High Dynamic Range Imaging and Low Dynamic Range Expansion for Generating HDR Content

Francesco Banterle [1] Kurt Debattista [1] Alessandro Artusi [1] Sumanta Pattanaik [2] Karol Myszkowski [3]

Patrick Ledda [1] Marina Bloj [4] and Alan Chalmers [1]

[1] The Digital Laboratory, WMG, University of Warwick
[2] University of Central Florida
[3] Max-Planck-Institut für Informatik
[4] Optometry Department, University of Bradford

Abstract

In the last few years, researchers in the field of High Dynamic Range (HDR) Imaging have focused on providing tools for expanding Low Dynamic Range (LDR) content for the generation of HDR images due to the growing popularity of HDR in applications, such as photography and rendering via Image-Based Lighting, and the imminent arrival of HDR displays to the consumer market. LDR content expansion is required due to the lack of fast and reliable consumer level HDR capture for still images and videos. Furthermore, LDR content expansion, will allow the re-use of legacy LDR stills, videos and LDR applications created, over the last century and more, to be widely available. The use of certain LDR expansion methods, those that are based on the inversion of tone mapping operators, has made it possible to create novel compression algorithms that tackle the problem of the size of HDR content storage, which remains one of the major obstacles to be overcome for the adoption of HDR. These methods are used in conjunction with traditional LDR compression methods and can evolve accordingly. The goal of this report is to provide a comprehensive overview on HDR Imaging, and an in depth review on these emerging topics.

1. Introduction

High Dynamic Range (HDR) Imaging has become one of the main areas of computer graphics. One major aspect of HDR imaging which is bound to become extremely relevant is the aspect of providing content for HDR displays. While content can be captured directly for HDR displays [DM97], this is typically not a straightforward process and may require specialised equipment to automate [Sph02, Pan02], just to obtain still images. The provision of animated HDR content is still in its infancy and few reliable methods exist to directly capture HDR video [Hoe07]. This has led to research into providing HDR content from Low Dynamic Range (LDR) originals. Such work makes it now possible to re-use the large amount of already existing legacy LDR in a way that makes full use of emerging HDR displays. Furthermore, several methods, based on LDR to HDR expansion, have been used for HDR compression and for enhancing the quality of rendered images based on HDR image-based lighting (IBL). While previous techniques dealing with general HDR methods have been collected and published, for example [RWPD05], at this time only a short survey without in-depth discussion and analysis of algorithmic aspects of LDR expansion techniques has been published by Myszkowski et al. [MMK08]. Myszkowski et al.'s work does not cover the association between LDR to HDR expansion and HDR compression techniques as is presented in this survey.

We begin this state-of-the-art report by giving a quick overview of the different aspects of HDR imaging. In Section 3 we present methods that expand LDR into HDR content with respect to still images, videos and the use of expansion in applications such as IBL. We classify these methods and present the work that has been done to validate such techniques. In Section 4 we show how LDR to HDR expansion methods have been used to compress HDR content, by taking advantage of already existing LDR compression

schemes. Finally, we conclude by discussing open problems and present future directions.

2. High Dynamic Range Imaging

The introduction of HDR imaging in the last two decades by the computer graphics community has revolutionised the field and other areas such as photography, virtual reality, visual effects, and the video-games industry. Physically-correct light values can now be captured and fully utilised for various applications without the need to linearise the signal and to deal with clamped values. The very dark and bright areas of a scene can be recorded at the same time into an image or a video, avoiding under-exposed and over-exposed areas. Traditional imaging methods do not use physical values and typically are constrained by limitations in technology that could only handle 8-bit per colour channel per pixel. Such imagery (8-bit or less per colour channel) is referred as LDR imagery. This change in how light can be recorded is comparable to the introduction of colour photography and has changed each stage of the imaging pipeline, see Figure 1. The four main stages are: capturing, storing, processing, and displaying.

2.1. Capturing

Currently, available consumer cameras are limited to capture only 8-bit images or 12-bit images in RAW format, which do not cover the full dynamic range of irradiance values in most environments in real world. The only possibility is to take a number of exposures of the same scene to capture details from very dark regions to very bright regions as proposed by Mann and Picard [MP95]. The problem with film and digital cameras is that they do not have a linear response, but a more general function h, called camera response function (CRF). Mann and Picard [MP95] proposed a simple method for calculating h, which consists of fitting the values of pixels at different exposures to a fixed CRF, $h(x) = ax^{\gamma} + b$. This parametric h is very limited and does not support most real CRFs. Debevec and Malik [DM97] proposed a simple method for recovering a CRF through a tabled h which is minimised using a squared error function. Mitsunaga and Nayar [MN99] improved this algorithm with a more robust method based on a polynomial representation of h. Note that the multiple exposure method assumes that images are perfected aligned, there are no moving objects, and CCD noise is not a problem. Robertson et al. [RBS99, RBS03] improved previous techniques for assembling HDR images from multiple exposures. They proposed an iterative calculation of the CRF in conjunction with a probabilistic weighting for merging different exposures.

Kang et al. [KUWS03] extended multiple exposure images methods for videos. They presented a system that had a programmed video-camera that temporally varies the shutter speed at each frame. The final video is generated aligning and warping corresponding frames at different shutter speeds and compositing them to recover the HDR one. However, the frame rate of this method is low, around 15 fps, and the scene has to contain low speed moving objects otherwise artifacts will appear. Nayar and Branzoi [NB03] developed an adaptive dynamic range camera, where a controllable liquid crystal light modulator is placed in front of the camera. This modulator adapts the exposure of each pixel on the image detector allowing to capture scenes with a very large dynamic range.

In the commercial field, few companies provide HDR cameras based on automatic multiple exposure capturing. The two main cameras are Spheron HDR VR camera [Sph02] by SpheronVR GmbH and Panoscan MK-3 [Pan02] by Panoscan Ltd , which are both full 360 degree panoramic cameras at high resolution. The two cameras capture full HDR images. For example, Spheron HDR VR can capture 26 f-stops of dynamic range at 50 Megapixels resolution in 24 minutes.

The alternative to automatic multiple exposure cameras is to use CCD sensors which can natively capture HDR values. In recent years, CCDs that record into 10/12-bit per channel in the logarithmic domain have been introduced by many companies such as Cypress Semiconductor [Cyp], Omron [Omr07], PTGrey [PtG04], Neuricam [Neu06], etc. The main problem with these sensors is that their resolution is low, VGA (640×480), and are noisy.Therefore, their applications are oriented to automotive, security, and automatisation use in factories.

In the cinema industry a few companies have proposed high quality solutions such as Viper camera [Tho05] by Thomson GV, Red One camera [Red06] by RED Company, and the Phantom HD camera [Vis05] by Vision Research, etc. All these video-cameras present high frame rates, low noise, full high definition ($1,920 \times 1,080$) resolution, and a good dynamic range (reaching the range of celluloid film), 10/12-bit per channel in the logarithmic/linear domain. However, they are extremely expensive (sometimes available only for renting) and they do not encompass the full dynamic range of the Human Visual System (HVS).

2.2. Storing

Once HDR images/videos are captured from the real world, or are synthesised using computer graphics, there is the need to store, distribute, and process these images. An uncompressed HDR pixel is represented using three single precision floating point numbers [Hou81], assuming three bands as for RGB colours. This means that a pixel uses 96 bits per pixel (bpp). Researchers have been working on compression methods to address the high demand on memory storage required for HDR content.

The early compression methods proposed an efficient and compact representation of floating point numbers, the main

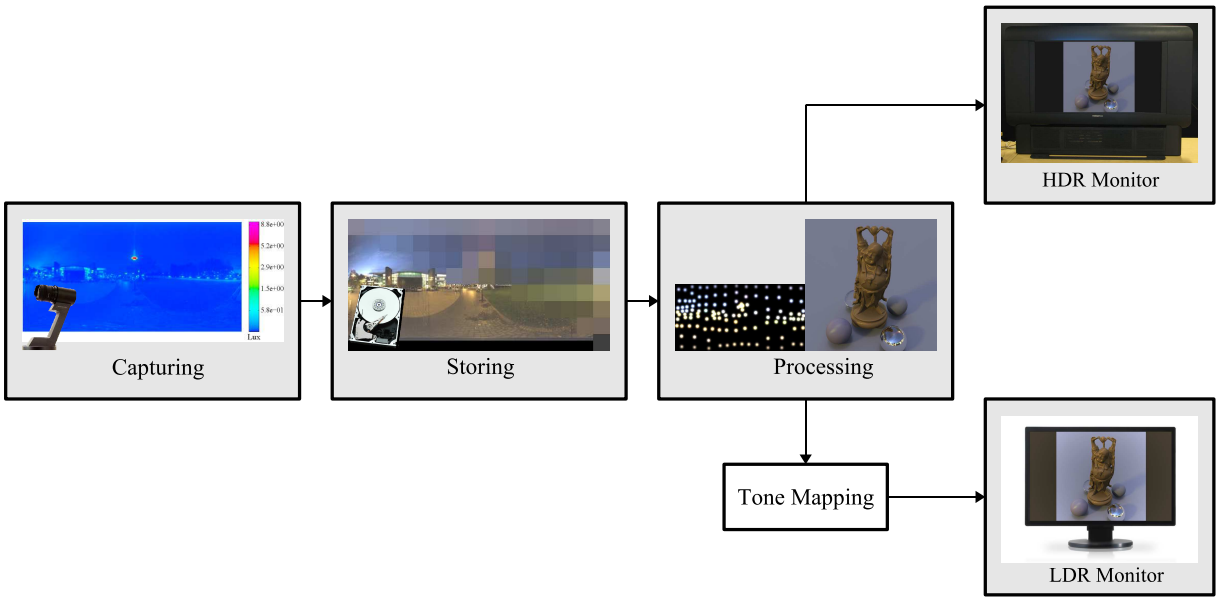

Figure 1: *The HDR pipeline in all its stages. Multiple exposure images are captured and combined obtaining an HDR image. Then this image is quantised, compressed, and stored on the hard disk. Further processing can be applied to the image. For example, areas of high luminance can be extracted and used to re-light a synthetic object. Finally, the HDR image or processed ones can be visualised using traditional LDR display technologies or native HDR monitors.*

formats are: RGBE/XYZE, LogLuv, and half precision numbers. RGBE/XYZE [War91] is an implementation of floating point where the exponent of the floating point is shared between RGB or XYZ values assuming the exponents have a similar magnitude. For this method the storage requirement is 32 bpp. LogLuv method [Lar98] proposed to separate luminance and chrominance where luminance is encoded in the logarithmic domain achieving 24/32 bpp. Finally, the OpenEXR file format [Ind02] proposed the use of half precision numbers, a 16-bit version of IEEE-754 standard [Hou81], maintaining the dynamic range in an extremely high quality.

LDR Imaging compression methods have also been extended to HDR. For example, block truncation methods in separate luminance and chrominance colour spaces [MCHAM06, RAI06, SLWL08] were applied to HDR textures, achieving 8 bpp. Moreover, Wang et al. [WWS*07] proposed a separated encoding for the HDR and LDR parts of a texture at 16 bpp. Other methods that exploit tone mapping, inverse tone mapping, and LDR de facto standards. They are reviewed in Section 4.

2.3. Image-Based Lighting

HDR content can simplify the definition and rendering process for re-lighting synthesised objects. In particular, IBL techniques are aimed at simulating light transport, defining light sources and surrounding environment. Blinn and Newell [BN76] first used IBL to simulate the optical effects such as reflection and refraction. This was extended by Miller and Hoffman [MH84] and Green [Gre86] for simulating diffuse effects by convolving environment maps. However, these methods were limited to pure diffuse or pure specular materials, without taking into account visibility or secondary rays. Debevec [Deb98] generalised these techniques using ray tracing framework applying the Rendering Equation [Kaj86]:

$$L(\mathbf{x},\omega) = L_e + \int_\Omega L(\omega') f_r(\omega',\omega) V(\mathbf{x},\omega') \mathbf{n}\cdot\omega' d\omega'$$

where $\mathbf{x}$ and $\mathbf{n}$ are respectively the position and normal of the hit object, L_e is the emitted radiance at point $\mathbf{x}$, L is the environment map, f_r is the BRDF, ω' is the out-going direction, and ω is the view vector. V is the visibility function, a Boolean function that determines if a ray is obstructed by an object or not. This technique can be applied to real-world objects or human beings for re-lighting them using HDR content [DHT*00]. Therefore, re-lighting using HDR images/videos is a very important application in many fields such as augmented reality, visual effects, and computer graphics. This is because the appearance of the image is transferred onto the re-lighted objects.

2.4. Displaying

Most of the display devices commercially available nowadays are not able to display HDR content. This is because current display technology has a low contrast ratio around $1,000:1$ and can process only 8/10-bit images for each colour channel. In the last two decades researchers spent significant efforts to compress the range of HDR images and videos in order to display them on LDR displays. Tone mapping is the operation that reduces the dynamic range of the input content to fit the dynamic range of the display technology. One of the important requirement is that this reduction of the range must preserve some properties of the original content such as local and global contrast, opponent brightness, etc. Tone mapping is performed using an operator, f, which is referred to as a Tone Mapping Operator (TMO). TMOs can be classified in different groups based on the underlying image processing techniques. The main groups of the taxonomy are: Global Operators (the same mapping, f, is applied to all pixels), Local Operators (the mapping, f of a pixel depends on its neighbour pixels), Segmentation Operators (the image is segmented in broad regions, and a different mapping is applied to each region), Frequency/Gradient Operators (low and high frequencies of the images are separated, while an operator is applied to the low frequencies, high frequencies are usually retained to preserve fine details), Perceptual Operators (f models some aspects of the HVS), Empirical Operators (f tries to create pleasant images inspired by other fields such as photography), and Temporal Operators (suitable for tone mapping HDR videos). For an in depth review on tone mapping see Reinhard et al. [RWPD05].

Only in the last few years, researches have been working on display technologies for a native visualisation of HDR images and videos without using TMOs. The two main devices are the HDR viewer [SHS*04] and the HDR Monitor [SHS*04]. Internally, both of these devices divide an HDR image into a detail layer with colours and a luminance layer that back-modulates the first one.

3. LDR to HDR Expansion

The capture of HDR via multiple exposures using a traditional camera is a very time consuming task, and on a movie set the time for capturing images is very limited. Moreover, HDR cameras, such as Spheron [Sph02] and Panoscan [Pan02], are currently quite expensive, limiting their availability to only a few customers.

In this section we cover the publications that recreate HDR images and videos from captured LDR content. This is an ill-posed problem since the information needed to generate full HDR content, that is, the information in over-exposed and under-exposed regions of the image/frame, is missing.

3.1. Decontouring Models

Daly and Feng [DF03, DF04] proposed a couple of methods for extending the bit-depth of classic 8-bit images and videos (effectively 6-bit due to MPEG-2 compression) for 10-bit monitors. New LCD monitors present higher contrast, typically around 1,000:1, and a high luminance peak, that usually is around 400cd/m^2. This means that displaying 8-bit data, without any refinement, would entail having the content linear expanded for higher contrast resulting in artifacts such as banding/contouring. The goal of their methods is to create a medium dynamic range image, removing contouring in the transition areas, without particular emphasis on over-exposed and under-exposed areas.

3.1.1. Amplitude Dithering for High Contrast Displays

The first algorithm, Daly and Feng [DF03], is based on amplitude dithering by Roberts [Rob62], see Figure 2. Amplitude dithering or noise modulation is a dithering technique which simply adds a noise pattern to an image before quantisation. This noise pattern is removed when the image needs to be visualised. The bit depth is perceived higher than the real one, because there is a subsequent averaging happening in the display and in the HVS. Roberts' technique was modified to apply it to high contrast displays by Daly and Feng. Subtractive noise was employed instead of additive, since during visualisation a monitor can not remove it. The authors modelled the noise combining the effect of fixed pattern display noise and the one perceived by the HVS, making the noise invisible. They used the contrast sensitivity function (CSF) which is a 2D and anisotropic function derived by psychophysical experiments [Dal93]. The CSF is extended in the temporal dimension [Wat86] for moving images, which allows the noise to have a higher variance, and furthermore, they show that the range can be extended by an extra bit.

3.1.2. Contouring Removal

The second algorithm, Daly and Feng [DF04], presents a different approach where contours are removed instead of being masked with invisible noise. The first step of the algorithm is to filter the starting image at p bit using a low-pass filter, see Figure 3. The filter needs to be wide enough to span across the false contours. Note that this operation increases the bit depth to $n > p$ because during averaging a higher precision is needed than the one for the original values. Then this image is quantised at p bit, where any contour that appears are false ones, because the image has no high frequencies. Subsequently, the false contours are subtracted from the original image, and the filtered image at p bit is added to restore low frequency components. The main limitation of the algorithm is that it does not remove artifacts at high frequencies, but they are hard to detect by HVS due to frequency masking [FSPG97].

