

An Investigation into High Dynamic Range Imaging Technologies

Yujie Mei, BSc

**Thesis Submitted to the University of Nottingham for the
Degree of Master of Philosophy**

School of Computer Science
The University of Nottingham, Nottingham, UK

March 2016

Abstract

This thesis studies high dynamic range imaging (HDR) technologies. It covers techniques for creating HDR radiance map from photographs, tone mapping HDR radiance map for display and the evaluation of the quality of tone-mapped images.

The influential technique introduced by Debevec and Malik has become the de facto standard for recovering high dynamic range radiance maps from photographs and has been widely used in research and commercial systems for over a decade. However, we have discovered an important defect in the original algorithm that will make this technique often fail to produce reasonable results in the extremely bright or dark regions of a scene. Therefore, we introduce a novel technique to correct this defect. Instead of the original algorithm where only pixel values from the photographs are used to guide the synthesis of the high dynamic range radiance map, we explicitly incorporate the shutter speed information of the camera. At each spatial pixel location, we estimate a “suitable shutter” that will make that location best exposed. A pixel’s contribution to the high dynamic range radiance value is not only a function of its value but also depends on the difference between shutter speed used to take the pixel and the estimated “suitable shutter” of that pixel. We also show that this new idea can be successfully used to directly fuse differently exposed photographs into a single low dynamic range image for display in conventional low dynamic range devices.

Then, we present a novel tone mapping framework. In this framework, firstly we introduce a tone mapping fidelity principle which explicitly stipulates that tone-mapped image data should not only be visually enhanced but should also stay faithful to the original image. Second, this principle naturally translates tone mapping into a constrained optimization problem where a two-term cost function, one measures the difference between the tone-mapped image and a visually enhanced version of the image, and the other measures the difference between the

tone-mapped image and the original image, is optimized. The relative weightings of the two terms in the cost function not only offers an insightful and simple mechanism to control the appearance of the tone-mapped image but also enables the introduction of spatially varying or uniform weighting functions thus unifying local and global tone mapping in a single framework.

The HDR image is not directly viewable and dynamic range compression will unavoidably lose information. A saliency map analyses the visual importance of the regions and can therefore direct the tone mapping operators to preserve the visual conspicuity of the regions that should more likely attract visual attention. Therefore, we present a novel tone mapping method - Saliency Modulated High Dynamic Range Image Tone Mapping (SMTM). In SMTM, we have developed a very fast algorithm to first compute the visual saliency map of the high dynamic range radiance map and then directly use the saliency of the local regions to control the local tone mapping curve such that highly salient regions will have their details and contrast better protected so as to remain salient and attract visual attention in the tone-mapped display. We present experimental results to show that SMTM provides competitive performances to state of the art tone mapping techniques in rendering visually pleasing low dynamic range displays. We also show that SMTM is better able to preserve the visual saliency of the HDR image and that SMTM renders high saliency regions to stand out to attract observers' attention.

Finally, to solve a difficult problem of evaluating tone mapping algorithms, we introduce a novel approach - pair comparison using Web 2.0 Technology. In this evaluation approach, we have developed a Web2.0 style system that enables Internet users from anywhere to evaluate tone-mapped HDR photos at any time. We adopt a simple paired comparison protocol, Internet users are presented a pair of tone-mapped images and are simply asked to select the one that they think is better or click a "no difference" button. These user inputs are collected in the web server and analysed by a rank aggregation algorithm which ranks the tone-mapped photos

according to the votes they received. The advantages of this approach include the potential of collecting large user inputs under a variety of viewing environments rather than limited user participation under controlled laboratory environments thus enabling more robust and reliable quality assessment. We also present data analysis to correlate user generated qualitative indices with quantitative image statistics which may provide useful guidance for developing better tone mapping operators.

List of Publications

1. **Yujie Mei** and Guoping Qiu, "Recovering High Dynamic Range Radiance Maps from Photographs Revisited: A Simple and Important Fix", the 7th International Conference on Image and Graphics (ICIG) 2013
2. **Yujie Mei**, Guoping Qiu and Kin-Man Lam, "Saliency Modulated High Dynamic Range Image Tone Mapping", Image and Graphics (ICIG), 2011 Sixth International Conference on pp.22-27, 12-15 Aug. 2011. doi: 10.1109/ICIG.2011.52
3. **Yujie Mei**, Guoping Qiu and Jiang Duan, "Evaluating HDR Photos using Web 2.0 Technology", Proceedings of SPIE Volume: 7867, SPIE Conference on Digital Photography VIII, 23. 27 January 2011, San Francisco, USA
4. **Yujie Mei**, Guoping Qiu, Kenneth Lam and Min Qiu, "Tone mapping HDR images using optimization: a general framework", ICIP2010, IEEE International Conference on Image Processing, September 2010, Hong Kong

Acknowledgements

Firstly, I would like to thank my supervisor Dr. Guoping Qiu for offering me the MPhil opportunity in School of Computer Science, the University of Nottingham, UK. He is always kind, helpful, and knowledgeable. He brings me to High Dynamic Range Imaging, this great field. I think my future work will surely focus on this field.

Thanks to all the members in IMA and VIPLAB. Special thanks to my colleagues and friends: Bozhi Liu, Jie Shu, Qian Zhang, Wenwen Bao, Hao Fu, Min Zhang, Mercedes Torres, Salvador Garcia Bernal, Bo Song, Chang Liu, and Minsheng Zhang.

Thanks to my parents and girlfriend in China, who have been giving me supports.

Declaration

I hereby declare that this thesis has not been submitted, either in the same or different form, to this or any other university for a degree.

Signature:

Contents

Abstract	I
List of Publications	IV
Acknowledgements	V
Declaration	VI
1 Introduction	1
2 Basics of High Dynamic Range Imaging	7
2.1 History of High Dynamic Range Imaging	7
2.2 Related Terms	9
“scene” and “image”	9
“colour model”	10
“High Dynamic Range Image”	10
“HD” and “HDR”	11
“WDR”, “FDR” and “HDR”	13
“Exposure Fusion”	14
2.3 Advantages of HDR Imaging	15
2.4 HDR Imaging Applications	17
2.5 Summary.....	17

3 HDR Imaging Workflow	18
3.1 Capture	18
3.2 Store	23
3.3 Display	25
3.4 Summary	26
4 Previous Tone Mapping Algorithms	27
4.1 Initial step	28
4.2 Global Tone Mapping Algorithms	29
4.2.1 Hierarchical Tone Mapping	29
4.2.2 Fast Tone Mapping	31
4.2.3 Other Global Tone Mapping Algorithms	34
4.2.3.1 Tumblin and Rushmeier’s Algorithms	34
4.2.3.2 Ward’s Algorithms	36
4.2.3.3 Ferschin et al’s Exponential Mapping	37
4.2.3.4 F. Drago et al’s adaptive logarithmic mapping	38
4.3 Local Tone Mapping Algorithms	39
4.3.1 Local Histogram Adjustment	39
4.3.2 Other Local Tone Mapping Algorithms	42
4.3.2.1 Gradient Domain	42
4.3.2.2 Bilateral Filtering	43
4.3.2.3 Tumblin et al’s Two Algorithms	43

5 Shutter Speed based HDR Radiance Maps Recovery	45
5.1 Introduction	45
5.2 Recovering High Dynamic Range Radiance Map Revisited	46
5.3 Camera Shutter Speed based Photograph Fusion	48
A. HDR Radiance Maps Recovery	49
B. Exposure Fusion	49
5.4 Experimental Results	50
5.5 Summary	51
6 A Comprehensive Optimization-Based Tone Mapping Algorithm	57
6.1 Introduction	57
6.2 Related works	58
6.2.1 Global Tone Mapping Curve Using Optimization	59
6.3 General tone mapping using optimization	60
6.4 Experimental Results	63
6.5 Summary	70
7 Saliency Modulated High Dynamic Range Image Tone Mapping	72
7.1 Introduction	72
7.2 Related work	73
7.3 Saliency modulated tone mapping	74
A. High Dynamic Range Image Saliency Map	75
B. Saliency Modulated Tone Mapping	76

7.4 Experimental Results	79
7.5 Summary	89
8 Pair Comparison Evaluating System for Tone Mapping Algorithms	90
8.1 Introduction	90
8.2 Related work	91
8.2.1 High dynamic range image display	91
8.2.2 Image quality metrics	91
8.2.3 Ranking with partial order and rank aggregation	92
8.2.4 Crowd sourcing, web2.0, social media, games with a purpose	93
8.2.5 Paired comparison and psychometrics	93
8.3 High dynamic range photo evaluation using web2.0	93
8.3.1 Technology Infrastructure and experiment design	93
8.3.2 Data collection and processing	94
8.4 Results and analysis	96
8.5 Summary	102
9 Conclusions and Future Work	104
9.1 Contributions	104
9.2 Future work and closing comment	105
Bibliography	107

Chapter 1

Introduction

Human eye is able to discriminate a considerable range of light intensity levels. However, most digital cameras and computer monitors can only represent the real world's scene within an extremely limited dynamic range.

