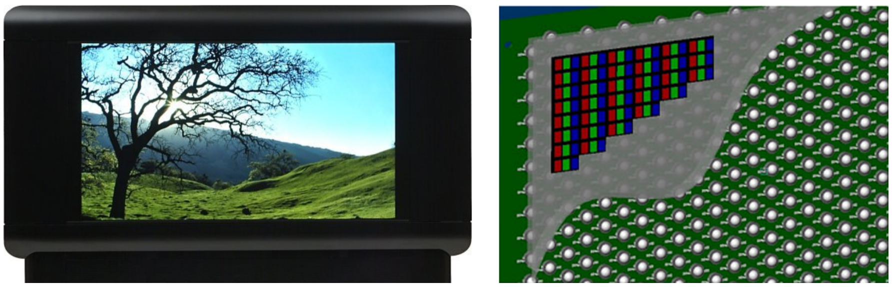


Fig. 5 Brightside DR 37P HDR display (left) and the LED array behind the LCD panel (right). Image courtesy of Brightside.

The company has later been acquired by Dolby and the technology patented as Dolby Vision [13]. This technology delivers more accurate, realistic colour and dynamic range compared to LDR technology and even its main rival HDR10 [51] by using 12-bit per colour channel support, wide colour gamut [26] with a peak brightness of 10000 nits. This technology allows much brighter and dimmer representations due to an LED-lit back panel, Figure 5 (right).

At the same time, SIM2 has developed their own technology and released three display models HDR47E3, HDR47ES4MB and HDR47ES6MB with 2500, 4500 and 6000 nits peak brightness respectively [48]. Nonetheless, this technology has still remained at the research and specialised domain, and could not truly penetrate the consumer market. This is partially due to the price of these units, but also lack of content and standardisation in the field. Instead, HDR10 technology, supporting 10-bit per channel, has been accepted as a de facto standard for HDR for the consumer market. There are many TV and computer screens supporting this “limited” HDR technology available at the moment. Although not as rich as “true” HDR, and introducing various artefacts, it is still a significant step forward from the conventional LDR technology which supports 8 bits per colour channel and has a narrower colour gamut (Rec. 709, BT 1886) [25].

When it comes to mobile devices, the market is more focused on better appearance rather than truthfulness of real luminance reproduction. Similarly to TV, the most prevalent HDR standard on high-end mobile devices is HDR10. There are a few examples of support for Dolby Vision or HDR10+, which increases the dynamic range of HDR10, but they suffer from similar limitations as TV devices with this technology. Only recently the Ultra HD Alliance, made up of movie producers and technology companies, proposed a new standard called ‘Ultra HD Premium’ with two HDR definitions, one for LED screens with the range of 0.05 – 1000 nits, and one for OLED displays with the range of 0.0005 – 540 nits. However, in most cases of mobile HDR, the decoding is done by software, rather than hardware. While OLED technology is promising for the mobile display industry, there are still challenges to be addressed.

Virtual Reality (VR) and Augmented Reality (AR) technologies have recently advanced and might have significant impact on our lives in the future. These technologies allow to immerse into another "reality" either as an augmentation of the physical environment (AR) with additional synthetic content, or as a separate computer-generated environment with any type of real of fictional content and various modes of interaction. Unfortunately, VR/AR hardware slightly lacks behind the computer and mobile screen technology. Only very recently new head-mounted displays (HMDs) have been released with reasonably good resolutions, e.g. HTC Vive Pro with 2880 x 1600 pixels (1400 x 1600 pixels per eye). However, increasing the dynamic range on these devices is something that is not currently considered, and is possibly not something that we might expect in the very near future. Recent work by Goudé et al. [18] proposes a TMO for 360° content viewed on HMDs that adjusts the content depending on the viewing angle, as well as a more globally consistent operator. Moving towards brighter screens in HMDs is not only a technological challenge but also raises many other considerations, such as health and safety, given the proximity of the displays to the eyes.

High dynamic range has proven a valuable and important tool for better representation and understanding of CH. If harnessed correctly with other stages throughout the pipeline it might create a great experience for both the general public and the academic community. Depending on the nature of the reconstruction and the system, different platforms could be considered. The fact that only desktop setups with dedicated display hardware can deliver "real" HDR experience is somewhat limiting. However, high-end mobile devices seem to be good alternatives for lower budgets and public displays. While VR and AR systems can make a significant user experience enhancement mainly through immersion and presence [49, 29], there are still a few limiting factors with available HMDs, such as price, safety, comfort, setup and maintenance, and the display technology.