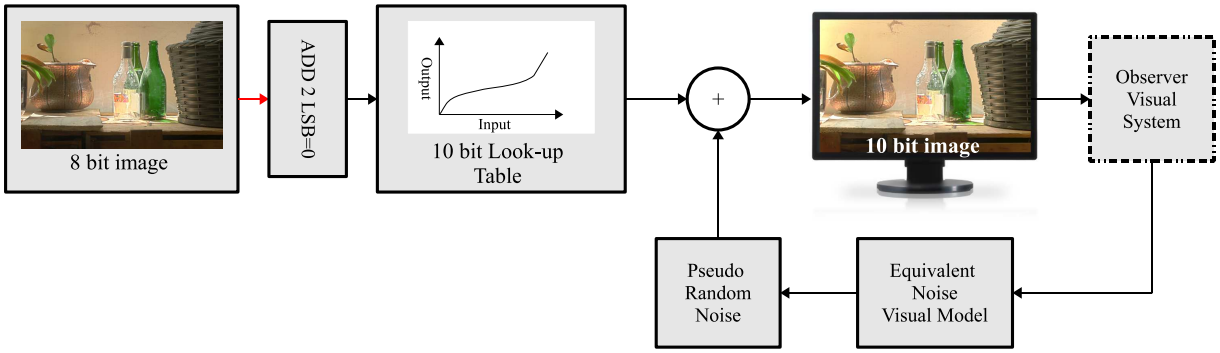

Figure 2: *The pipeline for bit depth extension using amplitude dithering by Daly and Feng [DF03].*

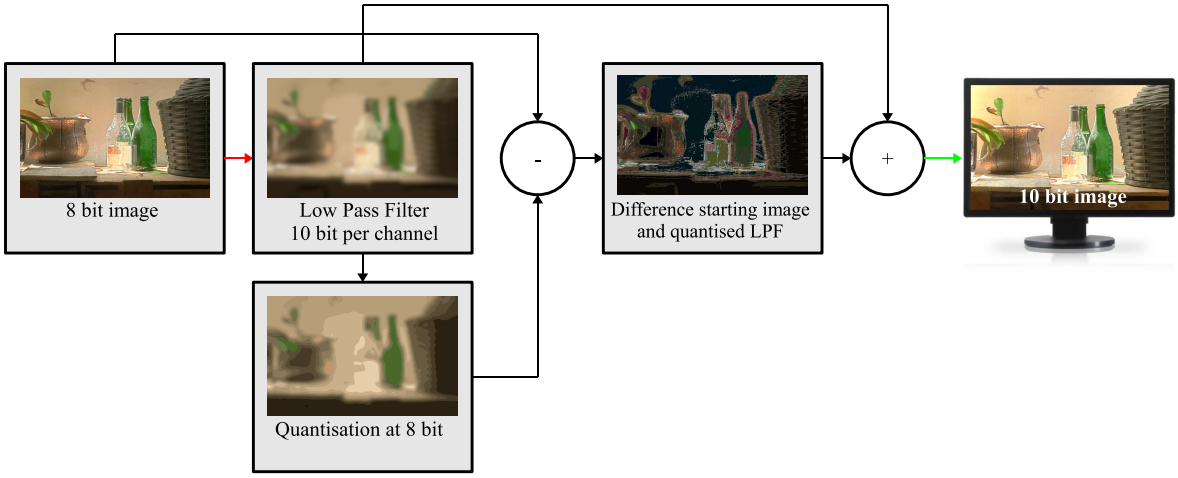

Figure 3: *The pipeline for bit depth extension using de-contouring by Daly and Feng [DF04].*

3.2. Global Models

Global models are those methods that apply the same single global expansion function on the LDR content at each pixel in the entire image.

3.2.1. A Power Function Model for Range Expansion

One of the first expansion methods was proposed by Landis [Lan02]. This expansion method, used primarily for IBL, is based on power functions. The luminance expansion is defined as:

$$L_{\mathrm{w}}(\mathbf{x}) = \begin{cases} (1-k)L_{\mathrm{d}}(\mathbf{x}) + kL_{\mathrm{w,\,Max}}L_{\mathrm{d}}(\mathbf{x}) & \text{if } L_{\mathrm{d}}(\mathbf{x}) \geq R \\ L_{\mathrm{d}}(\mathbf{x}) & \text{otherwise} \end{cases}$$

$$k = \left(\frac{L_{\mathrm{d}}(\mathbf{x}) - R}{1-R}\right)^{\alpha}$$

where R is the threshold for expansion which is equal to 0.5 in the original work, $L_{\mathrm{w,\,Max}}$ is the maximum luminance which the user needs for the expanded image, and α is the exponent of fall-off that controls the stretching curve.

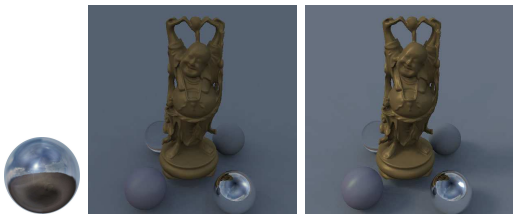

Figure 4: *An example of IBL using Landis' operator: a) The starting LDR lightprobe. b) The Happy Buddha is re-lighted using the image in a). c) The Happy Buddha is re-lighted using expanded lightprobe in a). Note that directional shadows from the sun are now present.*

While this technique produces suitable HDR light-probes

for IBL, see Figure 4, it may not produce good quality images/videos that can be visualised on HDR monitors. This is due to the fact that it does not handle artifacts such as exaggeration of compression or quantisation artifacts.

3.2.2. Linear Scaling for HDR Monitors

In order to investigate how well LDR content is supported by HDR displays, Akyüz et al. [AFR*07], run a series of psychophysical experiments. The experiments were run to evaluate tone mapped images, single exposure images and HDR images using the DR-37P HDR monitor. The experiment involved 22 naïve participants between 20 and 40 years old, and in all experiments 10 HDR images ranging from outdoor to indoor, from dim to very bright light conditions were used. The HDR images had around 5 orders of magnitude in order to be mapped to the DR-37P Dolby HDR monitor [Dol05].

The first experiment was a comparison between HDR and LDR images produced using various TMOs [LRP97, DD02, RSSF02], an automatic exposure (that minimises the number of over/under-exposed pixels), and an exposure chosen by subjects in a pilot study. Images were displayed on the DR-37P, using calibrated HDR images and LDR images calibrated to match the appearance on a Dell UltraSharp 2007FP 20.1" LCD monitor. Subjects had the task of ranking images which were looking best to them. For each original test image a subject had to watch a trial image for 2 seconds which was randomly picked between the different type of images. The experimental results showed that participants preferred HDR images. The authors did not find a large difference in participant preference between tone mapped and single exposure (automatic and chosen by the pilot)images.

In the second experiment the authors compared expanded single exposure with HDR and single exposure images (automatic and chosen by the pilot). To expand the single exposure images, they employed the following expansion method:

$$L_{\mathrm{w}}(\mathbf{x}) = k\left(\frac{L_{\mathrm{d}}(\mathbf{x}) - L_{\mathrm{d,\,Min}}}{L_{\mathrm{d,\,Max}} - L_{\mathrm{d,\,Min}}}\right)^{\gamma}$$

where k is the maximum luminance intensity of the HDR display, and γ is the non-linear scaling factor. For this experiment images with different γ values equal to 1, 2.2 and 0.45 were generated. The setup and the ranking task was the same as the first experiment. The results showed that brighter chosen exposure expanded images were preferred to HDR images, and vice versa when they had the same mean luminance. Authors suggested that mean luminance is preferable to contrast. Finally, another important result is that linear scaling, $\gamma = 1$, was the most favoured expansion, suggesting that a linear scaling may be enough for an HDR experience.

The authors worked only with high resolution HDR images, without compression artifacts, and artistically captured. While this works well under such ideal conditions, in more realistic scenarios, such as television programmes or DVDs, where compression is employed, this may not always be the case. In these cases a more accurate expansion needs to be done in order to avoid amplification of compression artifacts, and contouring.

3.3. Classification Models

The methods of Meylan et al. [MDS06, MDS07] and Didyk et al. [DMHS08] attempt to expand different aspects of the LDR content by identifying or classifying different parts in the image such as highlights and light sources.

3.3.1. Highlight Generation for HDR Monitors

Meylan et al. [MDS06, MDS07] presented an inverse Tone Mapping Operator (iTMO) with the specific task of representing highlights in LDR images when displayed on HDR monitors. The main idea is to detect the diffuse and specular part of the image and to expand these using different linear functions. The detection is based on the assumption that highlights are small and bright, which means that the maximum diffuse luminance value ω is obtained as the maximum of the low-pass filtered luminance channel L_{d} of the image. However, more processing is needed to avoid the case when white diffuse regions are next to regions with highlights, see Figure 5 for the complete pipeline for calculating ω.

After the calculation of ω, the luminance channel is expanded using the following function:

$$f(L_{\mathrm{w}}(\mathbf{x})) = \begin{cases} s_1 L_{\mathrm{d}}(\mathbf{x}) & \text{if } L_{\mathrm{d}}(\mathbf{x}) \leq \omega \\ s_1 \omega + s_2 (L_{\mathrm{d}}(\mathbf{x}) - \omega) & \text{otherwise} \end{cases}$$
$$s_1 = \frac{\rho}{\omega} \quad s_2 = \frac{1-\rho}{L_{\mathrm{d,\,Max}} - \omega} \qquad (1)$$

where $L_{\mathrm{d,\,Max}} = 1$ since the image is normalised, and ρ is the percentage of the HDR display luminance allocated to the diffuse part. However, a global application of f can produce quantisation artifacts around the enhanced highlights. This is reduced using a low pass filter only in the expanded regions, see Figure 6.

Finally, they ran a series of psychophysical experiments to determine the value of ρ for f using the DR-37P Dolby HDR monitor [Dol05]. The results showed that for outdoor scenes users preferred a high value of ρ, which means a small percentage of dynamic range allocated to highlights, while for indoor scenes this was the contrary. For indoor and outdoor scenes of equal diffuse brightness users chose a low value for ρ, so they preferred more range allocated to highlights. In conclusion from the analysis of the data, $\rho = 0.66$ is a good general estimate.

This algorithm is designed for a specific task, the reproduction of highlights on HDR monitors. The use for other

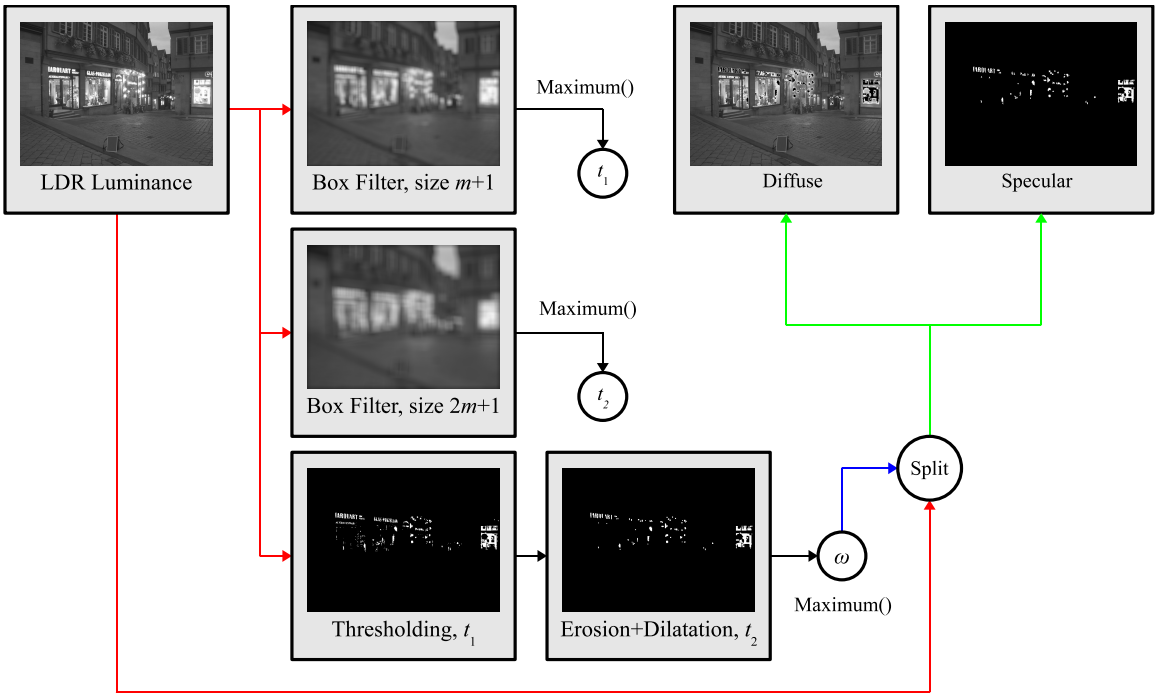

Figure 5: *The pipeline for the calculation of the maximum diffuse luminance value ω in an image in Meylan et al. [MDS07]. The image is firstly filtered using a box filter of size m to calculate value t_1 as the maximum of the luminance value. This operation is repeated for a filter of size $2m+1$ to calculate t_2. Then, t_1 is used as threshold on the original luminance resulting in a mask, subsequently an erosion and dilation filter is applied to the mask using t_2. While pixels active in the mask are considered specular pixels, black pixels are considered as diffuse ones.*

tasks, such as enhancement of videos, needs more processing and a classifier, which was underlined by authors' evaluation experiment.

3.3.2. Enhancement of Bright Video Features for HDR Display

Didyk et al. [DMHS08] proposed an interactive system for enhancing brightness of LDR videos, targeting and showing results for DVD content. The main idea of the system is to classify a scene into three components: diffuse, reflections, and light sources, and then to enhance only reflections and light sources. The authors explained that diffuse components are difficult to enhance without creating visual artifacts and it was probably the intention of film-makers to show them saturated as opposed to light sources and clipped reflections. The system works on non-linear values, because the goal is the enhancement and non-physical accuracy.

The system consists of three main parts: pre-processing, classification, and enhancement of clipped regions, see Figure 7 for the pipeline. The pre-processing step generates data needed during the classification. In particular, it determines clipped regions using a flood-fill algorithm. At least one channel must be saturated (over 230 for DVD content), and luma values must be greater than 222. Also, in this stage optical flow is calculated as well as other features such as image statistics, geometric features and neighbourhood characteristics.

Classification determines lights, reflections, and diffuse regions in a frame and relies on a training set of 2,000 manually classified regions. Primarily, a support vector machine [Vap95] with kernel $k(z,z') = \exp(-\gamma\|z-z'\|^2)$ performs an initial classification of regions. Subsequently, motion tracking improves the initial estimation, using a nearest neighbour classifier based on an Euclidean metric:

$$d^2((z,x,t),(z',x',t')) = 50\|z-z'\|^2 + \|z-z'\|^2 + 5(t-t')^2$$

where z are region features, x are coordinates in the image, and t is the frame number. This is allowed to reach a classification error of 12.6% on all regions used in the tests. Tracking of clipped regions using motion compensation further reduced the percentage of objects that require manual correction to 3%. Finally, the user can supervise classified regions, correcting wrong classifications using an intuitive user interface, see Figure 8.

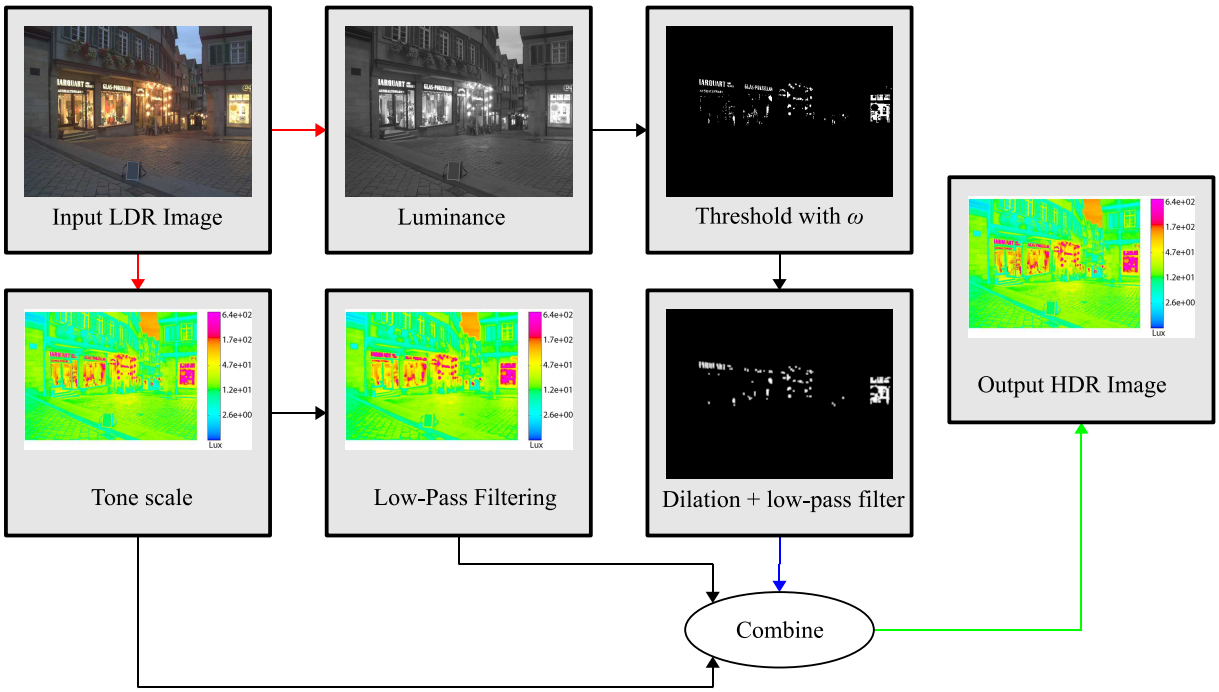

Figure 6: *The pipeline for the range expansion in Meylan et al.'s method [MDS07]. The original LDR image is expanded using Equation 1. Then, expanded luminance is filtered using a low pass filter. Finally, filtered expanded luminance and unfiltered one are linearly interpolated using a mask. This mask is calculated by thresholding LDR luminance with ω. To remove noise, the mask is filtered with a dilatation and low pass filter.*

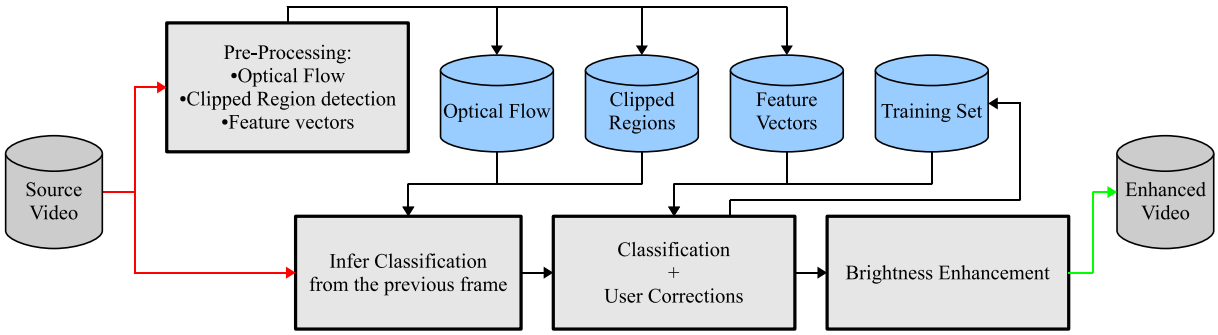

Figure 7: *The pipeline of the system proposed by Didyk et al. [DMHS08]: pre-processing (calculation of features vector, optical flow, and clipped regions), classification of regions using temporal coherence and a training set, user corrections (with updating of the training set), and brightness enhancement.*

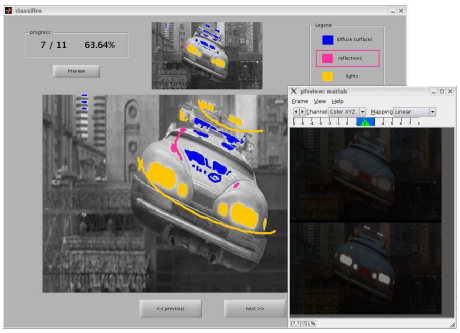

Figure 8: *The interface used for adjusting classification results.*

Clipped regions are enhanced by applying a non-linear adaptive tone curve, which is calculated based on partial derivatives within a clipped region stored in an histogram H. The tone curve is defined as an histogram equalisation on the inverted values of H:

$$f(b) = k \sum_{j=2}^{b} (1 - H[j]) + t_2$$

where t_2 is the lowest luma value for a clipped region, k is a scale factor that limits to the maximum boosting value m (equal to 150% for lights and 125% for reflections):

$$k = \frac{m - t_2}{\sum_{j=1}^{N} (1 - H[j])}$$

where N is the number of bins in H. To avoid contouring during boosting, the luma channel is filtered with bilateral filtering separating it into fine details and a base layer, which will be merged after luma expansion. See Section 3.6 for the validation part of this work. The method is semi-automatic, because intervention of the user is required.