The term “**dynamic range**” is defined as the ratio between the highest and lowest luminance level. It is measured as a ratio, or as a base-10 (decibel) or base-2 (bits) logarithmic value. The term “luminance” is defined as the intensity of humans' perceived brightness from real world scene. For images, dynamic range is the ratio between the possible lightest and darkest pixel defined by image file format. Therefore, the dynamic range of common image file formats is always 256:1 (around two orders of magnitude in each colour channel), such as JPEG (Joint Photographic Experts Group), BMP (Windows bitmap), GIF (Graphics Interchange Format), and PNG (Portable Network Graphics). Some medical image file formats may contain more dynamic range (2^{16} :1 in each colour channel), such as DICOM (Digital Imaging and Communications in Medicine), TIFF (Tagged Image File Format) and so on. Moreover, raw image formats (RAW) available on digital cameras and RGBE (Radiance HDR) mentioned in this thesis both have more dynamic range. Except for RGBE HDR image format, the rest of these image file formats are all discrete. In other words, their dynamic range has defined how many

levels of colour they can employ. While for HDR images, the data stored in them is not integer but float point numbers, therefore HDR images can linearly record the real world's light information without any compression.

For a display, the dynamic range is the full-on and full-off luminance ratio. For cameras, the dynamic range is intensity response range of a sensor. However, the maximum number of simultaneously emitted luminance levels is still 256. That means at a same time, a display can mostly represent only 256 possible levels in each colour channel even its dynamic range is 100,000:1. Similarly, the dynamic range of the final images produced by a camera is 256:1 using the common image file formats, or $2^{12} \sim 2^{14}$ using RAW image format.

For the human visual system, the simultaneous luminance adaptation range is around 5 orders of magnitude, and the overall adaptation range can reach 10 orders of magnitude [40]. For real world scenes, its dynamic range ratio often exceeds 100,000:1. For example, in a room with a window, the luminance intensity outside the window is much higher than that inside the room. However, we can still detect the details both inside and outside the window owing to the adaptive capacity of our human visual system (HSV).

Figure 1.1 shows the light intensity of various light sources. Maximum intensity of common CRT monitors is 10^2 [4].

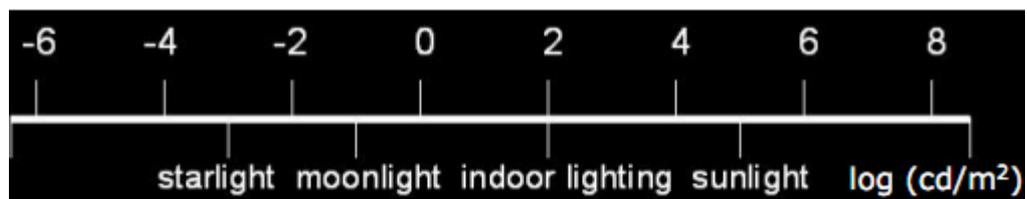


Figure 1.1: the light intensity of various light sources

Because of the common image file formats' limited dynamic range, lots of light information is always lost at capture time. The emergence of High Dynamic Range (HDR) Imaging technique has changed this. High Dynamic Range (HDR) Imaging

is an attractive way to capture the real world scene, since it allows recording the complete information on each pixel's luminance. HDR image owns much higher dynamic range than the traditional digital image, and much more accurately represents the luminance range from sunlight to darkness in night.

In Figure 1.2, the best exposed photos of a real world scene are shown on the left. In these scenes, light and dark areas both exist but cannot be recorded clearly together. As a result, bright areas are overexposed or dark areas are underexposed with few details. The images shown on the right are captured, stored, and represented with high dynamic range imaging techniques. In other words, these images are tone-mapped high dynamic range (HDR) images. It is seen that both light and dark areas of the images on the right have improved compared to those on the left.



Figure 1.2: Scenes captured in Nottingham. The left ones are photos taken by Nikon D700, and the right ones are HDR images' tone mapping results.

In recent years, HDR imaging technique has become the hot topic in the field of research and digital image application. It can be regarded as a lossless record of the real world scenes. However, an important problem it has to face is how to represent the HDR image data with huge range of luminance on the conventional display medium with limited dynamic range, such as hardcopy prints, monitors and projectors. Although a few high dynamic range display systems have been introduced, most of them only can be found in laboratory because of the high cost. Moreover, compared with traditional display devices, these HDR display systems only present a bit more dynamic range (for example, 12bits in Dolby PRM-4200, shown in figure 1.3). Therefore, these so-called HDR displays still need to solve the problem of how HDR images can be better represented. Therefore, a category of image processing technique named tone mapping or tone reproduction arises, which deal with the different dynamic ranges of the HDR images and conventional display devices.



Figure 1.3: Dolby PRM-4200

Tone mapping is an image processing and computer graphics technique used to present high dynamic range images approximately in a medium with limited dynamic range. Essentially, tone mapping strongly compress the high dynamic range of the real world scenes to displayable range, meanwhile it needs to ensure its fidelity and authenticity.

This thesis studies high dynamic range imaging techniques for radiance map generation, tone-mapped and quality evaluation. It is organized as follows:

Chapter 2 introduces the fundamental knowledge of high dynamic range imaging.

Chapter 3 describes the high dynamic range imaging workflow, including recovering HDR radiance maps, HDR radiance maps storage and RGBE file format, HDR image display and tone mapping.

Chapter 4 reviews previous tone mapping algorithms. These algorithms could be divided into two categories: global tone mapping algorithms and local tone mapping algorithms (also termed tone reproduction curve (TRC) based and tone reproduction operator (TRO) based).

Chapter 5 We discovered a defect in the excellent technique introduced by Debevec and Malik for recovering high dynamic range radiance maps from photographs, and propose a novel technique to correct this defect. In the original algorithm, only pixel values from the photographs are used to guide the synthesis of the high dynamic range radiance map; while in ours, each frame's shutter speed information and pixel values are applied together to make sure well exposed pixels get larger weighting, and produce more satisfactory results.

Chapter 6 presents a novel tone mapping framework. In this framework, a tone mapping fidelity principle is introduced: tone-mapped image data should stay faithful to the original high dynamic range image data. This principle naturally translates tone mapping into a constrained optimization problem where a two-term cost function, one measures the difference between the tone-mapped image and a visually enhanced version of the image, and the other measures the difference between the tone-mapped image and the original image, is optimized.

Chapter 7 High dynamic range image data is not directly visible but its visual attributes (such as saliency map) could be computed. Therefore, we introduce the

saliency map for HDR image data to create the local tone mapping curve so that highly salient regions will keep more details and contrast to remain salient, while lowly salient regions will has less details and contrast to avoid noises artefacts.

Chapter 8 introduces a novel approach - pair comparison using Web 2.0 technology for evaluating tone mapping algorithms. In this approach, a Web2.0 style system is implemented for any users to take part in the evaluation anywhere and anytime. The web server will collect users input data and analyse them by a rank aggregation algorithm.

Chapter 9 concludes the thesis and proposes future works.

Chapter 2

Basics of High Dynamic Range Imaging

In this chapter, the background information of high dynamic range imaging will be provided to form the basis of later chapters.

2.1 History of High Dynamic Range Imaging

In the early days, traditional photography is unable to record the real world's scene with high dynamic range, such as sky and sea on sunny days. The luminosity range between them is too extreme, therefore once sea is recorded, sky will be totally white; once the cloud in sky is recorded, the sea will be dark with nothing clear. In 1850s, Gustave Le Gray thought up an idea: he recorded the negative of the sky, and recorded the sea with long exposure, and combined them into one photo in positive as shown in Figure 2.1[48]. This can be considered to be one of the examples using HDR related techniques to record scenes of our real life.