7 Conclusions

High Dynamic Range imaging can provide solutions for a variety of Cultural Heritage applications due to its ability to capture, process, store and reproduce the real world at high fidelity. Disciplines related to CH, such as archaeology and architecture, can benefit greatly from using HDR, for example by accurately reproducing artefacts and structures in virtual environments or by capturing and documenting current traditions, buildings and artefacts.

Implementing HDR pipelines is challenging, with each stage presenting new obstacles. Capturing HDR images an videos is an open problem, given the limitations of our sensors and software. Displaying HDR content presents another challenge due to the limited capabilities of actual light reproduction of current displays, requiring tone mapping. Historical content is predominantly LDR, requiring dynamic range expansion for converting it to LDR, which is a problem of substantial difficulty.

It would however be great if we could accurately reproduce the Parthenon and see it in its original state, as it was thousands of years ago, in some virtual environment. That is however extremely hard due to the limited documentation and data availability. Our current heritage, captured with HDR technology, will be better preserved and then reproduced by future generations, giving a more complete picture of history and tradition.

References

[1] Akyüz, A.O., Fleming, R., Riecke, B.E., Reinhard, E., Bülthoff, H.H.: Do HDR displays support LDR content?: a psychophysical evaluation. ACM Transactions on Graphics (TOG) **26**(3), 38 (2007)

[2] Allen, N.P.: Is the Shroud of Turin the first recorded photograph? South African Journal of Art History **11**, 23–32 (1993)

[3] Banterle, F., Artusi, A., Debattista, K., Chalmers, A.: Advanced High Dynamic Range Imaging (2nd Edition). AK Peters (CRC Press), Natick, MA, USA (2017)

[4] Banterle, F., Ledda, P., Debattista, K., Chalmers, A.: Inverse tone mapping. Proceedings of GRAPHITE '06 p. 349 (2006). DOI 10.1145/1174429.1174489

[5] Boitard, R., Bouatouch, K., Cozot, R., Thoreau, D., Gruson, A.: Temporal coherency for video tone mapping. In: Applications of Digital Image Processing XXXV, vol. 8499, p. 84990D. International Society for Optics and Photonics (2012)

[6] Boitard, R., Cozot, R., Thoreau, D., Bouatouch, K.: Survey of temporal brightness artifacts in video tone mapping. In: HDRi2014-Second International Conference and SME Workshop on HDR imaging, vol. 9 (2014)

[7] CIE: Spatial distribution of daylight – cie standard general sky. vienna (austria). CIE Publication No. S 011/E (2003)

[8] Debattista, K., Bashford-Rogers, T., Selmanović, E., Mukherjee, R., Chalmers, A.: Optimal exposure compression for high dynamic range content. The Visual Computer **31**(6-8), 1089–1099 (2015)

[9] Debevec, P.: Rendering synthetic objects into real scenes: Bridging traditional and image-based graphics with global illumination and high dynamic range photography. In: ACM SIGGRAPH 2008 classes, p. 32. ACM (2008)

[10] Debevec, P., et al.: "making" the parthenon. In: 6th international symposium on virtual reality, archaeology, and cultural heritage, vol. 4 (2005)

[11] Debevec, P.E., Malik, J.: Recovering high dynamic range radiance maps from photographs. In: ACM SIGGRAPH 2008 classes, p. 31. ACM (2008)

[12] Didyk, P., Mantiuk, R.K., Hein, M., Seidel, H.P.: Enhancement of Bright Video Features for HDR Displays. Computer Graphics Forum **27**(4), 1265–1274 (2008)