3.4. Expand Map Models

The methods of Banterle et al. [BLDC06], its extensions [BLD*07, BLDC08], and Rempel et al. [RTS*07] use a guidance method to direct the expansion of the LDR content as opposed to global methods. Following the terminology used in Banterle et al. [BLDC06] we refer to these guidance methods as expand maps.

3.4.1. Non-Linear Expansion using Expand Maps

A general framework for expanding LDR content for HDR monitors and IBL was proposed by Banterle et al. [BLDC06, BLD*07]. The key points are the use of iTMO for expanding the range combined with a smooth field for the reconstruction of the lost over-exposed areas.

The first step of the framework is to linearise the input image, see Figure 9 for the pipeline. If the CRF is known, its inverse is applied to the signal. Otherwise, blind general methods can be employed such as Lin and et al.'s methods [LGYS04, LZ05]. Subsequently, the range of the image is expanded inverting a TMO. In their implementation, the inverse of the global Reinhard et al.'s operator [RSSF02] was used. This is because the operator has only two parameters, and range expansion can be controlled in a straightforward way. This iTMO is defined as:

$$L_w(\mathbf{x}) = \frac{1}{2} L_{w,\,\mathrm{Max}} L_{\mathrm{white}} \left(L_d(\mathbf{x}) - 1 + \sqrt{\left(1 - L_d(\mathbf{x})\right)^2 + \frac{4}{L_{\mathrm{white}}^2} L_d(\mathbf{x})} \right)$$

where $L_{w,\,\mathrm{Max}}$ is the maximum output luminance in cd/m^2 of the expanded image, and $L_{\mathrm{white}} \in (1, +\infty)$ is a parameter which determines the shape of the expansion curve. This is proportional to the contrast, authors suggested a value of $L_{\mathrm{white}} \approx L_{w,\,\mathrm{Max}}$ to increase the contrast while limiting artifacts due to expansion.

After range expansion, the expand map is computed. The expand map is a smooth field representing a low frequency version of the image in areas of high luminance. It has two main goals. The first is to reconstruct lost luminance profiles in over-exposed areas of the image. The second one is to attenuate quantisation or compression artifacts that can be enhanced during expansion. The expand map was implemented applying density estimation on samples generated using importance sampling (median-cut sampling [Deb05]). Finally, the expanded LDR image and the original one are combined using linear interpolation where the expand map acts as interpolation weight. Note that low luminance values are kept as in the original value. This avoids compression (for high L_{white} values) for low values which can result in contouring.

(a) (b)

Figure 10: *Application of Banterle et al.'s method [BLDC06, BLD*07] for re-lighting synthetic objects: a) Lucy's model is re-lighted using St. Peter's HDR lightprobe. b) Lucy's model is re-lighted using an expanded St. Peter's LDR lightprobe (starting at exposure 0).*

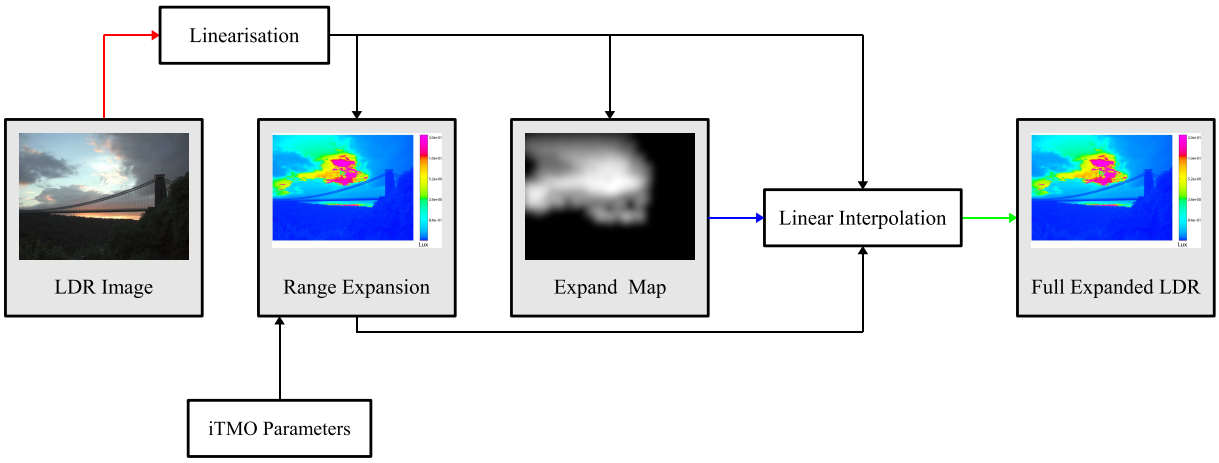

Figure 9: *The pipeline of Banterle et al.'s method [BLDC06, BLD*07].*

The framework was extended for automatically processing images and videos in Banterle et al. [BLDC08]. This is achieved using 3D sampling algorithms, volume density estimation, edge transfers, and a number of heuristics for determining the parameters of each component of the framework. Moreover, a coloured expand map was adopted, allowing the reconstruction of clipped colours. The main problem is the speed, but real-time performances on high definition content can be achieved using point-based graphics on GPU.

This algorithm presents a general solution for visualisation on HDR monitors and IBL, see Figure 10 for an example. Moreover, it was tested using HDR-VDP [MDMS05] for both tasks to prove its efficiency compared with simple exposure methods. The main limit of the framework is that large over-exposed areas (more than 30% of the image) can not be reconstructed using the expand map, producing grey smooth areas in the over-exposed areas. This is because there is not enough information to exploit.

3.4.2. LDR2HDR

A similar technique based on expand maps was proposed by Rempel et al. [RTS*07]. Their goal was real-time LDR expansion for videos. The algorithm pipeline is shown in Figure 11.

The first step of the LDR2HDR algorithm is to remove artifacts due to the compression algorithms of the media (such as MPEG) using a simple bilateral filter. Sophisticated artifact removal is not employed due to real-time constraints.

The next step of the method is to linearise the signal, using an inverse gamma function. Once the signal is linearised the contrast is stretched in an optimised way for the Dolby DR-37P HDR monitor [Dol05]. A simple linear contrast stretching is applied to boost values, however, they limited the maximum contrast to 5,000:1 to avoid artifacts. This means that the minimum value was mapped to 0.015 cd/m^2 while the maximum was mapped to 1,200 cd/m^2. To enhance brightness in bright regions a Brightness Enhance Function (BEF) is employed. This function is calculated applying a threshold of 0.92 (on a scale $[0,1]$ for LDR values). At this point the image is Gaussian filtered using a filter with a $\sigma = 30$ (150 pixels) which is chosen for 1920×1080 content. In order to increase contrast around edges an edge stopping function is used. Starting from saturated pixels, a flood-fill algorithm strategy is applied until an edge is reached, which is estimated using gradients. Subsequently, a morphological operator followed by a Gaussian filter with a smaller kernel is applied to remove noise. Finally, the BEF is mapped in the interval $[1,\alpha]$ where $\alpha = 4$ and finally, is multiplied with the scaled image to generate the HDR image, see Figure 12 for an example. To improve efficiency the BEF is calculated using Laplacian pyramids [BA87], which is implemented on the GPU or FPGA [Dol05].

The algorithm was evaluated using HDR-VDP [MDMS05] comparing the linearised starting image with the generated HDR image. This evaluation was needed to show that the proposed method does not introduce spatial artifacts during expansion of the content. Note that LDR2HDR processes each frame separately which may be not temporally coherent due to the nature of the BEF.

3.5. User Based Models

Since it may not always be possible to recover missing HDR content using automatic approaches, a different, user-based approach was proposed by Wang et al. [WWZ*07], whereby detailed HDR content can be added to areas that are meant to be expanded.

The authors demonstrated the benefits of an in-painting system to recover lost details in over-exposed and under-

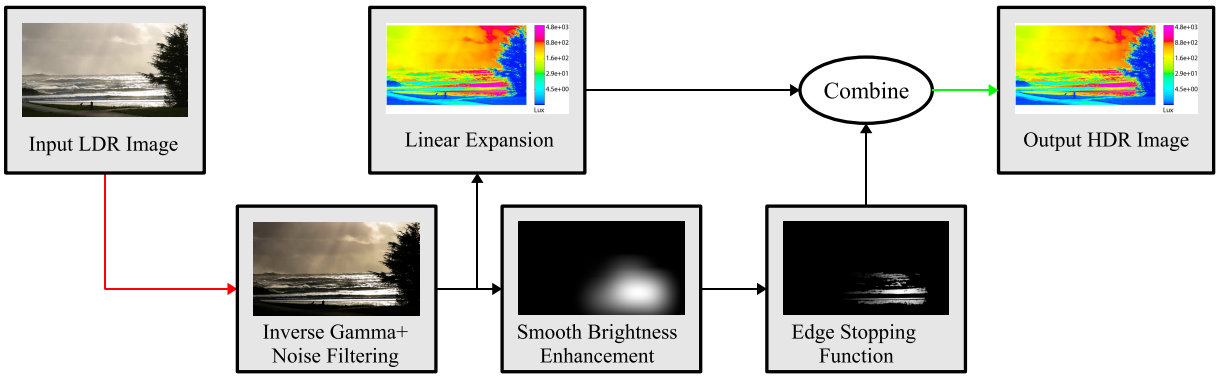

Figure 11: *The pipeline of Rempel et al.'s method [RTS*07].*

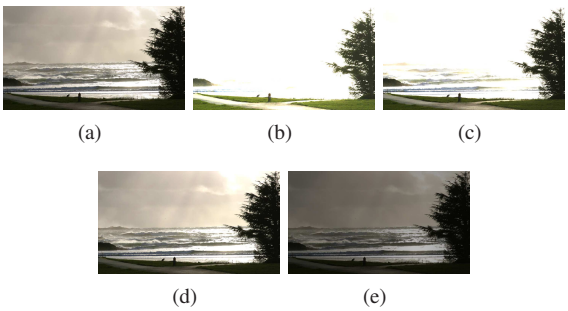

Figure 12: *Application of Rempel et al.'s method [RTS*07] to the Beach image from the authors' website: original LDR image (a) and different f-stops after expansion (b, c, d, e).*

exposed regions of the image, combined with a luminance boosting. The whole process was termed *hallucination*, and their system presents a mixture between automatic and user-based approaches.

The first step of hallucination is to linearise the signal, see Figure 13 for the complete pipeline. This is achieved with an inverse gamma function with $\gamma = 2.2$, which is the standard value for DVDs and television formats [ITU90]. After this step, the image is decomposed into large scale illumination and fine texture details. This is achieved by applying bilateral filtering to the image I obtaining a filtered version I_f. The texture details are obtained as $I_d = I/I_f$. Radiance for large scale illumination I_f is estimated using a linear interpolation of elliptical Gaussian kernels. Firstly, a weight map, w, is calculated for each pixel:

$$w(\mathbf{x}) = \begin{cases} \frac{C_{ue} - Y(\mathbf{x})}{C_{ue}} & Y(\mathbf{x}) \in [0, C_{ue}) \\ 0 & Y(\mathbf{x}) \in [C_{ue}, C_{oe}) \\ \frac{Y(\mathbf{x}) - C_{oe}}{1 - C_{oe}} & Y(\mathbf{x}) \in [C_{oe}, 1] \end{cases}$$

where $Y(\mathbf{x}) = R_s(\mathbf{x}) + 2G_s(\mathbf{x}) + B_s(\mathbf{x})$, and C_{ue} and C_{oe} are respectively the thresholds for under-exposed and over-exposed pixels. The authors suggested values of 0.05 and 0.85 for C_{ue} and C_{oe} respectively. Secondly, each over-exposed region is segmented and fitted with an elliptical Gaussian lobe G, where variance of the axis is estimated using region extents, and the profile is calculated using an optimisation procedure based on non over-exposed pixels at the edge of the region. The luminance is blended using a simple linear interpolation:

$$O(\mathbf{x}) = w(\mathbf{x})G(\mathbf{x}) + (1 - w(\mathbf{x}))\log_{10} Y(\mathbf{x})$$

Optionally users can add Gaussian lobes using a brush.

The texture details I_d are reconstructed using a texture synthesis technique similar to [BVSO03], where the user can select an area as a source region by drawing it with a brush. This automatic synthesis has some limits when scene understanding is needed, therefore a warping tool is included. This allows the user to select with a stroke-based interface a source region and a target region, and pixels will be transferred. This is a tool similar to the stamp and healing tools in Adobe Photoshop [Ado07].

Finally, the HDR image is built blending the detail and the large scale illumination, this is performed using Poisson image editing [PGB03] in order to avoid seams in the transition between expanded over-exposed areas and well-exposed areas.

This system can be used for both IBL and visualisation of images, and compared with other algorithms it may maintain details in clamped regions. However, the main problem of this approach is that it is user based and not automatic, which potentially limits its use to single images and not videos.

3.6. Validation

The development of methods for LDR expansion has produced various algorithms with different features. Therefore,

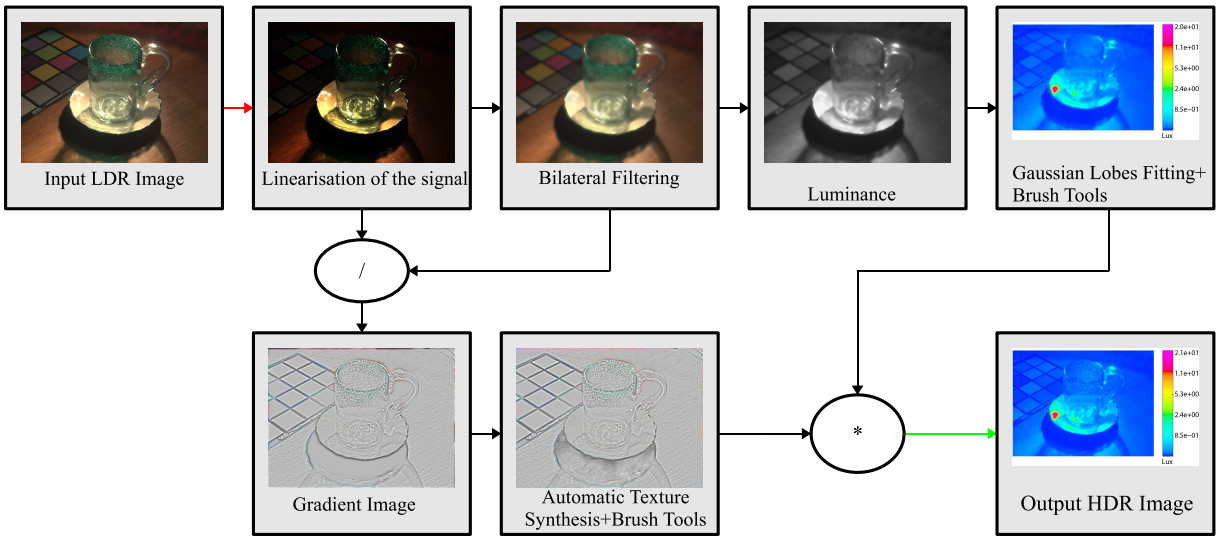

Figure 13: *The pipeline of the Wang et al. method [WWZ*07].*

there is a need to determine the quality performances of such algorithms to understand which method is more suitable for a given image and task. Moreover, the analysis of performances can help to highlight important features, such non-linearity, that can be important for the design of future expansion techniques.

3.6.1. HDR-VDP Comparisons

Banterle et. al [BLDC06, BLD*07] used the HDR Visual Difference Predictor (HDR-VDP) [MDMS05] for validating the quality of reconstruction against a ground truth and simple expansion operator without expand maps. The results showed that their proposed method reconstructs closer to the reference the missing HDR content.

Moreover, HDR-VDP was applied to compare re-lighted images with an HDR reference. This showed that LDR expansion allows a small error in the case of IBL.