Figure 2.1: Sky and sea photo by Gustave

Later in mid-twentieth century, a manual tone mapping technique – dodging and burning was introduced. Photographers selectively increase or decrease the exposure of the negative’s each region. Burning: they hold a card or other opaque object between the print and enlarger lens to allow light only fall on the selected dark regions as shown in left-top on Figure 2.2; Dodging: they hold a card or other opaque object to avoid light to fall on the bright regions as shown in left-bottom on Figure 2.2. After this processing, the dynamic range of the negative is compressed to be available on the final positive paper print. The work “Schweitzer at the Lamp” by W. Eugene Smith as shown in Figure 2.2 is a classical instance using this technique [Durand2001]. This photo cost him five days in the darkroom to reproduce the wide dynamic range between bright lamp and the dark regions in the room.

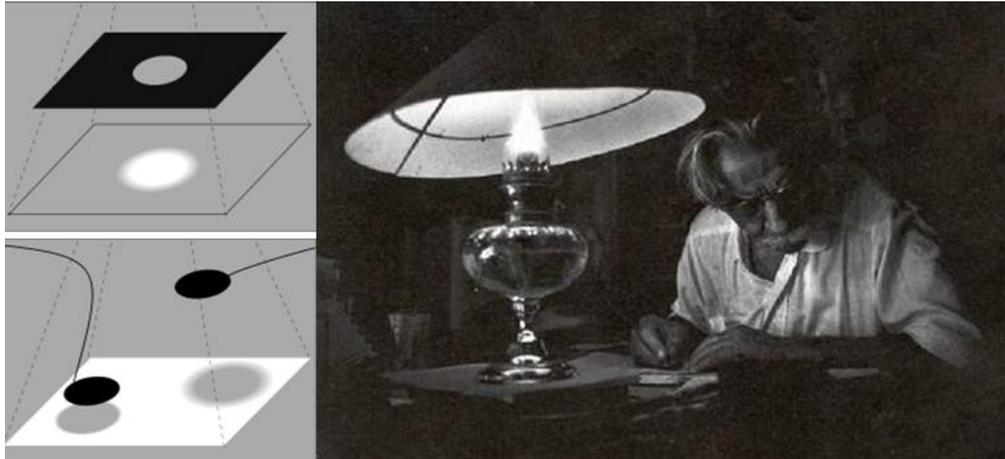


Figure 2.2: dodging and burning and “Schweitzer at the Lamp” by W. Eugene Smith

In the late 1980s High Dynamic Range (HDR) imaging has been introduced and in 1985 Gregory Ward created the Radiance RGBE image file format as the first (and still the most widely used) HDR imaging file format [30]. HDR Imaging file provides much more dynamic range information than the current imaging files and can display all parts of the scene clearly. Currently, RGBE is the most widely-used HDR imaging format. Beyond that, lots of other HDR image encodings are also employed: in 1995, Steve Mann and Rosalind Picard introduced their Global HDR in [75]; Adobe introduced TIFF Float format in 1992 [1] and in 2005, they released the CS2 version of Photoshop first time with HDR image support; in 2002, Industrial Light and Magic introduced their EXtended Range format [18].

2.2 Related Terms

It is necessary to explain a few related terms, because they are very important throughout the thesis.

“scene” and “image”

A scene can be described by its appearance and what impression it makes on people. What we see or we cannot see but existed in our real world, are all scenes. Image is the reproduction of these scenes. Images can be obtained through photography,

painting, computer software, and so on.

“colour model”

A colour model is an abstract mathematical model to describe the colour as a group of values. It is the basis of digital images. In this thesis, RGB and YUV colour model will be employed. RGB colour model comes from the three additive primary colours, red, green, and blue. This model uses different proportions of these three colours mixed to generate a variety of other colours. HDR images encoding and decoding mentioned in this thesis, employed similar colour model. But, unlike the traditional RGB model using integers, HDR images' RGB model use floating-point numbers.

YUV model is interchangeable with RGB. Its most outstanding feature is that it defines a luma (Y) value determining the luminance information, and another two values (chrominance) determining the colour information. In this thesis, we often deal with just the information on the brightness, but maintaining the colour information (RGB's scale) consistent to original, so YUV colour model is employed.

“High Dynamic Range Image”

Current HDR digital image has four main characteristics:

1. HDR image data is stored using floating point numbers. Relative to the conventional digital image (Low Dynamic Range Image, LDRI) stored using integer, and normally just 8bits (0 ~ 255) * 3 channels (RGB) to represent each pixel, HDR image uses 4 bytes floating point numbers to store the luminance information, and its represented range is extended up to $0 \sim 2^{127}$.
2. It linearly records the real world's radiance. The conventional digital image is obtained from a camera sensor with a nonlinear response function, which converts scene irradiance to the intensity values. And lots of scenes' detail information will always be lost after this response function. For example, the

white and bright areas will be both recorded as 255; the black and dark areas will be both recorded as 0.

3. There is no sensor could capture HDR image directly. Current so-called HDR camera could just capture images with rich tone or with relatively wider dynamic range (stored as raw file), but it is not the real HDR image at all. Nowadays, a widely used method to obtain HDR image is, capture a series of photos under different exposure, and then recover radiance maps from them. Each differently exposed photo will preserve different parts' light information. The more source images there are the more accurate HDR image data will be.
4. HDR image cannot be displayed directly on the conventional monitors. Always, before displaying, HDR image will be tone-mapped to the traditional 8bit per channel image. Moreover, a few tone mapping algorithms will generate surreal images as show in Figure 2.3, but whatever, these tone-mapped results will always represent more details than the traditional images.



Figure 2.3: Surreal tone-mapped result from High Dynamic Range Image

“HD” and “HDR”

To people who have never come into contact with HDR, they may be apt to confuse the two terms. HD (High Definition) depends from image's pixels numbers, and

decides image's clarity. In Figure 2.4, the second and third image is down sample of the first one, with different resolution. It clearly shows that, with higher resolution, image presents more details.



Figure 2.4: Same scene with different resolution: 320X240, 160X120, and 80X60

Currently, sensor's pixel number has reached the enough point so that people have gradually shifted focus to the imaging results from the resolution. This focus is actually the range of values that each pixel may represent. With traditional image's three bytes per pixel, each pixel can represent around 16.7 million different colours. This is known as "millions of colours." At first glance, this is an impressively large number. However, for each colour channel only 256 values will be employed, and it is obviously insufficient to represent many scenes. Now, we have the HDR technology to record these scenes, with theoretic 2^{127} values (4 bytes floating point numbers per pixel) per pixel. Figure 2.5 shows four images with different colour depth (each pixel's available colour numbers). It clearly shows that with more colour depth, each pixel can be represent with more colours, and images present more colour level. In summary, "HD" is for high resolution, and "HDR" is for colour depth (or dynamic range).



Figure 2.5: Four images with different colour depth: 8bit, 6bit, 4bit, and 3bit

“WDR”, “FDR” and “HDR”

“WDR” is short for “Wide Dynamic Range”. Any image with higher dynamic range than the traditional LDR image can be called WDR. For example, those images with more than 8 bits per colour channel: TIFF, JPEG, medical images, camera RAW file and so on. A few wide dynamic range display system which could only be found in laboratory may display these images, but most conventional display cannot represent them directly before tone mapping.

“FDR” is short for “Full Dynamic Range Image”. This is the same term as “HDR” in this thesis. But normally we call “HDR” more.

Nowadays, a lot of so-called HDR photos have appeared in various places, but actually they should be called “tone-mapped HDR images”. Although they are essentially the LDR images, but they are the tone-mapped result from HDR images. HDR tone mapping techniques have been used to reconstruct the whole image to represent more details within limited dynamic range. Also, there are lots of other so-called HDR photos obtained from WDR images’ tone mapping or exposure

fusion technique. Compared to HDR images' tone mapping results, they contain less details, but with more noise and unreal colour to hide the truth that they lack sufficient original image data.