[13] Dolby: Dolby vision: White paper (2016)
[14] Drago, F., Myszkowski, K., Annen, T., Chiba, N.: Adaptive Logarithmic Mapping For Displaying High Contrast Scenes. Computer Graphics Forum **22**(3), 419–426 (2003). DOI 10.1111/1467-8659.00689
[15] Eilertsen, G., Kronander, J., Denes, G., Mantiuk, R.K., Unger, J.: HDR image reconstruction from a single exposure using deep CNNs. ACM Transactions on Graphics (TOG) **36**(6), 178 (2017)
[16] Eilertsen, G., Mantiuk, R.K., Unger, J.: A comparative review of tone-mapping algorithms for high dynamic range video. Computer Graphics Forum **36**(2), 565–592 (2017). DOI 10.1111/cgf.13148
[17] Endo, Y., Kanamori, Y., Mitani, J.: Deep reverse tone mapping. ACM Transactions on Graphics (TOG) **36**(6), 171–177 (2017)
[18] Goudé, I., Cozot, R., Banterle, F.: Hmd-tmo: A tone mapping operator for 360° hdr images visualization for head mounted displays. In: Computer Graphics International Conference, pp. 216–227. Springer (2019)
[19] Greg Ward: Real Pixels. Graphics Gems II pp. 80–83 (1991)
[20] Grittmann, P., Pérard-Gayot, A., Slusallek, P., Křivánek, J.: Efficient caustic rendering with lightweight photon mapping. In: Computer Graphics Forum, vol. 37, pp. 133–142. Wiley Online Library (2018)
[21] Gutierrez, D., Sundstedt, V., Gomez, F., Chalmers, A.: Modeling light scattering for virtual heritage. Journal on Computing and Cultural Heritage (JOCCH) **1**(2), 8 (2008)
[22] Happa, J., Artusi, A., Czanner, S., Chalmers, A.: High dynamic range video for cultural heritage documentation and experimental archaeology. In: Proceedings of the 11th International conference on Virtual Reality, Archaeology and Cultural Heritage, pp. 17–24. Eurographics Association (2010)
[23] Hulusic, V., Rizvic, S.: Story guided virtual environments in educational applications. In: Transactions on Edutainment IX, pp. 132–149. Springer (2013)
[24] Inanici, M.: Evalution of high dynamic range image-based sky models in lighting simulation. Leukos **7**(2), 69–84 (2010)
[25] ITU-R: Parameter values for the hdtv standards for production and international programme exchange. ITU-R Recommendation BT.709-6 (2015)
[26] ITU-R: Parameter values for ultra-high definition television systems for production and international programme exchange. ITU-R Recommendation BT.2020-2 (2015)
[27] Jakubiec, J.A., Van Den Wymelenberg, K., Inanici, M., Mahic, A.: Improving the accuracy of measurements in daylit interior scenes using high dynamic range photography. In: Proceedings of the 32nd PLEA Conference, Los Angeles, CA, USA, pp. 11–13 (2016)
[28] Jensen, H.W., Durand, F., Dorsey, J., Stark, M.M., Shirley, P., Premože, S.: A physically-based night sky model. In: Proceedings of the 28th annual conference on Computer graphics and interactive techniques, pp. 399–408. ACM (2001)