3.6.2. Pairwise Comparisons Study for Video Sequences

In Section 3.3.2, Didyk et al. [DMHS08] presented a new operator for the expansion of LDR videos based on classification. Enhanced videos generated with this method were compared with the original videos and the ones generated, using the only method suitable for videos at the time, Rempel et al.'s [RTS*07] LDR2HDR. Comparisons were performed running a paired comparisons psychophysical experiment [Dav88, LCTS05] with 14 naïve subjects using a LCD Barco Coronis Color 3MP Diagnostic Luminance (12-bit per colour channel). The participants ranked 9 videos times 3 combinations: original video, Rempel et al. and their method. The study was analysed using a similar approach to Ledda et al. [OA07]. The experiment showed that for the overall scores Didyk et al.'s method was preferred with statistical significance compared to both the original video and the Rempel et al.'s one. However, there was no statistical significance between this method and other ones for six of the considered videos.

3.6.3. Dynamic Range Independent Image Quality Assessment

Aydin et al. [AMMS08] proposed a new perceptual metric (DI-IQA) which allows the comparison of images independently from their dynamic range. This metric can detect the loss and amplification of contrast, and the change of structure in images.

Due to the capabilities of this metric, quantisation artifacts and changes in the image details visibility can be quantified where they happen. Therefore, it can be employed to validate the quality of expansion algorithms avoiding time consuming psychophysical experiments. Authors presented a few examples where they applied the metric to expanded images showing when the signal starts to be distorted in function of the expansion, see Figure 14 for an example.

3.6.4. Pairwise Comparisons Studies for Image Visualisation and Image Based Lighting

Banterle et al. [BLD*08] proposed a psychophysical study for the evaluation of expansion algorithms based on pairwise comparisons methodology [Dav88, LCTS05] using an HDR reference image displayed on the Dolby DR-37p HDR monitor [Dol05]. The study involved 24 participants, and five algorithms were tested: Banterle et al. [BLDC06, BLD*07] (**B**), Meylan et al. [MDS06, MDS07] (**M**), Wang et al. [WWZ*07] (**W**), Rempel et al. [RTS*07] (**R**), and Akyüz

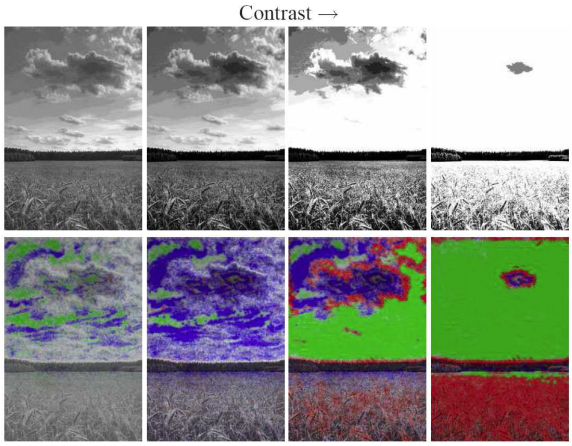

Figure 14: *An example of DI-IQA metric by Aydin et al. [AMMS08]. The response of the metric to simple contrast stretching with clipping. Contrast is increased from left to right, which results in more clipping and generates stronger visible contrast loss and reversal responses.*

et al. [AFR*07] (**A**). The study was divided in two experiments. The first one tested performances of various expansion algorithms for the recreation of eight HDR images starting from clipped ones. A subject had to choose the best picture in a pair which was closer to the reference on overall, in the dark areas, and in the bright ones. The second experiment had as goal to determine which expansion method performs better than the other for recreating six HDR environment maps for IBL for three different materials: pure diffuse, pure specular, and glossy. Each subject had to choose the best re-lighted object (a teapot) in a pair which was closer to the reference.

For the first experiment, the monotonically increasing functions **B**, **W** and **R** that enhance contrast non-linearly performed better overall and were grouped together in many of the results. The linear method **A**, and to a lesser extent **M**, performed worse overall, reflecting that for still images complex methods recreate HDR perceptually better.

For the second experiment, the diffuse results showed few differences. This is mostly due to the fact that rendering with IBL consists of the evaluation of an integral and during integration small details may be lost. This is less true for perfectly mirror-like or high glossy materials. However, in these cases details of the environment map reflected in the objects may be too small to be seen as was shown by the large groupings in the results. For more complex environment maps, the previously found ranking was reverted. Overall it is still clear that the operators, that perform best, as with the first experiment, are the non-linear operators.

This study showed that more advanced algorithms, that cater for quantisation errors introduced during expansion of an LDR image, such as **B**, **R** and **W**, can perform better than simple techniques that apply single or multiple linear scale expansions, such as **A** and **M**. The more computationally expensive methods **B**, **R** and **W**, are better at recreating HDR than simple methods. Even if a linear scale can elicit an HDR experience in an observer, as shown in [AFR*07], it does not correctly reproduce the perception of the original HDR image.

3.6.5. Exposure Understanding for Content Expansion

Martin et al. [MFS*08] presented an on going psychophysical study on evaluation of expansion algorithms. This study is divided in two parts. The first part of the study is focused on perception of exposure in images. The results of the first part showed that high-level semantics are needed for a proper classification of exposure. Moreover, an asymmetry in the perception of under-exposed and over-exposed images was found.

The second part consisted of side by side evaluation of the following expansion methods on a Dolby DR-37p monitor [Dol05] with images at different exposure levels: original LDR, Banterle et al. [BLDC06, BLD*07], Rempel et al. [RTS*07], and Akyüz [AFR*07]. The results of this experiment have not currently been published.

3.7. Overview

In Table 1 we present an overview of all the methods discussed in this section, summarising what techniques they use, and how they compare in terms of quality and performance. We find that most methods expand the dynamic range using either a linear or non-linear function, while Meylan et al. use a two-scale linear function. The reconstruction methods aim at smoothly expanding the dynamic range and a variety of methods are proposed. Unsurprisingly, the choice of expansion method and reconstruction influences the computational performance of the method and the quality. We present performance based on the timings from the individual papers and/or the complexity of the computation involved, where fast performance would make it possible to perform in real-time on current hardware while slow would require a handful of seconds. Wang et al.'s method requires a manual intervention somewhat hindering real-time performance. The quality results we present are based in other publications, primarily the psychophysical experiments shown in Banterle et al. [BLD*08]. It is clear that different methods are suitable for different applications, and the more straightforward methods are faster and more suitable for IBL or just improving highlights. For more complex still scenes and/or videos where further detail may be desirable, the more complex expansion methods may be preferable.

4. HDR Compression using Tone Mapping and Inverse Tone Mapping

HDR expansion methods have not only been employed for the generation of content from a single exposure image, but

Method	Expansion Function	Reconstruction+ Noise Reduction	Performance	Quality	Video
Daly and Feng 1 [DF03]	Linear	Additive Noise	Fast	Good$^\bullet$	Yes
Daly and Feng 2 [DF04]	Linear	Filtering	Fast	Good$^\bullet$	Yes
Landis [Lan02]	Non-Linear	N/A	Fast	Good for IBL	No
Akyüz et al. [AFR*07]	Linear	N/A	Fast	Average*	Yes
Meylan et al. [MDS06, MDS07]	Two Scale Linear	Filtering	Fast	Average* Good for Highlights*	Potential
Didyk et al. [DMHS08]	Non-Linear	Filtering+ Classification	Slow	Good$^\circ$	Yes
Banterle et al. [BLDC06, BLD*07, BLDC08]	Non-Linear	Expand Map	Slow	Good*	Yes
Rempel et al. [RTS*07]	Linear	Expand Map	Fast in hardware	Good*	Yes
Wang et al. [WWZ*07]	Non-Linear	Bilateral Filtering+ Texture Transfer	Manual	Good*	No

Table 1: *Classification of algorithms for expansion of LDR content. $^\circ$ is based on a psychophysical study in Didyk et al. [DMHS08]. $^\bullet$ is designed for medium dynamic range monitor, and not for IBL. * is based on a psychophysical study in Banterle et al. [BLD*08].*

they have proven beneficial for HDR content compression. These methods typically comprise of the compression of the dynamic range via tone mapping. The tone mapped image is subsequently encoded via traditional compression methods such as JPEG, in the case of images, or MPEG in the case of videos. These two steps comprise the encoding aspect of the compression. Decoding takes the role of the LDR compression's decoding method followed by an HDR expansion, usually inverting the method that was used for the dynamic range compression.

This approach to the compression of HDR content has the advantage of re-using previous compression schemes and standards. Also, it can allow backward-compatibility because the function for HDR expansion can be easily stored in an extension header of a standard. These functions require only a few parameters to be stored.

4.1. Backward Compatible JPEG-HDR

JPEG-HDR is an extension to the JPEG compression scheme for HDR images by Ward and Simmons [WS04, WS05]. The method does not use an explicit iTMO, nevertheless a spatial inverse function called Ratio Image (RI) is employed.

The encoding, see Figure 15, starts with the tone mapping of the HDR image discretised to 8-bit. After this, the original HDR image is divided by the tone mapped one obtaining the RI which is stored as a sub-band. The RI can be down-sampled reducing the sub-band size, because the HVS has a limited ability to detect large and high frequency changes in luminance. This fact was also exploited in Setzeen et al. [SHS*04] to improve the efficiency of HDR displays. However, down-sampling needs correction of the image, because the naïve multiplication of a down-sampled image times the tone mapped LDR image can produce halos/glare around edges. This problem can be solved in two ways: pre-correction and post-correction. The former method introduces corrections in the tone mapped image. This is achieved by down-sampling and afterwards up-sampling the RI image obtaining RI_d. Subsequently, the original HDR image is divided by RI_d, which is a tone mapped image with corrections. While this approach is effective, it can produce artifacts in the LDR image for the backward compatibility and this cannot be acceptable in many applications. The latter method consists of an up-sampling with guidance which is more expensive than the pre-correction one. While RI_d is discretised at 8-bit in the logarithmic domain and stored in application markers of JPEG, the tone mapped layer needs further processing for preserving colours. Two techniques are employed to solve this problem: compression of the gamut and a new $\mathrm{YC}_b\mathrm{C}_r$ encoding. The gamut compression produces a global desaturation. Given the following definition of saturation:

$$S(\mathbf{x}) = 1 - \frac{min(R(\mathbf{x}), G(\mathbf{x}), B(\mathbf{x}))}{L_\mathrm{w}(\mathbf{x})}$$

the desaturation of each colour channel is achieved by:

$$\begin{bmatrix} R_c(\mathbf{x}) \\ G_c(\mathbf{x}) \\ B_c(\mathbf{x}) \end{bmatrix} = \left(1 - S(\mathbf{x})'\right) \begin{bmatrix} L_\mathrm{w}(\mathbf{x}) \\ L_\mathrm{w}(\mathbf{x}) \\ L_\mathrm{w}(\mathbf{x}) \end{bmatrix} + S(\mathbf{x})' \begin{bmatrix} R_c(\mathbf{x}) \\ G_c(\mathbf{x}) \\ B_c(\mathbf{x}) \end{bmatrix} \quad (2)$$

where $\alpha \leq 1$ is a parameter which controls the level of saturation kept during colour encoding, β is a parameter which determines the colour contrast, and $S'(\mathbf{x}) = \alpha S(\mathbf{x})^{\beta-1}$ is the desaturation level. After this step, the image is encoded in a modified $\mathrm{YC}_b\mathrm{C}_r$ colour space, because it has a larger gamut than RGB colour space. Therefore, unused $\mathrm{YC}_b\mathrm{C}_r$ values can be exploited to preserve the original gamut of an HDR image. This is achieved by the following mapping:

$$R'(\mathbf{x}) = \begin{cases} 1.055R_c(\mathbf{x})^{0.42} - 0.055 & \text{if } R_c(\mathbf{x}) > t_r \\ 12.92R_c(\mathbf{x}) & \text{if } |R_c(\mathbf{x})| \leq t_r \\ -1.055(-R_c(\mathbf{x}))^{0.42} + 0.055 & \text{if } R_c(\mathbf{x}) < -t_r \end{cases}$$
$$t_r = 0.0031308$$

which is repeated for the green and blue channel. Finally, the standard mapping from RGB to YC_bC_r is used for the JPEG encoding.

The decoding for the pre-correction case consists of few steps, see Figure 16 for the complete pipeline. Firstly, the tone mapped layer is decoded using a JPEG decoder and the gamut is expanded inverting Equation 2. After this step, the RI_d image is decoded, expanded (from logarithmic domain to linear domain), and up-sampled to the resolution of the tone mapped layer. Finally, the image is recovered by multiplying the tone mapped layer by the RI_d image.

A first study [WS04] was conducted to determine a good TMO for compression purposes, which was based on comparison with the original HDR images using VDP [Dal93]. In this experiment different TMOs were compared such as histogram adjustment [LRP97], global photographic tone reproduction operator [RSSF02], fast bilateral filtering operator [DD02] and the gradient operator [FLW02]. Experiments showed that the fast bilateral filtering operator performed the best followed by the global photographic tone reproduction one. A second study was carried out to test image quality and compression rates on a data set of 217 HDR images. The data set was compressed using JPEG-HDR at different quality settings using the global photographic operator, RGBE, OpenEXR and LogLuv TIFF to study compression rates. HDR images compressed using JPEG-HDR were compared with original ones using VDP to quantify the quality of the resultant images. The study showed that the method can achieve a compression rate between 0.6-3.75 bpp for quality settings between $57-99\%$. However, quality degrades rapidly for JPEG quality below 60%, but only 2.5% of pixels were visibly different with a quality set at 90%, and only 0.1% with maximum quality.

Most importantly, the method is backward compatible, because RI_d is encoded using only extra application markers of JPEG. When an old application or one that is not designed for HDR imaging will open a JPEG-HDR file, it will display only the tone mapped layer allowing the user to have access to the LDR part of the content.

4.2. HDR-JPEG 2000

Xu et al. [XPH05] proposed a simple pre-processing technique which enables the JPEG 2000 standard [CSE00] to encode HDR images. The main idea is to transform floating point data in unsigned short integers (16-bit), that are supported by JPEG 2000 standard.

The encoding phase starts with the reduction of the dynamic range by applying a logarithm to the RGB values:

$$\begin{bmatrix} R'_{\mathrm{w}}(\mathbf{x}) \\ G'_{\mathrm{w}}(\mathbf{x}) \\ B'_{\mathrm{w}}(\mathbf{x}) \end{bmatrix} = \begin{bmatrix} \log R_{\mathrm{w}}(\mathbf{x}) \\ \log G_{\mathrm{w}}(\mathbf{x}) \\ \log B_{\mathrm{w}}(\mathbf{x}) \end{bmatrix}$$

Subsequently, the floating point values in the logarithm domain are discretised to unsigned short integers:

$$\begin{bmatrix} \overline{R}_{\mathrm{w}}(\mathbf{x}) \\ \overline{G}_{\mathrm{w}}(\mathbf{x}) \\ \overline{B}_{\mathrm{w}}(\mathbf{x}) \end{bmatrix} = \begin{bmatrix} f(R'_{\mathrm{w}}(\mathbf{x})) \\ f(G'_{\mathrm{w}}(\mathbf{x})) \\ f(B'_{\mathrm{w}}(\mathbf{x})) \end{bmatrix} \qquad f(x,n) = (2^n - 1)\frac{x - x_{\min}}{x_{\max} - x_{\min}} \tag{3}$$

where $x_{\max}$ and $x_{\min}$ are respectively the maximum and minimum values for the channel of x, and $n = 16$. Finally, the image is compressed using a classic JPEG 2000 encoder.

The decoding phase is quite straightforward. Firstly, the image is decompressed using a JPEG 2000 decoder, then the integer data is converted into floating point by inverting f in Equation 3:

$$\begin{bmatrix} R'_{\mathrm{w}}(\mathbf{x}) \\ G'_{\mathrm{w}}(\mathbf{x}) \\ B'_{\mathrm{w}}(\mathbf{x}) \end{bmatrix} = \begin{bmatrix} g(\overline{R}_{\mathrm{w}}(\mathbf{x})) \\ g(\overline{G}_{\mathrm{w}}(\mathbf{x})) \\ g(\overline{B}_{\mathrm{w}}(\mathbf{x})) \end{bmatrix}$$
$$g(x,n) = f^{-1}(x,n) = \frac{x}{2^n - 1}(x_{\max} - x_{\min}) + x_{\min}$$

Finally, values are exponentiated to get the final dynamic range:

$$\begin{bmatrix} R_{\mathrm{w}}(\mathbf{x}) \\ G_{\mathrm{w}}(\mathbf{x}) \\ B_{\mathrm{w}}(\mathbf{x}) \end{bmatrix} = \begin{bmatrix} e^{R'_{\mathrm{w}}(\mathbf{x})} \\ e^{G'_{\mathrm{w}}(\mathbf{x})} \\ e^{B'_{\mathrm{w}}(\mathbf{x})} \end{bmatrix}$$

The method using JPEG 2000 lossy mode was compared to JPEG-HDR [WS05] and HDRV [MKMS04], and when using JPEG 2000 lossless mode it was compared with RGBE [War91], LogLuv [Lar98], and OpenEXR [Ind02]. The metrics used for the comparison were RMSE in the logarithm domain and Lubin's VDP [Lub95]. The results of these comparisons showed that HDR-JPEG 2000 in lossy mode is superior to JPEG-HDR and HDRV, especially at low bit rates when these methods produce visible artifacts. Nevertheless, the method does not perform well when lossless JPEG 2000 is used, because the file size is higher than RGBE, LogLuv, and OpenEXR (these methods are lossy in terms of per-pixel float precision, but not spatially over pixel neighbourhoods).