In addition, a lot of so-called HDR products are now available in the market. They are actually the normal development of camera sensors. In contrast with the past, current high-end camera sensors are capable of sensing a wider range of light and the response function has also been optimized, therefore they have better imaging results. However, no matter how rapidly the camera technology develops, in recent years its sensing range is still difficult to reach the same level as human visual system. For some scenes with dark and bright regions existing together, people can see both clearly while these products still cannot record both because the dynamic range of the real world scene exceeds these products' capabilities

“Exposure Fusion”

Sometimes people just need the tone-mapped results, or just the imaging result better than the traditional ones. Then, a new application field-“exposure fusion” derived from HDR technology is introduced. It can rapidly generate a result similar to HDR image tone mapping result, but with less data processing. The HDR function of iPhone and Android phones are just this. Now we compare HDR technology with exposure fusion:

1. Ways of obtaining the images: Both of them need several source images from the same scene under different exposure. For HDR image, always it needs more than 7, to make its data meaningful; For exposure fusion normally 2 or 3 is needed (EV-1, EV+1 or EV-1, EV+0, EV+1). Then, HDR image will be obtained by radiance maps recovering algorithms; while exposure fusion technique will directly generate a traditional LDR image. It is obvious that HDR image will own more information with more source image data.
2. Result: HDR image will have much larger file size, while exposure fusion result

will be the same as the traditional image. HDR image could generate satisfactory result by different algorithms and it is possible for future processing; image generated by exposure fusion will not retain the source image, and it is one-off process.

In summary, exposure fusion technique derived from HDR technology improves the existing imaging techniques, but for higher imaging requirement HDR is needed.

2.3 Advantages of HDR Imaging

As shown in Figure 1.2, the best exposed photos of a real world scene are shown on the left. In these scenes, light and dark areas both exist but cannot be recorded clearly together. As a result, bright areas are overexposed or dark areas are underexposed with few details. The images shown on the right are captured, stored, and represented with HDR technology. It is seen that both light and dark areas' details have been retained well. It should be noticed the fact that the dynamic range of both sides' images is the same. Thus, using HDR imaging still has advantage even though there is a lack of display device capable of displaying HDR images directly.

HDR image has at least four advantages relative to LDR image:

1. HDR image could record the real luminance information as shown in Figure 2.6. Left image is the tone mapped result, and according to actual demand, we may get more details using different tone mapping operators; Middle image is obtained by mapping the HDR original image data to the right colour bar. It is seen that the middle image obviously represent the luminance information more close to the real world (red as bright, and blue as the dark).



Figure 2.6 Luminance information

2. HDR image could differentiate white and bright area; it could differentiate black and dark area. Because its image data is from different exposed result from the scene, low exposed source images will provide the bright area image data, while high exposed source images will provide the dark area image data. Figure 2.7 shows the conventional camera's result and the tone mapping result of HDR image. From it, we can see the white area in the left image is clearly seen in the HDR image's result.



Figure 2.7: conventional camera's result and HDR image's tone mapping result

3. Improved the image's dynamic range. In theory, HDR image's dynamic range is the same as the real world scene.
4. HDR image can be used to create amazing art work as shown in Figure 2.3.

2.4 HDR Imaging Applications

HDR imaging has wide applications including:

1. Because HDR image provides solid image data foundation, it is well suited for photography with high post-processing requirement, like professional photography, advertising photography, and so on.
2. Digitally record valuable articles like cultural treasures.
3. Digitally record important scenes.
4. Other application fields which need precise image data like monitoring.

In summary, this chapter introduces the history, advantages, and applications of high dynamic range imaging. This chapter also helps distinguish “scene” and “image”, “HDR” and “HD”, HDR image and current so-called HDR image. Moreover, this chapter introduces what is HDR, what is Exposure fusion, and also introduces the colour models employed in this thesis.

Chapter 3

HDR Imaging Workflow

The whole workflow of high dynamic range imaging contains three main stages: capture, storage, and display. This chapter will review the techniques used to capture high dynamic range image data, the file formats used to store high dynamic range image data, and the methods of displaying the high dynamic range images.

3.1 Capture

High dynamic range (HDR) image may be obtained from real world scenes and computer graphics (CG) rendering techniques. Right now most high-end graphics cards are able to generate HDR images directly. [19], [27], [35], [45], and [65] covered the topics of CG rendering very well. In this thesis, we only focus on the HDR image capture from real world scenes using conventional camera. This section will introduce the methods using to recover radiance map (HDR image data).

Traditional image data captured from conventional digital cameras is always stored as two-dimensional integer array with range 0~255 per each colour channel for recording intensity values. Lots of the scenes' detail information is lost during capture because of the non-linear mapping function in the cameras that converts scene irradiance to the intensity values in images. Moreover, the lost information is

impossible to be recovered from the resulting image data any more.

HDR image data (so-called radiance map) is a two-dimensional floating point values array, and each value records the relative radiance value in the real world scene. Comparing with the traditional LDR image data, radiance map is obviously able to records the scene's luminance information more accurately. Radiance map could be used to record high-quality scenes from the real world, and to reproduce the traditional 8-bits LDR digital image using tone mapping techniques. A radiance map could be considered as a lossless way to record the original scene.

As mentioned in Chapter 2, currently there is no sensor could capture HDR image directly. The most widely used approach is to take a sequence of different exposed low dynamic range (LDR) photos in the original scene and then recover radiance maps from them. High exposed photos will well preserve the information in dark regions, while low exposed photos will preserve more information in the bright regions. Figure 3.1 shows the process (just for illustrating, always 7 or more photos will be taken for meaningful HDR image data).

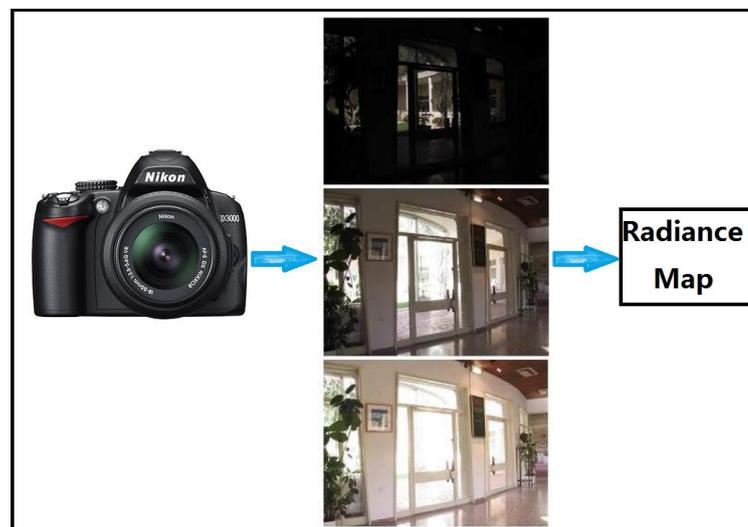


Figure 3.1: The process of producing high dynamic range map

Although in the hardware, there is no single sensor could capture HDR image, recently a few HDR imaging systems using multiple-sensor structure or with built-in

multiple-expose and recover radiance maps have been developed, such as [3], [77] and [83]. We also developed a HDR software - HDR Photographer, which could control Nikon SLR camera to capture HDR image directly. This is a full-function HDR software system, its functions contain: 1. HDR radiance map recovery (save as RGBE image format); 2. Directly capture HDR Image using Nikon DSLR and output the .hdr file; 3. HDR image tone mapping. This software supports most Nikon DSLR, like D3, D300, D700, D300S, and D3S. Its user interface is shown as Figure 3.2. Notice that, all source images should be registered before used to recover HDR radiance maps. My current work focus mainly not on this, therefore in this thesis it is assumed that all images are already registered.



Figure 3.2: HDR Photographer developed by the author

In a typical imaging system, image irradiance E is linearly related to scene radiance. The exposure X is related to irradiance E as shown in equation (3.1):

$$X = E * \Delta t \quad (3.1)$$

where Δt is the exposure time. The non-linear mapping relating the exposure X to the final digital value Z is usually referred as the response function (3.2) of a camera f :

$$Z = f(X) = f(E * \Delta t) \quad (3.2)$$

Then, the inverse of the function can be defined as:

$$X = f^{-1}(Z) = g(Z) \quad (3.3)$$

Once the response function of the imaging system is known, it is possible to estimate the radiance map based on a set of images captured under different exposures. Therefore recovering the radiance map of the scene becomes the question of how to recover the response function of the imaging system. In the following, Paul E. Debevec and Jitendra Malik's Method will be introduced [66].