[29] Jung, T., tom Dieck, M.C., Lee, H., Chung, N.: Effects of virtual reality and augmented reality on visitor experiences in museum. In: Information and communication technologies in tourism 2016, pp. 621–635. Springer (2016)
[30] Landis, H.: Production-ready global illumination. spherevfx.com (2002)
[31] Louvre-Lens-Multimedia: https://www.louvrelens.fr/informations-pratiques/tarifs-abonnement/ (2019)
[32] Malzbender, T., Gelb, D., Wolters, H.: Polynomial texture maps. In: Proceedings of the 28th annual conference on Computer graphics and interactive techniques, pp. 519–528. ACM (2001)
[33] Mantiuk, R.K., Daly, S.J., Kerofsky, L.: Display adaptive tone mapping. ACM Transactions on Graphics (TOG) **27**(3), 1 (2008). DOI 10.1145/1360612.1360667
[34] Marco, J., Guillén, I., Jarosz, W., Gutierrez, D., Jarabo, A.: Progressive transient photon beams. Computer Graphics Forum (2019)
[35] Marnerides, D., Bashford-Rogers, T., Hatchett, J., Debattista, K.: ExpandNet: A Deep Convolutional Neural Network for High Dynamic Range Expansion from Low Dynamic Range Content. Computer Graphics Forum **37**(2), 37–49 (2018). DOI 10.1111/cgf.13340
[36] Meylan, L.: Tone mapping for high dynamic range images. Tech. rep., EPFL (2006)
[37] and accurate Meylan, L., Daly, S.J., Süsstrunk, S.: The Reproduction of Specular Highlights on High Dynamic Range Displays. In: 14th Color Imaging Conference, pp. 333–338. Scottsdale, AZ, USA (2006)
[38] NHM: http://naturalhistory.si.edu/vt3/ (2019)
[39] Perez, R., Seals, R., Michalsky, J.: All-weather model for sky luminance distribution—preliminary configuration and validation. Solar energy **50**(3), 235–245 (1993)
[40] Pietroni, E., Forlani, M., Rufa, C.: Livia's villa reloaded: An example of re-use and update of a pre-existing virtual museum, following a novel approach in storytelling inside virtual reality environments. In: 2015 Digital Heritage, vol. 2, pp. 511–518. IEEE (2015)
[41] Reinhard, E., Devlin, K.: Dynamic range reduction inspired by photoreceptor physiology. IEEE Transactions on Visualization and Computer Graphics **11**(1), 13–24 (2005)
[42] Reinhard, E., Khan, E.A., Akyüz, A.O., Johnson, G.: Color imaging: fundamentals and applications. AK Peters/CRC Press (2008)
[43] Reinhard, E., Stark, M., Shirley, P., Ferwerda, J.: Photographic Tone Reproduction for Digital Images. ACM Transactions on Graphics (TOG) **21**(3), 267–276 (2002). DOI 10.1145/566654.566575
[44] Reinhard, E., Ward, G., Pattanaik, S., Debevec, P.: High Dynamic Range Imaging: Acquisition, Display and Image-based Lighting, second edn. Elsevier (2010)
[45] Satỳlmỳs, P., Bashford-Rogers, T., Chalmers, A., Debattista, K.: A machine-learning-driven sky model. IEEE computer graphics and applications **37**(1), 80–91 (2016)

[46] Selmanovic, E., Rizvic, S., Harvey, C., Boskovic, D., Hulusic, V., Chahin, M., Sljivo, S.: VR Video Storytelling for Intangible Cultural Heritage Preservation. In: R. Sablatnig, M. Wimmer (eds.) Eurographics Workshop on Graphics and Cultural Heritage. The Eurographics Association (2018). DOI 10.2312/gch.20181341

[47] SGG: `http://h.etf.unsa.ba/srp/` (2019)

[48] SIM2: `http://hdr.sim2.it/productslist` (2019)

[49] Slater, M., Linakis, V., Usoh, M., Kooper, R.: Immersion, presence and performance in virtual environments: An experiment with tri-dimensional chess. In: Proceedings of the ACM symposium on virtual reality software and technology, pp. 163–172. ACM (1996)

[50] Stanco, F., Tanasi, D.: Experiencing the past: computer graphics in archaeology. Digital Imaging for Cultural Heritage pp. 1–37 (2011)

[51] Standard, S.: High dynamic range electro-optical transfer function of mastering reference displays. SMPTE ST **2084**, 1–14 (2014)

[52] Tanselle, G.T., et al.: Literature and artifacts. Bibliographical Society of the University of Virginia (1998)

[53] Tursun, O.T., Akyüz, A.O., Erdem, A., Erdem, E.: The State of the Art in HDR Deghosting: A Survey and Evaluation. Computer Graphics Forum **34**(2), 683–707 (2015). DOI 10.1111/cgf.12593

[54] Unesco: `http://www.unesco.org` (2019)

[55] Unger, J.: Incident light fields. Dissertation no. 1233, Linköping Studies in Science and Technology (2009)

[56] Unger, J., Gustavson, S., Larsson, P., Ynnerman, A.: Free form incident light fields. Computer Graphics Forum (Proceedings of EGSR 2008) **27**(4), 1293–1301 (2008)

[57] Urbano, C., Magalhães, L., Moura, J., Bessa, M., Marcos, A., Chalmers, A.: Tone mapping operators on small screen devices: An evaluation study. Computer Graphics Forum **29**(8), 2469–2478 (2010). DOI 10.1111/j.1467-8659.2010.01758.x

[58] V-Must.net: http://v-must.net (2019)

[59] VaticanMuseum: `http://www.museivaticani.va/content/museivaticani/en/collezioni/musei.html` (2019)

[60] Ward, G.: Fast, Robust Image Registration for Compositing High Dynamic Range Photographs from Hand-Held Exposures. Journal of Graphics Tools **8**(2), 17–30 (2003). DOI 10.1080/10867651.2003.10487583