The HDR-JPEG 2000 algorithm is a straightforward method for lossy compression of HDR images at high quality, without artifacts at low bit rates. However, the method

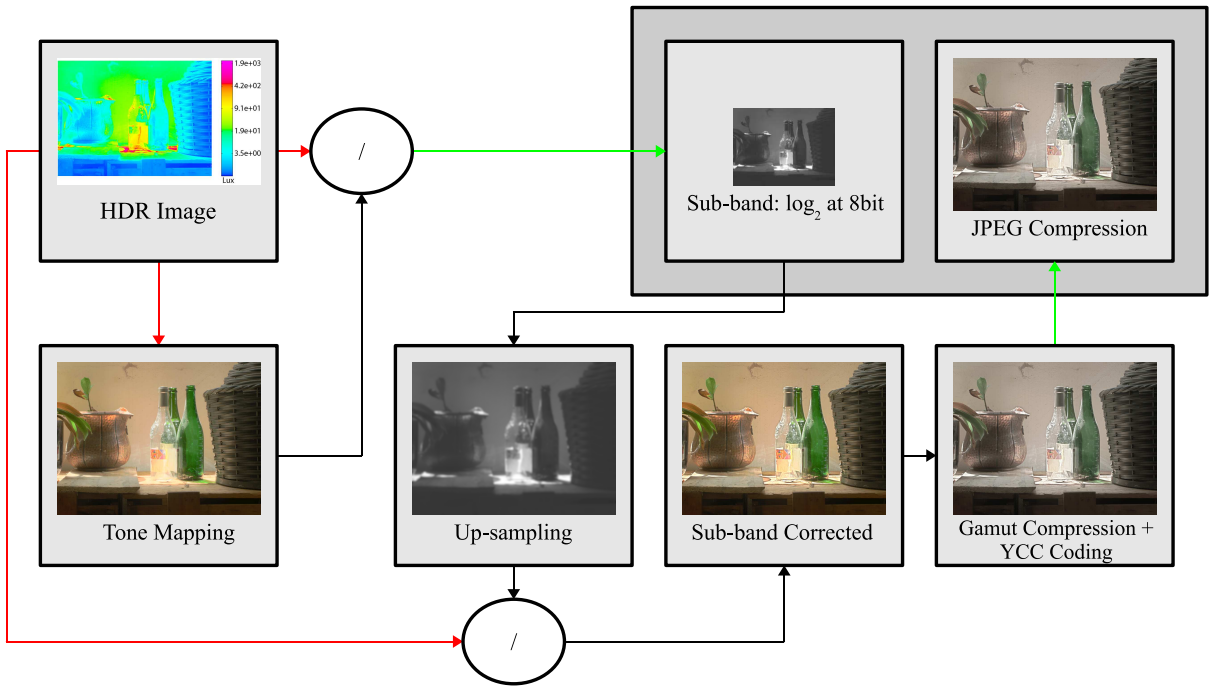

Figure 15: *The encoding pipeline for JPEG-HDR for pre-correction case by Ward and Simmons [WS04, WS05].*

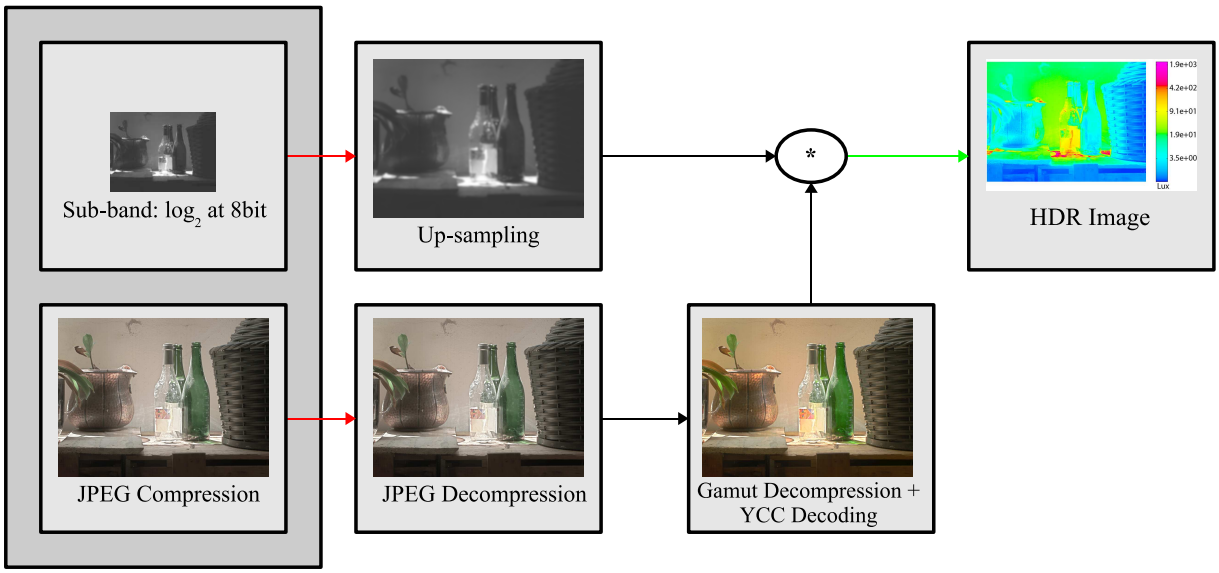

Figure 16: *The decoding pipeline for JPEG-HDR by Ward and Simmons [WS04, WS05].*

is not suitable for real-time graphics, because fixed time look-ups are needed. Also, the method does not exploit all the compression capabilities of JPEG 2000 because it operates at high level. For example, separate processing for luminance and chromaticity could reduce the size of the final image while keeping the same quality.

4.3. Compression and Companding High Dynamic Range Images with Sub-bands Architectures

Li et al. [LSA05] presented a general framework for tone mapping and inverse tone mapping of HDR images based on multi-scale decomposition. While the main goal of the algorithm is tone mapping, in addition, the framework can also compress HDR images. A multi-scale decomposition splits a signal $s(x)$ (1D in this case) into n sub-bands $b_1(x),...,b_n(x)$ with n filters $f_1,...,f_n$, in a way the signal can be reconstructed as:

$$s(x) = \sum_{i=1}^{n} b_i(x)$$

Wavelets [SDS95] and Laplacian pyramids [BA87] are examples of multi-scale decomposition that can be used in Li et al.'s framework.

(a) (b) (c)

Figure 17: *An example of tone mapping using the multi-scale decomposition with Haar Wavelets. a) Activity map, b) Gain map and c) Tone mapped luminance.*

The main concept is to apply a gain control to each sub-band of the image to compress the range. For example, a sigmoid expands low values and flats peaks, however it introduces distortions that can appear in the final reconstructed signal. In order to avoid such distortions, a smooth gain map inspired by neurons was proposed. The first step, is to build an activity map, reflecting the fact that the gain of a neuron is controlled by the level of its neighbours. The activity map is defined as:

$$A_i(\mathbf{x}) = G(\sigma_i) \otimes |B_i(\mathbf{x})|$$

where $G(\sigma_i)$ is a Gaussian kernel with $\sigma_i = 2^i\sigma_1$ which is proportional to i, the sub-band's scale. The activity map is used to calculate the gain map, which turns gain down where activity is high and vice versa:

$$G_i(\mathbf{x}) = p(A_i\mathbf{x}) = \left(\frac{A_i\mathbf{x}+\varepsilon}{\delta_i}\right)^{\gamma-1}$$

where $\gamma \in [0,1]$ is a compression factor, and ε is the noise level that prevents the noise from being seen. $\delta_i = \alpha_i \sum_{\mathbf{x}} A_i(\mathbf{x})/(M)$ is the gain control stability level where M is the number of pixels in the image, $\alpha_i \in [0.1,1]$ is a constant related to spatial frequency. Once the gain maps are calculated, sub-bands can be modified:

$$B_i'(\mathbf{x}) = G_i(\mathbf{x})B_i(\mathbf{x}) \tag{4}$$

Note that it is possible to calculate a single activity map for all sub-bands by pooling all activity maps:

$$A_{ag}(\mathbf{x}) = \sum_{i=1}^{n} A_i(\mathbf{x})$$

From A_{ag}, a single gain map $G_{ag} = p(A_{ag})$ is calculated for modifying all sub-bands. The tone mapped image is finally obtained summing all modified sub-bands B_i', see Figure 17. The compression is applied only to the V channel of an image in the HSV colour space. Finally, to avoid oversaturated images S can be reduced by $\alpha \in [0.5,1]$. The authors presented a comparison with the fast bilateral filter operator [DD02] and gradient domain operator [FLW02].

The framework can be additionally used for the compression task, applying expansion after compression, called *companding*. The expansion operation is obtained by a straightforwad modification of Equation 4:

$$B_i'(\mathbf{x}) = \frac{B_i(\mathbf{x})}{G_i(\mathbf{x})}$$

A straightforward companding operation is not sufficient for compression especially if the tone mapped image is compressed using lossy codecs. Therefore, the companding operation needs to be iterative to determine the best values for the gain map, see Figure 18. The authors proposed to compress the tone mapped image using JPEG. In this case a high bit-rate is needed (1.5 bpp - 4bpp) with chrominance subsampling disabled to avoid that JPEG artifacts are amplified during expansion, because a simple up-sampling strategy is adopted.

4.4. Backward Compatible HDR-MPEG

Backward compatible HDR-MPEG is a codec for HDR videos that was introduced by Mantiuk et al. [MEMS06].

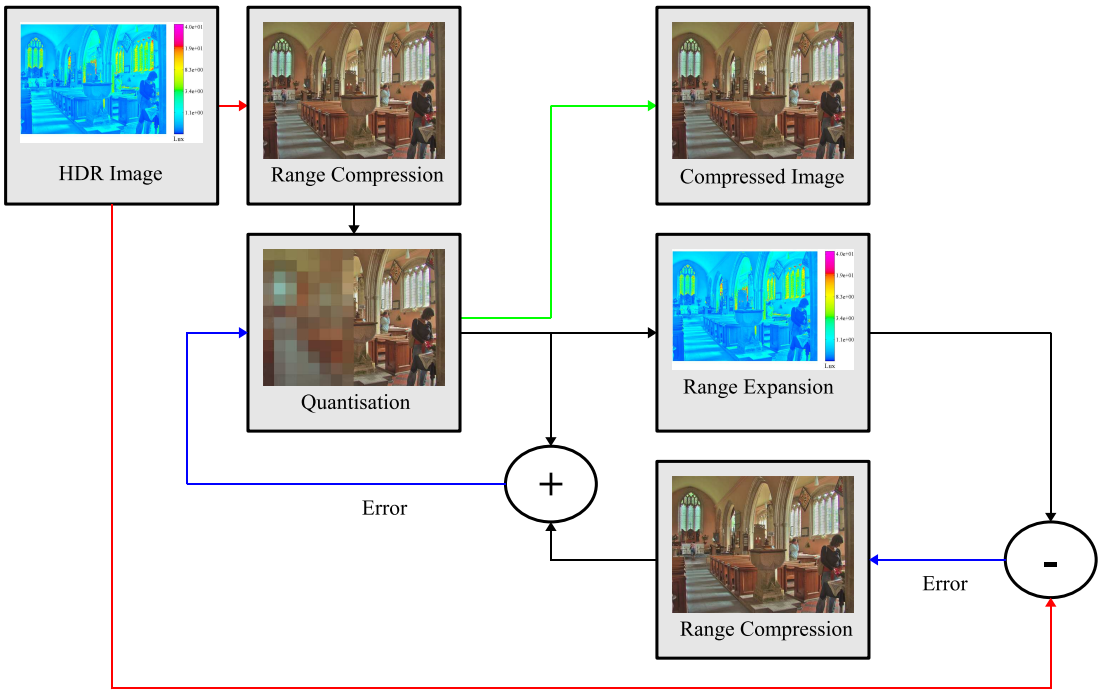

Figure 18: *The optimisation companding pipeline of Li et al. [LSA05] .*

As in the case of JPEG-HDR this algorithm is an extension to the standard MPEG-4 codec (H.264) [WSBL03] that works on top of the standard encoding/decoding stage allowing backward compatibility. In a similar way to JPEG-HDR each frame is divided into an LDR part, using tone mapping, and an HDR part. However, in this method, the reconstruction function (RF) a tabled iTMO is employed instead of a RI. HDR-MPEG is a natural extension of perception motivated video encoding (HDRV) [MKMS04]. However, the primary features of the HDRV codec design is that it is a modification of standard MPEG-4 with new steps in the encoding/decoding stage such as the perceptual luminance encoding. Morevoer, HDRV was designed for a target of 10-11 bit for luminance, a format that is rarely supported in software and hardware, which compromises its backward compatibility.

The decoding stage takes as input an HDR video in the XYZ colour space and it applies tone mapping to each HDR frame obtaining LDR frames as a first step, see Figure 19 for the complete pipeline. These are coded with MPEG-4, stored in an LDR stream, and finally decoded to obtain a uncompressed and MPEG quantised frames. Subsequently, the LDR frame and the HDR frame are converted to a common colour space. For both HDR and LDR frames CIE 1976 Uniform Chromaticity (u',v') coordinates are used to code chroma. While non-linear luma of sRGB is used for LDR pixels, a different luma coding is used because sRGB non-linearity is not suitable for high luminance ranges $[10^{-5},10^{10}]$, see [MEMS06]. This luma coding, at 12-bit, for HDR luminance values is given as:

$$l_{\mathrm{w}} = f(L_{\mathrm{w}}) = \begin{cases} 209.16\log(L_{\mathrm{w}}) - 731.28 & \text{if } L_{\mathrm{w}} \geq 10469 \\ 826.81 L_{\mathrm{w}}^{0.10013} - 884.17 & \text{if } 5.6046 \leq L_{\mathrm{w}} < 10469 \\ 17.554 L_{\mathrm{w}} & \text{if } L_{\mathrm{w}} < 5.6046 \end{cases}$$

where its inverse transform, $g(l_{\mathrm{w}}) = f^{-1}(l_{\mathrm{w}})$, is:

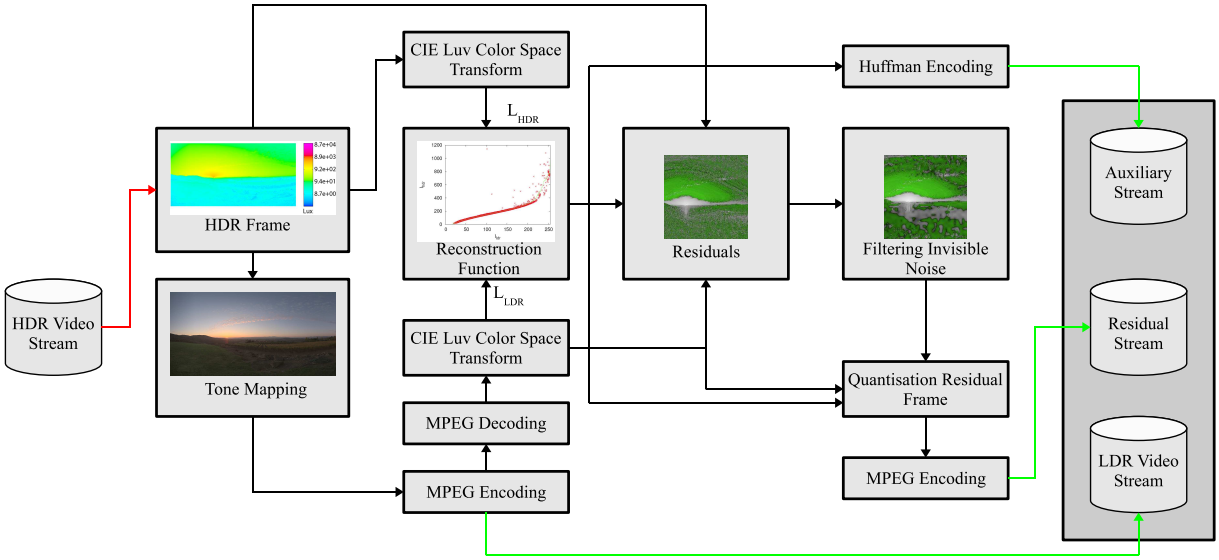

Figure 19: *The encoding pipeline for Backward Compatible HDR-MPEG by Mantiuk et al. [MEMS06].*

$$L_{\mathrm{w}} = g(l_{\mathrm{w}}) = \begin{cases} 32.994\exp(0.0047811 l_{\mathrm{w}}) & \text{if } l_{\mathrm{w}} \geq 1204.7 \\ 7.3014e-30(l_{\mathrm{w}}+884.17)^{9.987} & \text{if } 98.381 \leq l_{\mathrm{w}} < 1204.7 \\ 0.056968 l_{\mathrm{w}} & \text{if } l_{\mathrm{w}} < 98.381 \end{cases} \quad (5)$$

At this point both the HDR and the LDR frames are in a comparable colour space, and an RF, that maps l_{d} to l_{w}, is calculated in a straightforward way by averaging l_{w}, which falls into one of 256 bins representing the l_{d} values:

$$RF(i) = \frac{1}{|(\Omega_l)|} \sum_{\mathbf{x}\in\Omega_l} l_{\mathrm{w}}(\mathbf{x}) \qquad \text{where } \Omega_l = \left\{ i | l_{\mathrm{d}}(\mathbf{x}) = l \right\}$$

where $l \in [0,255]$ is an index of a bin, $l_{\mathrm{d}}(\mathbf{x})$ and $L_{\mathrm{w}}(\mathbf{x})$ are respectively the luma for LDR and HDR pixel at $\mathbf{x}$. RF for chromaticity is approximated imposing $(u'_{\mathrm{d}}, v'_{\mathrm{d}}) = (u'_{\mathrm{w}}, v'_{\mathrm{w}})$. Once RFs are calculated for all frames, they are stored in an auxiliary stream using Huffman encoding.

After this stage a residual image is calculated for improving overall quality:

$$r_l(\mathbf{x}) = l_{\mathrm{w}}(\mathbf{x}) - RF(l_{\mathrm{d}}(\mathbf{x}))$$

The residual image is discretised at 8 bit, using a quantisation factor different for each bin based on its maximum magnitude value, which leads to:

$$\hat{r}_l(\mathbf{x}) = \left\lfloor \frac{r_l(\mathbf{x})}{q(m)} \right\rfloor_{-127}^{127} \qquad \text{where } m = k \Leftrightarrow i \subset \Omega_k$$

where $q(m)$ is the quantisation factor which is calculated for a bin Ω_l as:

$$q(m) = max\left(q_{min}, \frac{max_{\mathbf{x}\in\Omega_l}(|r_l(\mathbf{x})|)}{127} \right)$$

$\hat{r}_l$ needs to be compressed in a stream using MPEG, but a naïve compression would generate a low compression rate, because a large amount of high frequencies are present in $\hat{r}_l$. In order to improve the compression rate, the image is filtered removing frequencies in regions that are not distinguishable by the HVS. This is achieved by using the original HDR frame as guidance to the filtering. The filtering is performed in the wavelet domain, and it is applied only to the three finest scales modeling contrast masking, and lower sensibility to high frequencies.