In this method, the inputs are a sequence of differently exposed images with the accurate known exposure duration. Denote pixel values as Z_{ij} where i represents a spatial index and j represents images' indexes over exposure times Δt_j . The response function f can be written as:

$$Z_{ij} = f(E_j \Delta t_j) \quad (3.4)$$

The inverse of the function can be written as:

$$f^{-1}(Z_{ij}) = E_j \Delta t_j \quad (3.5)$$

Then, take the natural logarithm for each side,

$$\ln f^{-1}(Z_{ij}) = \ln E_j + \ln \Delta t_j \quad (3.6)$$

Define $g = \ln f^{-1}$, then:

$$g(Z_{ij}) = \ln E_i + \ln \Delta t_j \quad (3.7)$$

Now, In equation (3.7), Z_{ij} and Δt_j are already known, and recovering the camera's response function is equal to the recovery of the irradiance E_i together with $g(Z)$ as shown in Figure 3.3.

$$g(Z_{ij}) = \ln E_i + \ln \Delta t_j$$

Known: Z_{ij} and Δt_j

Unknown: g and E_i

Recovering the camera's response function

=

To the recovery of the irradiance E_i together with $g(Z)$ function

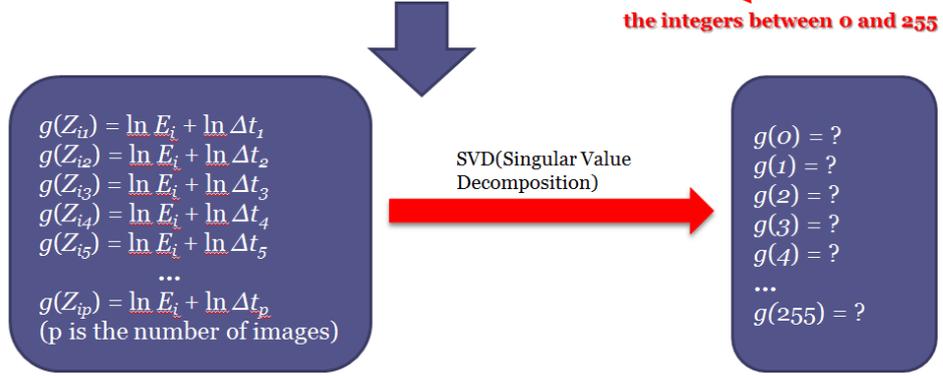


Figure 3.3: Solve the simultaneous equations to obtain the response function

To solve this problem, SVD (Singular Value Decomposition) algorithm is used to defined function $g(Z)$. In theory, when the function $g(Z)$ is solved, each pixel's irradiance in each image can be computed using the following equation:

$$\ln E_{ij} = g(Z_{ij}) - \ln \Delta t_j \quad (3.8)$$

Then, in order to achieve a better estimation of irradiance, a reasonable solution is to use all available exposures as a weighting. Equation (3.9) is applied in Paul E. Debevec's method:

$$\ln E_i = \frac{\sum_{j=1}^P w(Z_{ij})(g(Z_{ij}) - \ln \Delta t_j)}{\sum_{j=1}^P w(Z_{ij})} \quad (3.9)$$

where P is the number of images, and $w(Z)$ is defined as:

$$w(Z) = Z - Z_{min} \quad \text{for } Z \leq \frac{1}{2}(Z_{min} + Z_{max}) \quad (3.10)$$

$$w(Z) = Z_{max} - Z \quad \text{for } Z > \frac{1}{2}(Z_{min} + Z_{max}) \quad (3.11)$$

Apart from the method of [66], other techniques for capturing HDR images have also been developed e.g., [75] and [80]. More recently, technology has also been

developed to capture high dynamic range videos [72], but is immature.

3.2 Store

The traditional low dynamic range image pixel is represent by only 256 discrete integer values per colour channel, while high dynamic range image pixel could employ a range up to $0 \sim 2^{127}$ with floating point values. At present, for efficient storage, a few HDR image encoding techniques have been introduced, such as IEEE TIFF Format [1], RGBE Format [29], SGI LogLuv TIFF [27], and EXR Format [18]. And the RGBE format is the most commonly used.

The Radiance RGBE picture file format proposed by Greg Ward [29] stores pixels in 32 bits (4 bytes). It includes 1 byte mantissa for each colour channel (red, green, and blue) and 1 byte for a shared exponent. This file format's most important feature is that it allows having pixels to have precision of floating point values and the extended range. Because of the shared exponent use, RGBE format gains the advantage that it does not need 12 bytes for each pixel (in IEEE TIFF Format) and 4 bytes is enough. The RGBE file data is always compressed using run-length encoding.

The transformation formula between RGBE and HDR image floating point data (RGB) is as following:

1. RGBE \rightarrow HDRI Floating Point

If $e = 0$, $R = G = B = 0$, otherwise:

$$R = r * 2^{(e-128-8)}$$

$$G = g * 2^{(e-128-8)}$$

$$B = b * 2^{(e-128-8)}$$

2. HDRI Floating Point \rightarrow RGBE

Define $v = \max(R, G, B)$

If $v < 10^{-32}$, $r = g = b = e = 0$, otherwise:

Transform v to base 2 (bits) logarithmic value:

$v = m * 2^n$ ($0 < m < 1$), then:

$$r = R * m * 256 / v$$

$$g = G * m * 256 / v$$

$$b = B * m * 256 / v$$

$$e = n + 128$$

where r, g, b, e is the RGBE file data, and R, G, B is the HDR image floating point data.

The header of RGBE file includes:

- ASCII information header, begins with "#?RADIANCE"
- One line includes "FORMAT=32-bit_rle_rgbe"
- Image information lines begin with "EXPOSURE=" or "CAPDATE=" and so on, if present
- One empty line
- Resolution information line telling resolution and pixel scanning order, -Y M +X N

For example:

```
#?RADIANCE
```

```
CAPDATE= 2003:07:14 17:25:14
```

```
PRIMARIES= 0.6400 0.3300 0.3
```

```
FORMAT= 32-bit_rle_rgbe
```

-Y 2592 +X 1944

(the rest data in this .hdr file)

It tells us that the capture date is 14/07/2003 and the resolution is 2592X1944.

Then, the rest data lines record the RGB channels' information using run-length encoding.

3.3 Display

The captured radiance maps record the relatively accurate information of the original scenes. Most of the current visualization devices cannot display the HDR image file directly because the conventional monitors have a dynamic range of only two orders of magnitude which is much less than that of HDR images. Although currently a few high dynamic range display systems have been introduced [34], they only can be found in laboratory because of the high cost. Moreover, they just present a bit more dynamic range than the conventional ones, and still cannot display all information contained in HDR image.

To solve this problem, a category of image processing techniques called tone mapping are introduced, which copes with the huge range of real world luminance and the limited dynamic range of the monitors. Tone mapping is an important component in the digital imaging pipeline. The objective of tone mapping is to manipulate the image data such that a visually pleasing image can be reproduced either on an electronic display (e.g. an LCD panel) or hardcopy prints. Often, the original image maybe too dark, various details may not be apparent, or the overall contrast may not match that of the reproduction media. In these cases, it is necessary

to adjust the image data, i.e. tone map the image, so that in the reproduction (display or printing) the details are visible, the colours are correct, the contrasts are appropriate and the reproduction conveys a visual appearance that reflects as truthfully as possible the world scene. In the case of reproducing high dynamic range (HDR) images in conventional low-dynamic range (LDR) media, tone mapping is used to compress the dynamic range of the image data such that it can be fitted within the range of the reproduction media.

3.4 Summary

This chapter describes the high dynamic range imaging workflow, including recovering HDR radiance maps, HDR radiance maps storage and RGBE file format, HDR image display and tone mapping

Chapter 4

Previous Tone Mapping Algorithms

There are three possible ways to display a high dynamic range image.

1. Represent on high dynamic range display system. As mentioned in the introduction, most of these displays can only be found in laboratory because of the high cost. They just present a bit more dynamic range, but are far from the requirement of displaying HDR image.
2. Display part of the HDR image data each time by different exposing. Just like taking different exposed photos for creating HDR image, each time only display part of the HDR image data, and rest data will be displayed as totally white or black.
3. Display after tone mapping. Tone mapping will compress the HDR image's dynamic range to fit the conventional display devices. Therefore, it is also-called tone reproduction, HDR compression, and HDR optimization. With tone mapping, HDR image can be simply viewed on conventional devices.

Currently, a lot of tone mapping algorithms have been introduced for displaying high dynamic range images. They can be classified into two groups: global tone mapping algorithms and local tone mapping algorithms [41]. Global tone mapping algorithms refer to techniques that manipulate the pixel distributions, representative examples

include [16], [24], [28], [32], [42], [51] and [68]. Local tone mapping algorithms involve the spatial manipulation of local neighbouring pixel values, often at multiple scales; examples include [9], [11], [14], [17], [50], [52], and [70]. In the following sections, some commonly used algorithms will be introduced.

4.1 Initial step

In a few tone mapping algorithms, there is an initial step to scale the HDR image luminance. One approach is Logarithmic Luminance Mapping (or similar methods). In this approach, these algorithms always only work on the luminance of the image, which can be calculated using $I = 0.299*R+0.587*G+0.114*B$ where R, G, B is the floating point values of HDR image luminance. Then following function is used to initially compress the luminance of the high dynamic range image to the dynamic range DR :

$$I_{DR} = DR * \frac{\log(I+\tau)-\log(I_{min}+\tau)}{\log(I_{max}+\tau)-\log(I_{min}+\tau)} \quad (4.1)$$

where I is the original HDR image luminance value, I_{min} is the minimum luminance value and I_{max} is the maximum luminance value; I_{DR} is the compressed luminance value of dynamic range DR ; adjusting τ will appropriately tune the overall brightness of the reproduced image, and τ is also used to avoid $\log 0$. This computing formula is used to initially convert the HDR image data to luminance value of dynamic range DR . After logarithmic luminance mapping, the lower dynamic range luminance data will be modified using different algorithms and scaled to $[0, 255]$. At the end, the LDR luminance ($DR = 256$) data will be used to calculate the final output pixels.

The value of DR is normally set to 10000, or even larger. Because the obtained luminance data is linear, the I_{DR} can be directly used to produce the final output result if setting DR to 256.

The formula to calculate the final output pixels is as following:

$$R_L = \left(\frac{R_H}{L_H}\right)^\gamma * I_{DR} \quad (4.2)$$

$$G_L = \left(\frac{G_H}{L_H}\right)^\gamma * I_{DR} \quad (4.3)$$

$$B_L = \left(\frac{B_H}{L_H}\right)^\gamma * I_{DR} \quad (4.4)$$

where L_H is the original HDR luminance value; the parameter γ which can be set between 0.4 and 0.6, controls the colour appearance of the output image; R_H, G_H, B_H is the original HDR colour channel value; R_L, G_L, B_L is the final output image data.

4.2 Global Tone Mapping Algorithms

4.2.1 Hierarchical Tone Mapping

This algorithm is termed hierarchical nonlinear linear (HNL) tone-mapping operator in the original paper [23]. It maps the pixels in two hierarchical steps. The first step allocates appropriate numbers of LDR display levels to different HDR intensity intervals according to the pixel densities of the intervals. The number of HDR intensity intervals is less than the number of available LDR display levels (usually 8, 16, or 32). The second step linearly maps the HDR intensity intervals to their corresponding LDR display levels. In this algorithm, the assignment of LDR display levels to HDR intensity intervals is controlled by a quite simple and flexible formula with a single adjustable parameter. The entire process is described as follow:

Firstly, logarithmic luminance mapping is used to compress the HDR luminance to the linear lower dynamic range luminance value (set the dynamic range to more than 10,000) and construct the luminance histogram. Then divide the mapped luminance values into a certain number (suppose this number is K) of equal length intervals as Figure 4.1

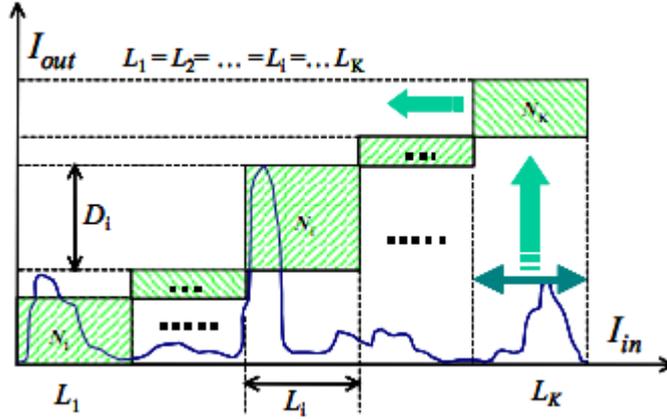


Figure 4.1: To estimate the density across the intensity range, we can divide the intensity range into equal length and then count the number of pixels falling into each interval. Pixels falling into interval L_i are assigned D_i Display levels, which is proportional to N_i . In the second step, L_i will be linearly mapped to D_i .

K is the number of equal length intervals and N_i is the number of pixels falling onto interval L_i , then the number of display levels Num_{L_i} assigned to the i^{th} interval is:

$$Num_{L_i} = \frac{N_i^\beta}{\sum N_i^\beta} * L \quad (4.5)$$

where L is the total number of different intensity levels that LDR monitor can accommodate and usually it is 256. A control parameter β is used and $\beta \geq 0$. When $\beta = 0$, it is linear mapping because each equal length interval is given equal number of display levels (i.e. $\frac{L}{K}$). With the increase of β , intervals with larger N_i are assigned more display levels whereas intervals with smaller N_i are given fewer display levels. When $\beta = 1$, the number of display levels assigned to each interval L_i is exactly proportional to the ratio of N_i to the total number of pixels. At the extreme β tend to positive infinite, all the display levels will be assigned to the interval with the largest N_i since the larger β is the more display levels are allocated to the intervals with larger N_i .

After allocating the LDR display levels to each logarithmic luminance mapped HDR luminance interval, the next step is to linearly map the HDR luminance intervals to

their corresponding LDR display levels. For instance, if L_i is allocated 10 LDR display levels (suppose from 89 to 98), the interval L_i in HDR luminance should be divided into 10 segments with equal length and then pixels within each segments will be assigned one display level (e.g. pixels with the first segments are assigned 89, and with the third segments are assigned 91). After this step, the logarithmic luminance mapped lower dynamic range luminance data has been scaled to $[0, 255]$, and can be used to calculate the final output pixels.

Some results of this algorithm are shown in Figure 4.2:



Figure 4.2 Hierarchical tone mapping results

4.2.2 Fast Tone Mapping [42]

There are two extreme forms to quantize the output of the logarithmic luminance

mapping algorithm. When linearly quantizing the output to display intensity levels, the obtained images show well preserved overall brightness, but usually lack details and appear to have low contrast; when quantizing the output using histogram equalization, the obtained images show much more details but cause artefact in the uniform luminance area. In this algorithm, a linear to equalized quantization strategy will be used:

$$le_n = l_n + \beta (e_n - l_n) \quad (4.6)$$

where l_n is the cut for linear quantization luminance axis, and e_n is the cut for histogram equalized quantization luminance axis ($n = 1, 2, 3, \dots, 255$). $0 \leq \beta \leq 1$ is a controlling parameter. If $\beta = 0$, the quantization is linear, while $\beta = 1$, the quantization is histogram equalized, usually β is always set to 0.4 to 0.6. By setting the β , it can strike a balance between the above two extreme forms.

To implement this algorithm, a highly efficient recursive binary cut approach will be used. Firstly, find the cut that divides the luminance range into two equal length intervals (linear quantization to two levels), then find another cut such that two sides have the same amount of pixel (equal to half of the total pixel population, i.e. histogram equalized quantization to two levels). Then find a cut between these two cuts using the linear to equalized quantization strategy. After this step, apply the same process recursively to cut the luminance axis into 256 intervals and map the HDR luminance value to the 256 low dynamic range display value. The detailed steps of the recursive binary cut approach to implement this algorithm are shown in Figure 4.3.