The decoding stage is quite straightforward. MPEG streams (tone mapped video and residuals) and RF streams are decoded, see Figure 20 for the complete pipeline. Then, an HDR frame is reconstructed applying firstly its RF to the LDR decoded frame, and secondly adding residuals to the expanded LDR frame. Finally, CIE Luv values are converted to XYZ ones using Equation 5 for luminace.

HDR-MPEG was evaluated using three different metrics: HDR VDP [MDMS05], universal image quality index (UQI) [WB02], and classic Signal to Noise Ratio (SNR).

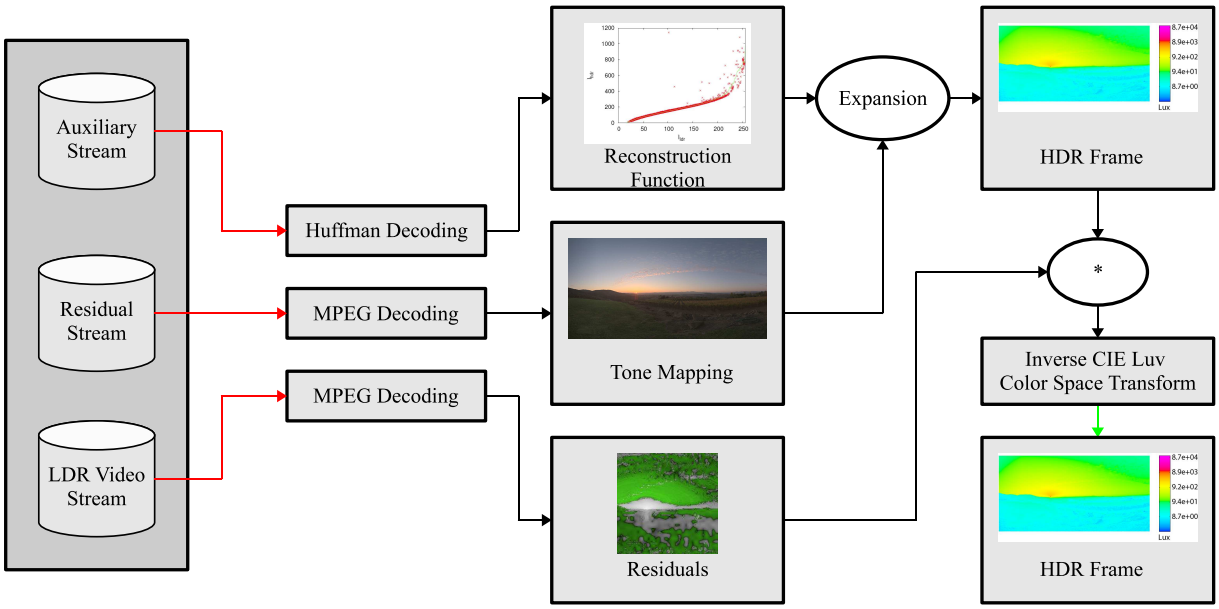

Figure 20: *The decoding pipeline for Backward Compatible HDR-MPEG by Mantiuk et al. [MEMS06].*

As in the case of JPEG-HDR, there was first a study that explored the influence of a TMO on quality/bit-rate. This experiment was performed using different TMOs such as time dependent visual adaption [PTYG00], fast bilateral filtering [DD02], photographic tone reproduction [RSSF02], gradient domain [FLW02] and adaptive logarithmic mapping [DMAC03]. These TMOs were modified to avoid temporal flickering and applied to a stream using default parameters. The study showed that most of these TMOs have the same performances except the gradient domain one, which creates larger streams than others. However, this TMO generated more attractive images for backward compatibility, therefore the choice of a TMO for the video compression depends by the trade-off between bit-rate and the needed backward compatible quality. The second study compared HDR-MPEG against HDRV [MKMS04] and JPEG-HDR [WS05] using the photographic tone reproduction operator as the TMO [RSSF02]. The results showed that HDR-MPEG has a better quality than JPEG-HDR, but a worse PSNR/bit rate ratio than HDRV.

4.5. Encoding of High Dynamic Range Video with a Model of Human Cones

Similarly to Li et al. [LSA05], Van Hateren [Hat06] proposed a new TMO based on a model of human cones [HL06] which can be inverted to encode HDR images and videos. The TMO and iTMO work in troland units (td), the measure of retinal illuminance I, which is derived by the scene luminance in cd/m^2 multiplied by the pupil area in mm^2. Van Hateren proposed a temporal and a static version of its TMO.

The temporal TMO is designed for HDR videos and presents low-pass temporal filters for removing photon and source noise, see Figure 21. The TMO starts by simulating the absorption of I by visual pigment, which is modelled by two low-pass temporal filters that are described in terms of a differential equation:

$$\frac{dy}{dt} + \frac{1}{\tau} y = \frac{1}{\tau} x$$

where τ is a time constant, $x(t)$ is the input at time t, and $y(t)$ is the output. At this point, a strong non linearity is applied to the result of low-pass filters E^* for simulating the breakdown of cyclic guanosine monophosphate (cGMP) by enzymes (cGMP is a nucleotide, that controls the current across the cell membranes):

$$\alpha = \frac{1}{\beta} = \left(c_\beta + k_\beta E^* \right)^{-1}$$

where $k_\beta E^*$ is the light-dependent activity of an enzyme, and c_β the residual activity. The breakdown of cGMP is counteracted by the production of cGMP; a highly non-linear feedback loop under control of inter-cellular calcium. This system is modelled by a filtering loop which outputs the current across cell membrane, I_{os} (the final tone mapped value), by the outer segment of a cone.

Van Hateren showed that the range expansion is quite straightforward by inverting the feedback loop. However, the

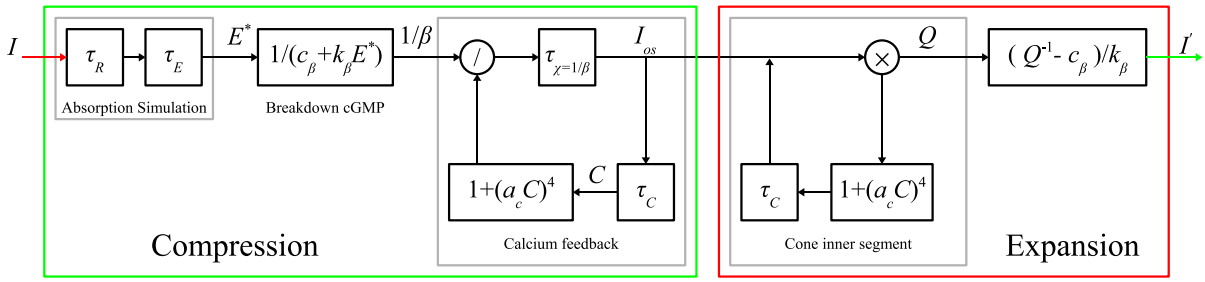

Figure 21: *The pipeline for range compression (red) and range expansion (green) proposed by Van Hateren [Hat06].*

process cannot be fully inverted because the first two low-pass filters are difficult to invert, so this results in $I' \approx I$. In order to fully invert the process for inverse tone mapping purposes Van Hateren proposed a steady version of the TMO, a global TMO, defined as:

$$I_{\mathrm{os}} = \frac{1}{(1+(a_C I_{\mathrm{os}})^4)(c_\beta + k_\beta I)} \tag{6}$$

where a_C is a scaling constant. Equation 6 can be easily inverted as:

$$I = \frac{1}{k_\beta}\left[\frac{1}{I_{\mathrm{os}}(1+(a_C I_{\mathrm{os}})^4)} - c_\beta\right]$$

(a) (b) (c)

Figure 22: *An example of Van Hareten's algorithm with a frame of the RNL sequence: a) original f-stop 0. b) Tone mapped frame using cone model. c) Reconstructed frame using proposed iTMO.*

Van Hateren applied its TMO and iTMO to uncalibrated HDR movies and images, which were scaled by a harmonic mean. The results showed that the method does not need gamma correction, removes noise, and accounts for light adaptation, see Figure 22. The main drawbacks of the TMO are the introduction of motion blur in movies, and the limited dynamic range that can handle, 10,000:1, which causes saturation in very dark and bright regions. Finally, the author does not provide any study on companding and further quantisation for his TMO/iTMO.

4.6. Two-layer Coding Algorithm for High Dynamic Range Images Based on Luminance Compensation

A similar compression scheme to HDR-MPEG was proposed by Okuda and Adami [OA07] for compression of HDR images. Similar to JPEG-HDR and HDR-MPEG, this scheme is designed to be backward compatible. The main differences from HDR-MPEG are the presence of a minimisation step for optimising tone mapping parameters, the compression of residuals using wavelets, and the use of the Hill function for tone mapping and its analytic inverse instead of a tabled function. The Hill function is a generalised sigmoid function which is defined as:

$$L_{\mathrm{d}}(\mathbf{x}) = f(L_{\mathrm{w}}(\mathbf{x})) = \frac{L_{\mathrm{w}}(\mathbf{x})^n}{L_{\mathrm{w}}(\mathbf{x})^n + k^n}$$

where n and k are a constants that depend by the image. The inverse g of f is given by:

$$L_{\mathrm{w}}(\mathbf{x}) = g(L_{\mathrm{d}}(\mathbf{x})) = f^{-1}(L_{\mathrm{d}}(\mathbf{x})) = k\left(\frac{L_{\mathrm{d}}(\mathbf{x})}{1-L_{\mathrm{d}}(\mathbf{x})}\right)^{\frac{1}{n}}$$

The encoding is divided into a few steps, see Figure 23. Firstly, a minimisation process using the original HDR image is performed in the logarithm domain to match HVS perception and avoid outliers at high values. The error to minimise is given by:

$$E(I) = \sum_{\mathbf{x}\in I}\left(\log(L_{\mathrm{w}}(\mathbf{x})) - \log(g(L_{\mathrm{d}}(\mathbf{x})))\right)^2 \tag{7}$$

for determining n and k. The optimum solution is uniquely determined imposing the partial derivatives of E for k and n equal to zero, leading to:

$$k = \exp\left(\frac{\left(\sum_{\mathbf{x}} B(\mathbf{x})^2\right)\left(\sum_{\mathbf{x}} A(\mathbf{x})\right) - \left(\sum_{\mathbf{x}} B(\mathbf{x})\right)\left(\sum_{\mathbf{x}} A(\mathbf{x})B(\mathbf{x})\right)}{M\left(\sum_{\mathbf{x}} B(\mathbf{x})^2\right) - \left(\sum_{\mathbf{x}} B(\mathbf{x})^2\right)}\right)$$

$$n = \frac{M\left(\sum_{\mathbf{x}} B(\mathbf{x})^2\right) - \left(\sum_{\mathbf{x}} B(\mathbf{x})\right)^2}{M\left(\sum_{\mathbf{x}} A(\mathbf{x})B(\mathbf{x})\right) - \left(\sum_{\mathbf{x}} A(\mathbf{x})\right)\left(\sum_{\mathbf{x}} B(\mathbf{x})\right)}$$

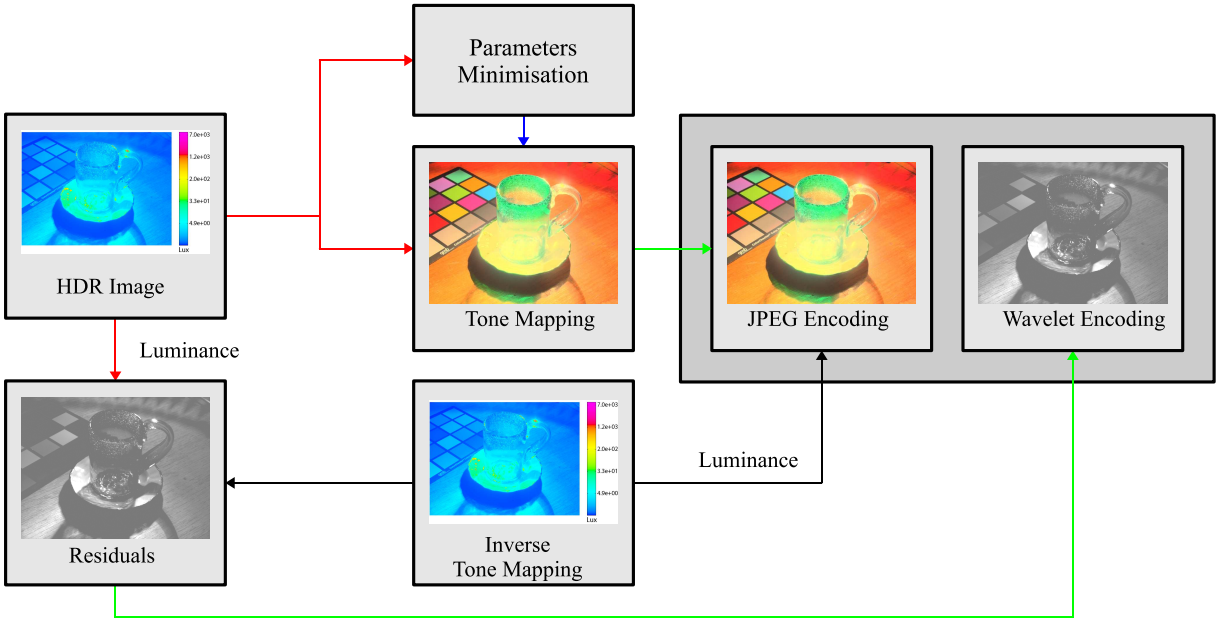

Figure 23: *The encoding pipeline presented in Okuda and Adami [OA07].*

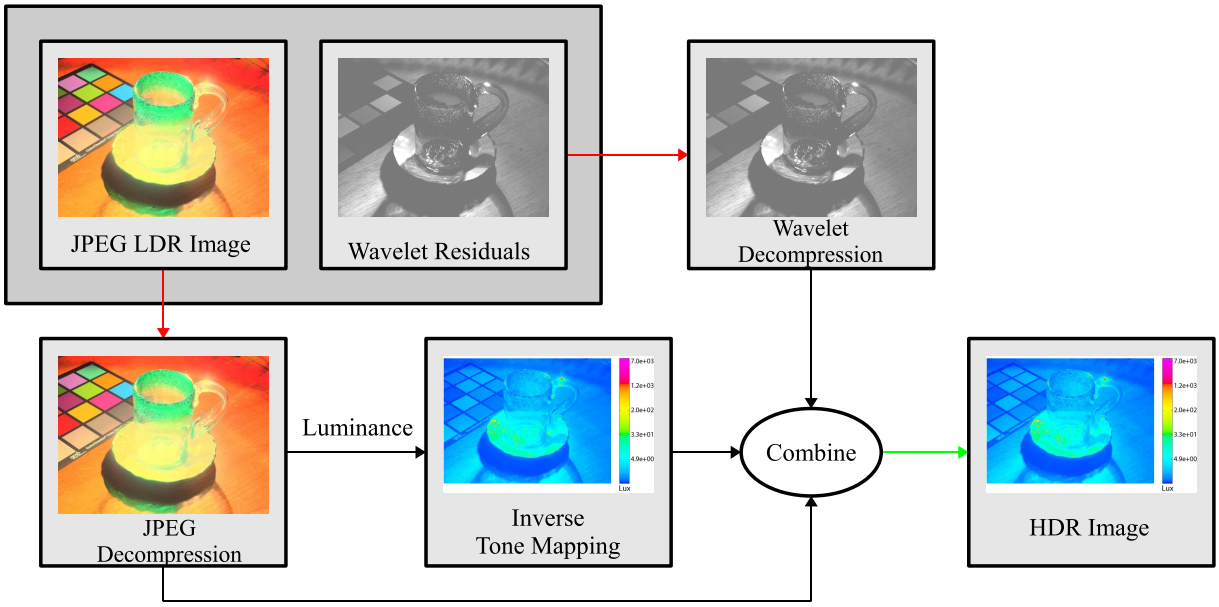

Figure 24: *The decoding pipeline presented in Okuda and Adami [OA07].*

where M is the number of pixels, and A and B are defined as:

$$A(\mathbf{x}) = \log L_{\mathrm{w}}(\mathbf{x}) \qquad B(\mathbf{x}) = \log\left(\frac{L_{\mathrm{d}}(\mathbf{x})}{1-L_{\mathrm{d}}(\mathbf{x})}\right)$$

Once the parameters are determined the image is tone mapped and encoded using JPEG. Then residuals are calculated to improve quality. They are calculated as:

$$R(\mathbf{x}) = \left(\frac{L_{\mathrm{w}}(\mathbf{x})}{g(L_{\mathrm{d}}(\mathbf{x}))+\varepsilon}\right)^{\gamma}$$

where $\gamma \in (0,1]$ is a constant, and ε is a small value to avoid discontinuities chosen by the user. Finally, R is encoded using a wavelet image compression scheme.

The decoding stage is straightforward. Once the LDR image and residuals are decoded using a JPEG decoder and a wavelet decoder, final HDR values are recovered by:

$$L_{\mathrm{w}}(\mathbf{x}) = R(\mathbf{x})^{\frac{1}{\gamma}}\left(g(L_{\mathrm{d}}(\mathbf{x}))+\varepsilon\right)$$

Two colour compensation methods are presented to preserve distortions caused by tone mapping. The first one is a modification of Ward and Simmons [WS05] where α and β are calculated with a quadratic minimisation using an error function similar to Equation 7. The second method is to apply a polynomial $P(x)$ for each LDR colour channel, assuming that a polynomial relationship exists between LDR and HDR values. Coefficients of $P(x)$ are fitted using the Gaussian weighted difference between the original and the reconstructed HDR channels.