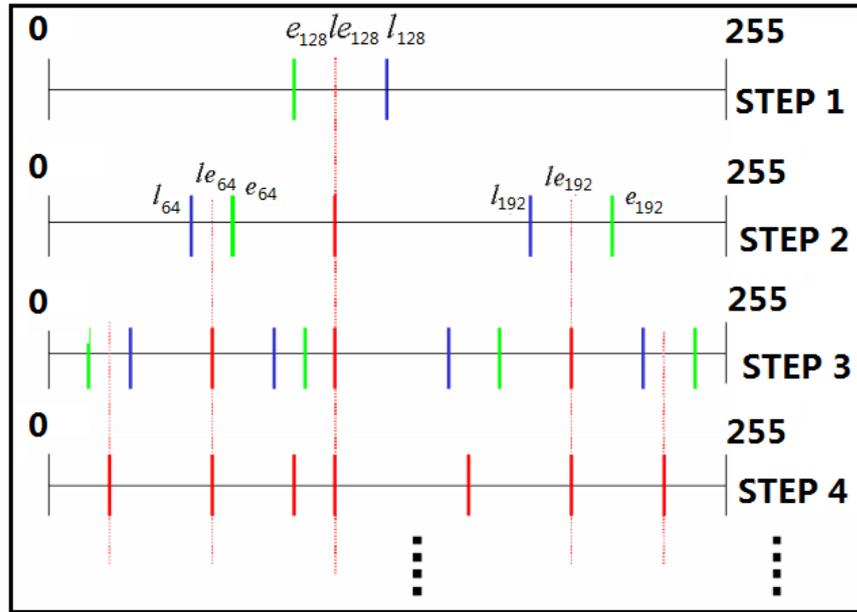


Figure 4.3 Recursive binary cut approach implementation of histogram adjustment approach based the linear to equalized quantization strategy. The algorithm first finds the le_{128} and divides the luminance into two segments. Then these two segments are each independently divided into 2 segments. The process is then applied recursively onto each obtained segment to divide it into 2 new segments based until the 256 segments (LDR levels) are created.

After the recursive binary cut process, map the HDR luminance value within the same group to the same low dynamic range display value, and then calculate the final output pixels. Some results are as shown in Figure 4.4:

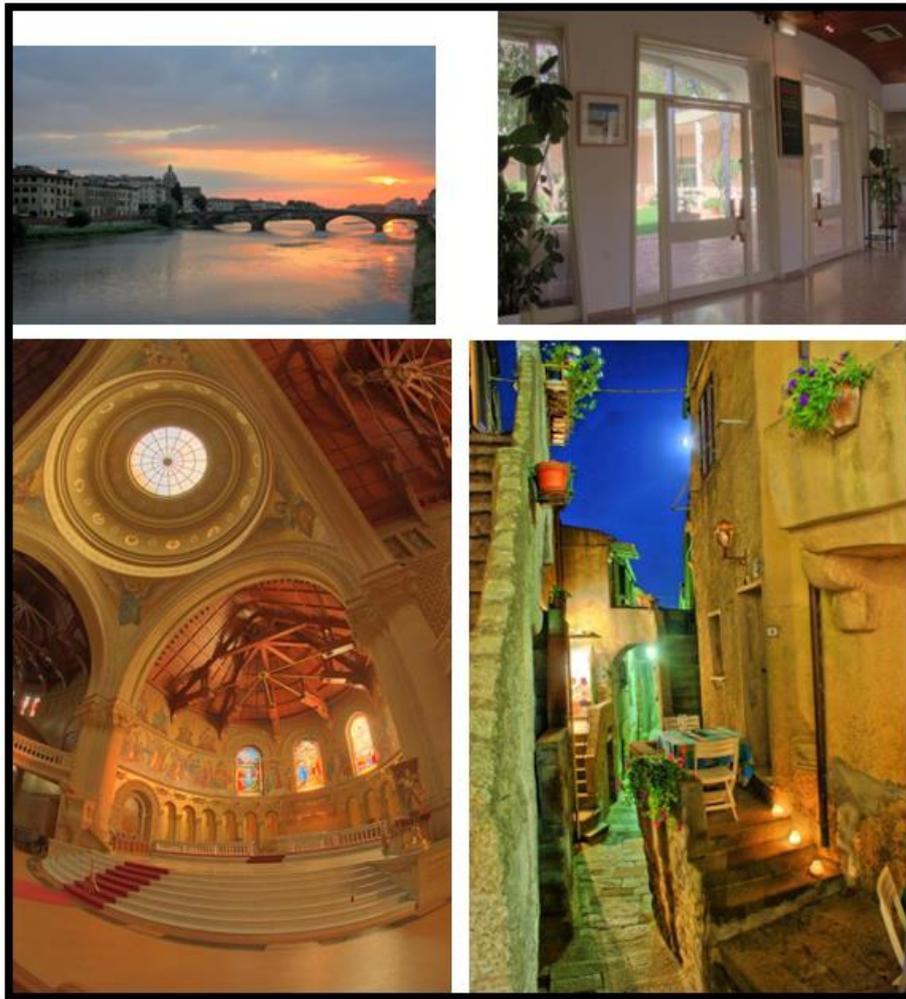


Figure 4.4: Fast tone mapping results

4.2.3 Other Global Tone Mapping Algorithms

This section will briefly introduce another four commonly used global tone mapping algorithms: Tumblin and Rushmeier's Algorithms [51], Ward's Algorithms [32], Ferschin et al's Exponential Mapping [68], and F. Drago et al's adaptive logarithmic mapping [16].

4.2.3.1 Tumblin and Rushmeier's Algorithms

Early in HDR imaging study, researchers always focus on the problem how to map real world luminance values to display luminance values. Tumblin and Rushmeier introduced a framework to solve this [51].

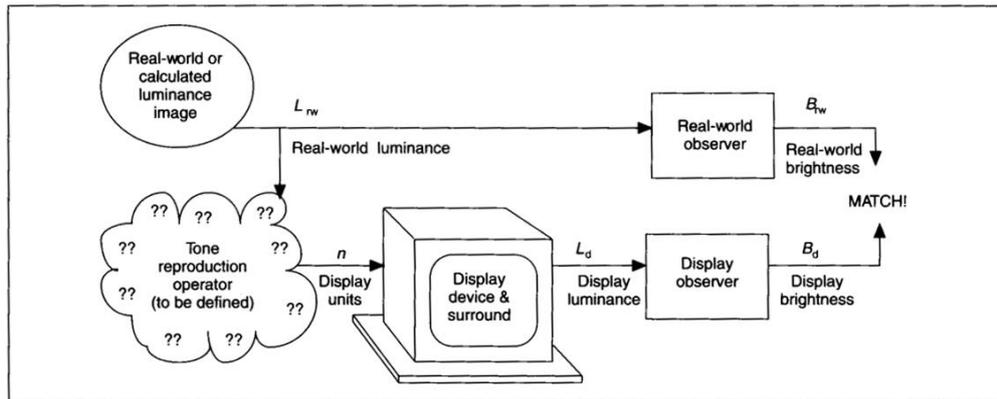


Figure 4.5: Tone mapping problem illustration from [51]

They provide the illustration in Figure 4.5 to show the real world observer and display observer model is a mathematical model of how human visual system converts real world luminance to perceived brightness. However, it is impossible to physically obtain the exact luminance match because of the huge difference of dynamic range between the real world and conventional visualization devices. Moreover, a quantity of factors like surround background will always affect the perceived brightness; different viewing conditions will sometimes cause that different luminance evoke the same perceived brightness for humans. Based on this, Tumblin and Rushmeier introduced a simple tone mapping operator by concatenating a real-world observer model, and inverse display observer model, and an inverse display device model as shown in Figure 4.6.