The compression scheme was evaluated on a data set of 12 HDR images and compared with JPEG-HDR and HDR-MPEG using two metrics. The first one is the mean square difference (MSE) in CIELAB colour space [Fai05] to test overall quality:

$$MSE = \frac{1}{M}\sqrt{\sum_{\mathbf{x}\in I_1}\Delta_{Lab}(\mathbf{x})}$$

$$\Delta_{Lab}(\mathbf{x}) =$$

$$\left(L_{\mathrm{w},1}(\mathbf{x})-L_{\mathrm{w},2}(\mathbf{x})\right)^2+\left(a_1(\mathbf{x})-a_2(\mathbf{x})\right)^2+\left(b_1(\mathbf{x})-b_2(\mathbf{x})\right)^2$$

The second metric is MSE in the Daly's non-linearity domain [Dal93] to test reconstructed luminance:

$$L_{MSE}(I_1,I_2) = \frac{1}{M}\sqrt{\sum_{\mathbf{x}\in I_1}\left(Dn(L_{\mathrm{w}1}(\mathbf{x}))-Dn(L_{\mathrm{w}2}(\mathbf{x}))\right)^2}$$

$$Dn(x) = \frac{x}{x+12.6x^{0.63}}$$

In their experiments the proposed method achieved better results for both metrics in comparison with JPEG-HDR and HDR-MPEG at different bit rates. While the quality of this method is up to two times better than HDR-MPEG and JPEG-HDR at high bit rates (around 8-10 bits), it is comparable for low bit rates (around 1-4 bits).

4.7. HDR Texture Compression using Inverse Tone Mapping

A similar compression scheme to Okuda and Adami [OA07] was presented by Banterle et al. [BDLC08] for HDR texture compression. This method was designed to take advantage of graphics hardware. The generalised framework presented use a minimisation process that takes into account the compression scheme for tone mapped images and residuals. Moreover, it was shown that up-sampling of tone mapped values before expansion does not introduce visible errors.

Authors employed the global Reinhard et al. operator [RSSF02] and its inverse [BLDC06] in their implementation. The forward operator is defined as:

$$\begin{cases} f(L_{\mathrm{w}}(\mathbf{x})) = L_{\mathrm{d}}(\mathbf{x}) = \frac{\alpha L_{\mathrm{w}}(\mathbf{x})(\alpha L_{\mathrm{w}}(\mathbf{x})+L_{\mathrm{w,H}}L_{\mathrm{white}}^2)}{L_{\mathrm{w,H}}L_{\mathrm{white}}^2(\alpha L_{\mathrm{w}}(\mathbf{x})+L_{\mathrm{w,H}})} \\ \left[R_{\mathrm{d}}(\mathbf{x}),G_{\mathrm{d}}(\mathbf{x}),B_{\mathrm{d}}(\mathbf{x})\right]^{\top} = \frac{L_{\mathrm{d}}(\mathbf{x})}{L_{\mathrm{w}}(\mathbf{x})}\left[R_{\mathrm{w}}(\mathbf{x}),G_{\mathrm{w}}(\mathbf{x}),B_{\mathrm{w}}(\mathbf{x})\right]^{\top} \end{cases}$$

where L_{white} is the luminance white point, $L_{\mathrm{w,H}}$ is the harmonic mean, and α is the scale factor. While the inverse is given by:

$$\begin{cases} g(L_{\mathrm{d}}(\mathbf{x})) = f^{-1}(L_{\mathrm{d}}(\mathbf{x})) = L_{\mathrm{w}}(\mathbf{x}) = \\ = \frac{L_{\mathrm{white}}^2 L_{\mathrm{w,H}}}{2\alpha}\left(L_{\mathrm{d}}(\mathbf{x})-1+\sqrt{(1-L_{\mathrm{d}}(\mathbf{x}))^2+\frac{4L_{\mathrm{d}}(\mathbf{x})}{L_{\mathrm{white}}^2}}\right) \\ \left[R_{\mathrm{w}}(\mathbf{x}),G_{\mathrm{w}}(\mathbf{x}),B_{\mathrm{w}}(\mathbf{x})\right]^{\top} = \frac{L_{\mathrm{w}}(\mathbf{x})}{L_{\mathrm{d}}(\mathbf{x})}\left[R_{\mathrm{d}}(\mathbf{x}),G_{\mathrm{d}}(\mathbf{x}),B_{\mathrm{d}}(\mathbf{x})\right]^{\top} \end{cases}$$

The first stage of encoding is to estimate parameters of the TMO, similarly to [Rei02], and to apply a colour transformation, see Figure 25 for the encoding pipeline. However, this last step can be skipped because S3TC does not support colour spaces with separated luminance and chromaticity. Subsequently, the HDR texture and estimated values are used as input in a Levenberg-Marquadt minimisation loop which ends when the local optimum for TMO parameters is reached. In the loop, the HDR texture is firstly tone mapped and encoded with S3TC. Secondly, residuals are calculated and encoded using S3TC. Finally, the image is reconstructed, and error is calculated and new TMO parameters are estimated. When local optimum is reached, the HDR texture is tone mapped with these parameters and encoded using S3TC with residuals in the alpha channel.

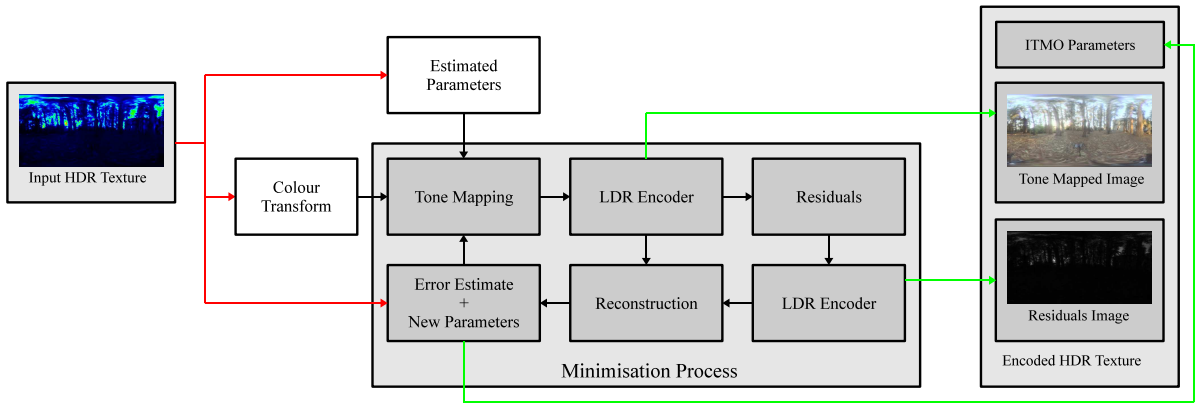

Figure 25: *The encoding pipeline presented in Banterle et al. [BDLC08].*

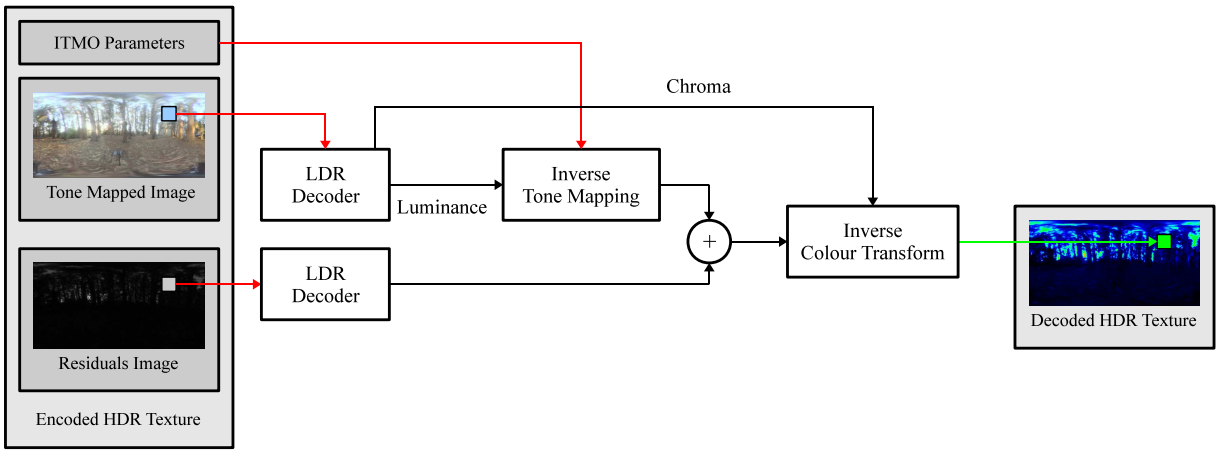

Figure 26: *The decoding pipeline presented in Banterle et al. [BDLC08].*

The decoding stage is straightforward and can be implemented in a simple shader on GPU, see Figure 26 for the decoding pipeline. When a texel is needed in a shader, the tone mapped texture is fetched and its luminance is calculated. The inverse tone mapping, uses these luminance values, combined with the TMO parameters, to obtain the expanded values which are then added to the residuals. Finally, luminance and colours are recombined. Note that the inverse operator can be pre-computed into a 1D texture to speed-up the decoding. Moreover, computations can be sped-up applying filtering during the fetch of the tone mapped texture. This is because the filtering is applied to coefficients of a polynomial function. Authors proposed a bound of this error, showing that is not significant in many cases.

This proposed scheme was compared to RGBE [War91], Munkberg et al.'s method [MCHAM06], Roimela et al.'s scheme [RAI06], and Wang et al.'s scheme [WWZ*07] using HDR-VDP [MDMS05], mPSNR [MCHAM06], and RMSE in the logarithm domain [XPH05]. The results showed that

(a) (b)

Figure 27: *A comparison of real-time decoding schemes on current graphics hardware applied to St. Peter's Cathedral lightprobe: a) Banterle et al.'s scheme [OA07]. b) Wang et al.'s one [WWZ*07] showing visible contouring artifacts.*

the new schemes presents a good trade-off between quality and compression, as well as the ability to decode textures in real-time. Moreover, it has a better quality on average than Wang et al.'s method, the other real-time decoding scheme, avoiding contouring artifacts, see Figure 27. The main disadvantage of this method is not being able to efficiently encode the luminance and chromaticity due to limits of S3TC.

4.8. Validation

The evaluation of quality for image compression schemes is usually performed using image metrics such as: HDR Visual Difference Predictor (HDR-VDP) [MDMS05], a perceptual metric, Root Mean Squared Error (RMSE) in the $\log_2$[RGB] domain [XPH05], and multi-exposure Peak Signal Noise Ratio (mPSNR).

4.8.1. Root Mean Squared Error in the $\log_2$[RGB] domain

The RMSE in the $\log_2$[RGB] domain was proposed by Xu et al. [XPH05], which is defined as follows:

$$RMSE(I,\hat{I}) = \sqrt{\frac{1}{n}\sum_{\mathbf{x}} \Delta_{RGB}(\mathbf{x})}$$

$$\Delta_{RGB}(\mathbf{x}) = \left(\log_2 \frac{R(\mathbf{x})}{\hat{R}(\mathbf{x})}\right)^2 + \left(\log_2 \frac{G(\mathbf{x})}{\hat{G}(\mathbf{x})}\right)^2 + \left(\log_2 \frac{B(\mathbf{x})}{\hat{B}(\mathbf{x})}\right)^2$$

where I is the reference image and (R,G,B) its red, green and blue channels, $\hat{I}$ the comparison image and $(\hat{R},\hat{G},\hat{B})$ its channels, n the number of pixels of the image. A small RMSE value means that image $\hat{I}$ is close to the reference, zero means that they are the same, while a high value means that they are very different.

4.8.2. mPSNR

mPSNR is an extension of PSNR metric to HDR domain by Munkberg et al. [MCHAM06]. This takes a series of exposures which are tone mapped using a simple gamma curve:

$$T(X,c) = \left[255(2^c X)^{\frac{1}{\gamma}}\right]_0^{255} \tag{8}$$

where c is the current f-stop, X is a colour channel, and $\gamma = 2.2$. Then the classic Mean Square Error (MSE) is computed:

$$MSE(I,\hat{I}) = \frac{1}{n \times p}\sum_{c=p_{\text{Min}}}^{p_{\text{Max}}}\sum_{\mathbf{x}}\left(\Delta R_c^2(\mathbf{x}) + \Delta G_c^2(\mathbf{x}) + \Delta B_c^2(\mathbf{x})\right) \tag{9}$$

where p_{Min} and p_{Max} are respectively the minimum and maximum exposures, p is the number of used exposures, n is the number of pixels in the image, $\Delta R_c(\mathbf{x}) = T(R(\mathbf{x}),c) - T(\hat{R}(\mathbf{x}),c)$ for the red colour channel, and similarly for green and blue channels. Finally, the m-SNR is calculated using the standard formula:

$$mPSNR(I,\hat{I}) = 10\log_{10}\left(\frac{3\times 255^2}{MSE(I,\hat{I})}\right) \tag{10}$$

5. Conclusions

This state-of-the art report has presented a comprehensive overview of the current research that expands LDR content for the generation of HDR images and videos. The LDR to HDR expansion methods fill the void between classic imaging and HDR imaging, allowing existing LDR content to be used for HDR applications. Even if HDR imaging has become very popular outside the computer graphics community, there are no still HDR cameras and video-cameras for the general consumer. A certain knowledge, that a general consumer is not expected to have, is needed to capture HDR images. Moreover, only capturing still images requires certain conditions such as static scenes, a tripod to avoid misalignments, and no variation in the lighting conditions. To meet all these conditions in real scenarios is very difficult and requires time and expertise.The discussed methods for LDR to HDR expansion result in a good compromise between HDR imaging and available camera technology.

We have also shown how LDR to HDR expansion methods and related techniques can be suitable for HDR compression. These methods' encoding procedure first uses a luminance compression step, generally via some tone mapping operation, followed by standard LDR compression. The decoding is performed via an LDR decoding stage, followed by an expansion, potentially the inverse of the tone mapping operator used. These compression methods are particularly useful because most of them are backwards-compatible and also, as compression methods improve in the LDR fields, there is an immediate and automatic benefit for such HDR compression techniques.

The main limitations of most LDR to HDR expansion methods occurs when trying to expand large over-exposed areas. This issue clearly depends on the size of over-exposed areas in the image or video. The quality is inversely proportional to the area of over-exposed regions since large over-exposed areas imply more information to reconstruct than smaller ones. As an example of reconstruction that highlights these limitations, using Banterle et al.'s method [BLDC06], see Figure 28. While the method of Wang et al. [WWZ*07] would be more suitable in such a situation, the manual method may be detrimental for many applications. Further research in this area would be required to tackle this problem. When considering video sequences, exposure changes in between frames, showing details in over-exposed or under-exposed areas which become well-exposed may be exploited as a solution to such problems. It may be possible

to project well-exposed areas from other (previous or successive) frames onto the current one where that same area is over-exposed or under-exposed.

(a) (b)

Figure 28: *An example of the limits of LDR to HDR expansion, using Banterle et al.'s LDR to HDR expansion [BLDC06]: a) The original LDR image. Note that it contains large over-exposed areas. b) The expanded image in a) at f-stop −3. The reconstruction produced a smooth gray gradient pattern, because there is not enough information in these areas.*

We have presented some validation methods for identifying the quality of the expansion methods. These validation methods currently only cater for still images and IBL applications and no validation study on expanded videos yet exists. As HDR video becomes more important some form of validation for video will naturally be required. Psychophysical studies on videos with complex stimuli, such as a shot from a movie, may not be easily run, possibly automated methods may provide a more straightforward solution.

HDR technology has not yet reached its full maturity. Capturing HDR videos is still an open issue. Moreover, capturing HDR still images can be a long and tedious task when using classic LDR technology, and is still expensive with HDR cameras. The presented methods for the reconstruction of HDR content from LDR content have made the capturing of HDR content for consumers a more straightforward process. Moreover, all LDR legacy content can be utilised for HDR media, or used for HDR processing and re-lighting real and synthetic objects.

6. Acknowledgement

We thank Ahmet Oğuz Akyüz, Rafał Mantiuk, Tunç Ozan Aydin, Piotr Didyk, Paul Debevec, Greg Ward, Edward H. Adelson, Allan Rempel, and Hans van Hateren for the LDR and HDR images used in our paper. We thank Piotr Didyk and Rafał Manituk for their insightful comments on the draft of this STAR. Finally we thank Stanford's Graphics Group for the Happy Buddha and Lucy models from the Stanford 3D repository.

This work reported in this STAR has formed part of EPSRC grants EP/D032148 and EP/G001634/1 whose funding and support are gratefully acknowledged.

References

[Ado07] ADOBE: Adobe photoshop.

[AFR*07] AKYÜZ A. O., FLEMING R., RIECKE B. E., REINHARD E., BÜLTHOFF H. H.: Do HDR displays support LDR content?: a psychophysical evaluation. In *SIGGRAPH '07: ACM SIGGRAPH 2007 papers* (New York, NY, USA, 2007), ACM, p. 38.

[AMMS08] AYDIN T. O., MANTIUK R., MYSZKOWSKI K., SEIDEL H.-P.: Dynamic range independent image quality assessment. In *SIGGRAPH '08: ACM SIGGRAPH 2008 papers* (New York, NY, USA, 2008), ACM, pp. 1–10.

[BA87] BURT P. J., ADELSON E. H.: The Laplacian pyramid as a compact image code. *Readings in computer vision: issues, problems, principles, and paradigms* (1987), 671–679.

[BDLC08] BANTERLE F., DEBATTISTA K., LEDDA P., CHALMERS A.: A GPU-friendly method for high dynamic range texture compression using inverse tone mapping. In *GI '08: Proceedings of graphics interface 2008* (Toronto, Ont., Canada, Canada, 2008), Canadian Information Processing Society, pp. 41–48.

[BLD*07] BANTERLE F., LEDDA P., DEBATTISTA K., CHALMERS A., BLOJ M.: A framework for inverse tone mapping. *The Visual Computer 23*, 7 (2007), 467–478.

[BLD*08] BANTERLE F., LEDDA P., DEBATTISTA K., ARTUSI A., BLOJ M., CHALMERS A.: A psychophysical evaluation of inverse tone mapping techniques. *To Appear in Computer Graphics Forum* (2008).

[BLDC06] BANTERLE F., LEDDA P., DEBATTISTA K., CHALMERS A.: Inverse tone mapping. In *GRAPHITE '06: Proceedings of the 4th international conference on Computer graphics and interactive techniques in Australasia and Southeast Asia* (New York, NY, USA, 2006), ACM, pp. 349–356.

[BLDC08] BANTERLE F., LEDDA P., DEBATTISTA K., CHALMERS A.: Expanding low dynamic range videos for high dynamic range applications. In *SCCG '08: Proceedings of the 4th Spring Conference on Computer Graphics* (New York, NY, USA, 2008), ACM, pp. 349–356.

[BN76] BLINN J. F., NEWELL M. E.: Texture and reflection in computer generated images. In *SIGGRAPH '76: Proceedings of the 3rd annual conference on Computer Graphics and Interactive Techniques* (New York, NY, USA, 1976), ACM, pp. 266–266.

[BVSO03] BERTALMIO M., VESE L., SAPIRO G., OSHER S.: Simultaneous structure and texture image inpainting. *IEEE Transactions on Image Processing 12*, 8 (August 2003), 882–889.

[CSE00] CHRISTOPOULOS C., SKODRAS A., EBRAHIMI T.: The JPEG2000 still image coding system: an overview. *IEEE Transactions on Consumer Electronics 46*, 4 (November 2000), 1103–1127.

[Cyp] CYPRESSSEMICONDUCTOR: Lupa 1300-2. *http://www.cypress.com/*.

[Dal93] DALY S.: The visible differences predictor: an algorithm for the assessment of image fidelity. *Digital images and human vision* (1993), 179–206.

[Dav88] DAVID H. A.: *The Method of Paired Comparisons, 2nd ed.* Oxford University Press, 1988.

[DD02] DURAND F., DORSEY J.: Fast bilateral filtering for the display of high-dynamic-range images. In *SIGGRAPH '02: Proceedings of the 29th annual conference on Computer graphics and interactive techniques* (New York, NY, USA, 2002), ACM, pp. 257–266.

[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 '98: Proceedings of the 25th annual conference on Computer graphics and interactive techniques* (New York, NY, USA, 1998), ACM, pp. 189–198.

[Deb05] DEBEVEC P.: A median cut algorithm for light probe sampling. In *SIGGRAPH '05: ACM SIGGRAPH 2005 Posters* (New York, NY, USA, 2005), ACM, p. 66.

[DF03] DALY S., FENG X.: Bit-depth extension using spatiotemporal microdither based on models of the equivalent input noise of the visual system. In *Proceedings of Color Imaging VIII: Processing, Hardcopy, and Applications* (June 2003), vol. 5008, 455, SPIE.

[DF04] DALY S., FENG X.: Decontouring: prevention and removal of false contour artifacts. In *Proceedings of Human Vision and Electronic Imaging IX* (June 2004), vol. 5008, 455, SPIE.

[DHT*00] DEBEVEC P., HAWKINS T., TCHOU C., DUIKER H.-P., SAROKIN W., SAGAR M.: Acquiring the reflectance field of a human face. In *SIGGRAPH '00: Proceedings of the 27th annual conference on Computer graphics and interactive techniques* (New York, NY, USA, 2000), ACM Press/Addison-Wesley Publishing Co., pp. 145–156.

[DM97] DEBEVEC P. E., MALIK J.: Recovering high dynamic range radiance maps from photographs. In *SIGGRAPH '97: Proceedings of the 24th annual conference on Computer graphics and interactive techniques* (New York, NY, USA, 1997), ACM Press/Addison-Wesley Publishing Co., pp. 369–378.

[DMAC03] DRAGO F., MYSZKOWSKI K., ANNEN T., CHIBA N.: Adaptive logarithmic mapping for displaying high contrast scenes. In *Proc. of EUROGRAPHICS 2003* (Granada, Spain, 2003), Brunet P., Fellner D. W., (Eds.), vol. 22 of *Computer Graphics Forum*, Blackwell, pp. 419–426.

[DMHS08] DIDYK P., MANTIUK R., HEIN M., SEIDEL H.-P.: Enhancement of bright video features for hdr displays. In *Proceeding of Eurographics Symposium on Rendering 2008* (2008), Computer Graphics Forum, Eurographics, Blackwell Ltd.

[Dol05] DOLBY: http://www.dolby.com/promo/hdr/technology.html, 2005.

[Fai05] FAIRCHILD M. D.: *Color Appearance Models 2nd Edition*. Wiley-IS&T, 2005.

[FLW02] FATTAL R., LISCHINSKI D., WERMAN M.: Gradient domain high dynamic range compression. In *SIGGRAPH '02: Proceedings of the 29th annual conference on Computer graphics and interactive techniques* (New York, NY, USA, 2002), ACM, pp. 249–256.

[FSPG97] FERWERDA J. A., SHIRLEY P., PATTANAIK S. N., GREENBERG D. P.: A model of visual masking for computer graphics. In *SIGGRAPH '97: Proceedings of the 24th annual conference on Computer graphics and interactive techniques* (New York, NY, USA, 1997), ACM Press/Addison-Wesley Publishing Co., pp. 143–152.

[Gre86] GREENE N.: Environment mapping and other applications of world projections. *IEEE Computer Graphics and Applications 6*, 11 (November 1986), 21–29.

[Hat06] HATEREN J. H. V.: Encoding of high dynamic range video with a model of human cones. *ACM Trans. Graph. 25*, 4 (2006), 1380–1399.

[HL06] HATEREN J. H. V., LAMB T. D.: The photocurrent response of human cones is fast and monophasic. *BMC Neuroscience 7*, 34 (April 2006).

[Hoe07] HOEFFLINGER B. (Ed.): *High-Dynamic-Range (HDR) Vision*, vol. 26 of *Springer Series in Advanced Microelectronics*. Springer, 2007.

[Hou81] HOUGH D.: Applications of the proposed IEEE-754 standard for floating point arithmetic. *Computer 14*, 3 (Mar. 1981), 70–74.

[Ind02] INDUSTRIAL LIGHT & MAGIC: OpenEXR. *http://www.openexr.org* (2002).

[ITU90] ITU: ITU-R BT.709, basic parameter values for the HDTV standard for the studio and for international programme exchange. In *Standard Recommendation 709, International Telecommunication Union.* (1990).

[Kaj86] KAJIYA J. T.: The rendering equation. *SIGGRAPH Computer Graphics 20*, 4 (1986), 143–150.

[KUWS03] KANG S. B., UYTTENDAELE M., WINDER S., SZELISKI R.: High dynamic range video. In *SIGGRAPH '03: ACM SIGGRAPH 2003 Papers* (New York, NY, USA, 2003), ACM, pp. 319–325.

[Lan02] LANDIS H.: Production-ready global illumination. In *Siggraph Course Notes 16* (2002).

[Lar98] LARSON G. W.: Logluv encoding for full-gamut, high-dynamic range images. *Journal of Graphics Tools 3*, 1 (1998), 15–31.

[LCTS05] LEDDA P., CHALMERS A., TROSCIANKO T., SEETZEN H.: Evaluation of tone mapping operators using a high dynamic range display. In *SIGGRAPH '05: ACM SIGGRAPH 2005 Papers* (New York, NY, USA, 2005), ACM, pp. 640–648.

[LGYS04] LIN S., GU J., YAMAZAKI S., SHUM H.-Y.: Radiometric calibration from a single image. In *CVPR 2004: Proceedings of the 2004 IEEE Conference on Computer Vision and Pattern Recognition (CVPR2004) - Volume 2* (Washington, DC, USA, 2004), IEEE Computer Society, pp. 938–945.

[LRP97] LARSON G. W., RUSHMEIER H., PIATKO C.: A visibility matching tone reproduction operator for high dynamic range scenes. *IEEE Transactions on Visualization and Computer Graphics 3*, 4 (1997), 291–306.

[LSA05] LI Y., SHARAN L., ADELSON E. H.: Compressing and companding high dynamic range images with subband architectures. *ACM Trans. Graph. 24*, 3 (2005), 836–844.

[Lub95] LUBIN J.:. World Scientific Publishers, 1995, ch. A Visual Discrimination Model for Imaging System Design and Evaluation, pp. 245–283.

[LZ05] LIN S., ZHANG L.: Determining the radiometric response function from a single grayscale image. In *CVPR '05: Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05) - Volume 2* (Washington, DC, USA, 2005), IEEE Computer Society, pp. 66–73.

[MCHAM06] MUNKBERG J., CLARBERG P., HASSELGREN J., AKENINE-MÖLLER T.: High dynamic range texture compression for graphics hardware. *ACM Trans. Graph. 25*, 3 (2006), 698–706.

[MDMS05] MANTIUK R., DALY S., MYSZKOWSKI K., SEIDEL H.-P.: Predicting visible differences in high dynamic range images - model and its calibration. In *Human Vision and Electronic Imaging X, IST SPIE's 17th Annual Symposium on Electronic Imaging* (2005), Rogowitz B. E., Pappas T. N., Daly S. J., (Eds.), vol. 5666, pp. 204–214.

[MDS06] MEYLAN L., DALY S., SÄIJSSTRUNK S.: The Reproduction of Specular Highlights on High Dynamic Range Displays. In *IST/SID 14th Color Imaging Conference* (2006).

[MDS07] MEYLAN L., DALY S., SÜSSTRUNK S.: Tone Mapping For High Dynamic Range Displays. In *Electronic Imaging* (2007), vol. 6492.

[MEMS06] MANTIUK R., EFREMOV A., MYSZKOWSKI K., SEIDEL H.-P.: Backward compatible high dynamic range mpeg video compression. In *SIGGRAPH '06: ACM SIGGRAPH 2006 Papers* (New York, NY, USA, 2006), ACM, pp. 713–723.

[MFS*08] MARTIN M., FLEMING R., SORKINE O., , GUTIERREZ D.: Understanding exposure for reverse tone mapping. In *CEIG 2008* (September 2008), Matey L., Torres J., (Eds.), pp. 1–9.

[MH84] MILLER ., HOFFMAN C.: Illumination and reflection maps: Simulated objects in simulated and real environments. In *Siggraph '84 Advanced Computer Graphics Animation seminar note* (New York, NY, USA, July 1984), ACM Press.

[MKMS04] MANTIUK R., KRAWCZYK G., MYSZKOWSKI K., SEIDEL H.-P.: Perception-motivated high dynamic range video encoding. In *SIGGRAPH '04: ACM SIGGRAPH 2004 Papers* (New York, NY, USA, 2004), ACM, pp. 733–741.

[MMK08] MYSZKOWSKI K., MANTIUK R., KRAWCZYK G.: *High Dynamic Range Video*. Synthesis Digital Library of Engineering and Computer Science. Morgan & Claypool Publishers, San Rafael, USA, 2008.

[MN99] MITSUNAGA T., NAYAR S.: Radiometric Self Calibration. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (June 1999), vol. 1, pp. 374–380.

[MP95] MANN S., PICARD R. W.: Being "undigital" with digital cameras: Extending dynamic range by combining differently exposed pictures. In *In Proceedings of IS&T 46th annual conference* (May 1995), pp. 422–428.

[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.

[Neu06] NEURICAM: Nc1805 - pupilla. *http://www.neuricam.com/* (2006).

[OA07] OKUDA M., ADAMI N.: Two-layer coding algorithm for high dynamic range images based on luminance compensation. *J. Vis. Comun. Image Represent. 18*, 5 (2007), 377–386.

[Omr07] OMROM: Fz3 series. *http://www.ia.omron.com/* (2007).

[Pan02] PANOSCAN: Panoscan mk-3, 2002.

[PGB03] PÉREZ P., GANGNET M., BLAKE A.: Poisson image editing. *ACM Trans. Graph. 22*, 3 (2003), 313–318.

[PtG04] PTGREYRESEARCH: Firefly mv. *http://www.ptgrey.com/* (2004).

[PTYG00] PATTANAIK S. N., TUMBLIN J., YEE H., GREENBERG D. P.: Time-dependent visual adaptation for fast realistic image display. In *SIGGRAPH '00: Proceedings of the 27th annual conference on Computer graphics and interactive techniques* (New York, NY, USA, 2000), ACM Press/Addison-Wesley Publishing Co., pp. 47–54.

[RAI06] ROIMELA K., AARNIO T., ITÄRANTA J.: High dynamic range texture compression. In *SIGGRAPH '06: ACM SIGGRAPH 2006 Papers* (New York, NY, USA, 2006), ACM, pp. 707–712.

[RBS99] ROBERTSON M. A., BORMAN S., STEVENSON R. L.: Dynamic range improvement through multiple exposures. In *Proceedings of the 1999 International Conference on Image Processing (ICIP-99)* (Los Alamitos, CA, 1999), pp. 159–163.

[RBS03] ROBERTSON M. A., BORMAN S., STEVENSON R. L.: Estimation-theoretic approach to dynamic range enhancement using multiple exposures. *Journal of Electronic Imaging 12*, 2 (April 2003), 219–228.

[Red06] REDCOMPANY: Read one. *http://www.red.com/* (2006).

[Rei02] REINHARD E.: Parameter estimation for photographic tone reproduction. *Journal Graphics Tools 7*, 1 (2002), 45–52.

[Rob62] ROBERTS L.: Picture coding using pseudo-random noise. *IEEE Transactions on Information Theory 8*, 2 (February 1962), 145–154.

[RSSF02] REINHARD E., STARK M., SHIRLEY P., FERWERDA J.: Photographic tone reproduction for digital images. In *SIGGRAPH '02: Proceedings of the 29th annual conference on Computer graphics and interactive techniques* (New York, NY, USA, 2002), ACM, pp. 267–276.

[RTS*07] REMPEL A. G., TRENTACOSTE M., SEETZEN H., YOUNG H. D., HEIDRICH W., WHITEHEAD L., WARD G.: Ldr2hdr: on-the-fly reverse tone mapping of legacy video and photographs. In *SIGGRAPH '07: ACM SIGGRAPH 2007 papers* (New York, NY, USA, 2007), ACM Press.

[RWPD05] REINHARD E., WARD G., PATTANAIK S., DEBEVEC P.: *High Dynamic Range Imaging: Acquisition, Display and Image-Based Lighting*. Morgan Kaufmann Publishers, December 2005.

[SDS95] STOLLNITZ E. J., DEROSE T. D., SALESIN D. H.: Wavelets for computer graphics: A primer. *IEEE Comput. Graph. Appl. 15*, 3 (1995), 76–84.

[SHS*04] SEETZEN H., HEIDRICH W., STUERZLINGER W., WARD G., WHITEHEAD L., TRENTACOSTE M., GHOSH A., VOROZCOVS A.: High dynamic range display systems. *ACM Trans. Graph. 23*, 3 (2004), 760–768.

[SLWL08] SUN W., LU Y., WU F., LI S.: Dhtc: an effective dxtc-based hdr texture compression scheme. In *GH '08: Proceedings of the 23rd ACM SIGGRAPH/EUROGRAPHICS symposium on Graphics hardware* (Aire-la-Ville, Switzerland, Switzerland, 2008), Eurographics Association, pp. 85–94.

[Sph02] SPHERON: Spheronhdr, 2002.

[Tho05] THOMSONGRASSVALLEY: Viper filmstream. *http://www.thomsongrassvalley.com/* (2005).

[Vap95] VAPNIK V. N.: *The nature of statistical learning theory*. Springer-Verlag New York, Inc., New York, NY, USA, 1995.

[Vis05] VISIONRESEARCH: Phantom hd. *http://www.visionresearch.com/* (2005).

[War91] WARD G.: Real pixels. *Graphics Gems 2* (1991), 15–31.

[Wat86] WATSON A.: *Temporal Sensitivity Chapter 6 in Handbook of Perception, Vol. I*. Springer Verlag, 1986.

[WB02] WANG Z., BOVIK A.: A universal image quality index. *IEEE Signal Processing Letters 9*, 3 (March 2002), 81–84.

[WS04] WARD G., SIMMONS M.: Subband encoding of high dynamic range imagery. In *APGV '04: Proceedings of the 1st Symposium on Applied perception in graphics and visualization* (New York, NY, USA, 2004), ACM Press, pp. 83–90.

[WS05] WARD G., SIMMONS M.: JPEG-HDR: A backwards-compatible, high dynamic range extension to JPEG,. In *CIC 13th: Proceedings of the Thirteenth Color Imaging Conference* (2005), The Society for Imaging Science and Technology.

[WSBL03] WIEGAND T., SULLIVAN G., BJONTEGAARD G., LUTHRA A.: Overview of the H.264/AVC video coding standard. *IEEE Transactions on Circuits and Systems for Video Technology 13*, 7 (July 2003), 560–576.

[WWS*07] WANG L., WANG X., SLOAN P.-P., WEI L.-Y., TONG X., GUO B.: Rendering from compressed high dynamic range textures on programmable graphics hardware. In *I3D '07: Proceedings of the 2007 symposium on Interactive 3D graphics and games* (New York, NY, USA, 2007), ACM, pp. 17–24.

[WWZ*07] WANG L., WEI L.-Y., ZHOU K., GUO B., SHUM H.-Y.: High dynamic range image hallucination. In *Proceedings of Eurographics Symposium on Rendering* (Jun 2007).

[XPH05] XU R., PATTANAIK S. N., HUGHES C. E.: High-dynamic-range still-image encoding in jpeg 2000. *IEEE Comput. Graph. Appl. 25*, 6 (2005), 57–64.