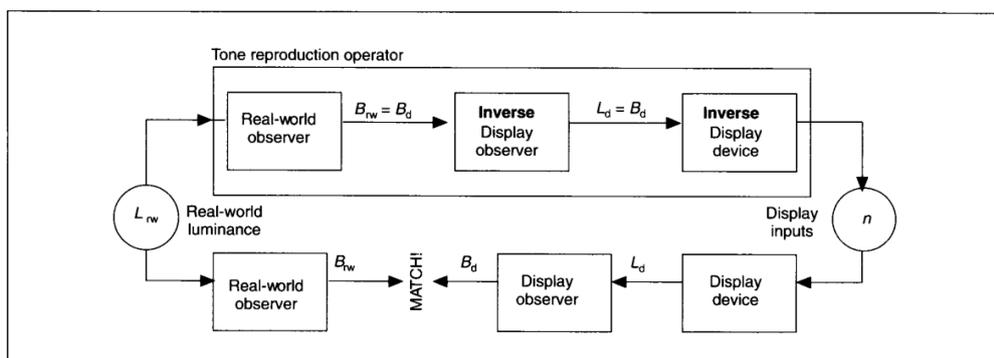


Figure 4.6: Tumblin and Rushmeier's solution from [51]

Authors chose Stevens et al's observer model [76], which assigned a luminance to

the brightness using a power function to relate luminance I_{in} to brightness B at an adaptation level of I_a is:

$$B = 10^\beta I_{in}^\alpha \quad (4.7)$$

where α and β are calculated as:

$$\alpha = 0.4 \log_{10}(I_a) + 2.92 \quad (4.8)$$

$$\beta = -0.4(\log_{10}(I_a))^2 + (-2.584 \log_{10}(I_a)) + 2.208 \quad (4.9)$$

where I_a is the adaptation level. Authors also suggest the following calculation for world adaptation level:

$$\log_{10}(I_{aw}) = E[\log_{10}(I_{in})] + 0.84 \quad (4.10)$$

where the $E[\log_{10}(I_{in})]$ is the expected value of $\log_{10}(I_{in})$

However, it is a tricky “chicken and egg” problem because to work out the adaptation level, the display luminance is always needed and vice versa. To solve this, they assumed that the distribution of display luminance was uniform on a logarithmic scale and the midpoint in this scale could be used as display adaptation level. Tumblin and Rushmeier’s operator reproduces the overall brightness well, however it still lose the spatial details especially in the bright regions.

4.2.3.2 Ward’s Algorithms

Ward [32] introduced a similar strategy to solve the problem shows in Figure 4.5. But Ward’s approach is simpler and it preserves most details in the images. In his approach based on Blackwell’s method [32], he describes that the minimum discernible difference in luminance ΔI is related with adaptation luminance I_a as following:

$$\Delta I(I_a) = 0.0594(1.219 + I_a^{0.4})^{2.5} \quad (4.11)$$

Then, correlate the discernible difference of real world luminance and display by:

$$\Delta I(I_{da}) = m \Delta I(I_{wa}) \quad (4.12)$$

Where m is the scale factor. I_{da} represents the display adaptation luminance. I_{wa} is the realworld adaptation luminance.

Finally, the display luminance can be calculated as:

$$I_d = mI_w \quad (4.13)$$

where I_d is the display luminance and I_w is the real world luminance.

Compared with Tumblin and Rushmeier's operator, Ward's operator successfully preserves spatial details, but performances worse for simulating the perceived brightness.

4.2.3.3 Ferschin et al's Exponential Mapping

Ferschin et al introduced a simple tone mapping curve - exponential mapping [68]. They used the following function to calculate the final output luminance value from the real world luminance value:

$$I_{DR} = DR * (1 - e^{-\frac{I}{I_{average}}}) \quad (4.14)$$

where the DR is the compressed dynamic range (should be 256 normally in current display devices); the I_{DR} is the compressed luminance value of dynamic range DR ; the $I_{average}$ is the average luminance value of the real world scene. Figure 4.7 shows the exponential tone mapping curve.

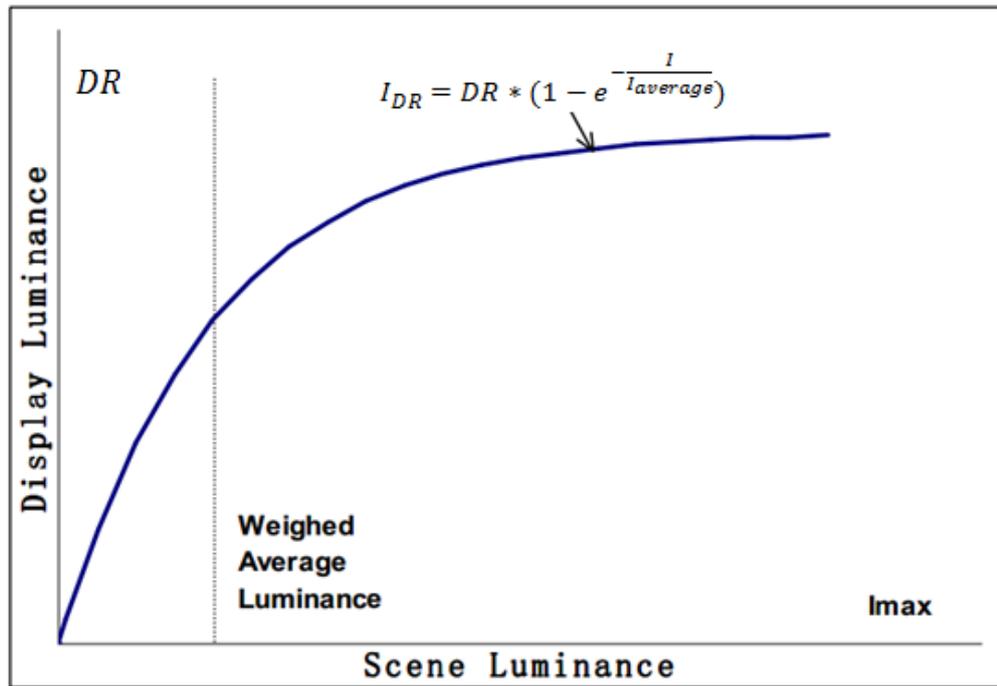


Figure 4.7: Exponential mapping

The exponential mapping was targeted to provide a smooth transition between the real world luminance value and the display luminance value. This function always gains the average real world luminance value to $0.632 * DR$, therefore this method would reduce the bright regions' brightness and increase the dark regions' brightness. However, this simple method will considerably lose some spatial details.

4.2.3.4 F. Drago et al's adaptive logarithmic mapping

Another common approach is the adaptive logarithmic mapping introduced by F. Drago et al. [16]. This author adopted a simple logarithmic mapping in equation (4.15) from Stockham [81]:

$$I_d = \frac{\log(I_w + 1)}{\log(I_{max} + 1)} \quad (4.15)$$

where I_d is the display luminance; I_w is the scene luminance; I_{max} is the maximum scene luminance. Then F. Drago et al. changed the base of the logarithmic function to achieve different representation of the mapping images. From the experiments using equation (4.15), it is obvious that no matter what the value is, a satisfactory

result cannot always be obtained if using a fixed base. After changing base value, F. Drago et al. found that with smaller base the function will increase the contrast and brightness for low luminance values; with larger base the function will extremely compress the higher luminance values; with different base, the logarithmic function will achieve that particular luminance range being best mapped. Therefore they introduced a central idea: smaller scene luminance needs smaller logarithmic base, vice versa. They proposed mapping function borrowed from “bias” power function by [53]. After the development of the operator, they also proposed an approach to accelerate the algorithm.

4.3 Local Tone Mapping Algorithms

It is easy to see that global tone mapping methods keep less local contrast information of the original scenery. When people see a scene, the eyes adapt to different local regions in the scenery and are adjusted to a state that best capture the contrast and details within it. Since the focus of our eyes navigates rapidly among different small regions in the scene and quickly adapts to them, it appears that we simultaneously see details everywhere in HDR scenery. To simulate this, local methods always segment the scene to several regions and do tone mapping in each region, and then merge them and avoid boundary artefact.

4.3.1 Local Histogram Adjustment [44]

This algorithm is based on fast tone mapping. Firstly, the image is divided into a number of regular blocks (the number is set to 32X32 in this method) and apply fast tone mapping in each block, the results are as Figure 4.8:



Figure 4.8: Applying fast tone mapping in divided blocks

Obviously, these images show more details and local contrast in either dark or bright regions. However, the local area causes sharp jumps among different blocks. The result is that boundary artefacts are presented in the images. In order to solve the boundary artefacts, an approach as illustrated in Figure 4.9 is introduced. For each pixel $D(x,y)$ in the image, use the developed fast tone mapping in each block to transfer it to $\text{HALEQ}_1[D(x,y)]$, $\text{HALEQ}_2[D(x,y)]$, $\text{HALEQ}_3[D(x,y)]$ (HALEQ is short for Histogram Adjustment using Linear and Equalized Quantization strategy, i.e. fast tone mapping mentioned above), etc.

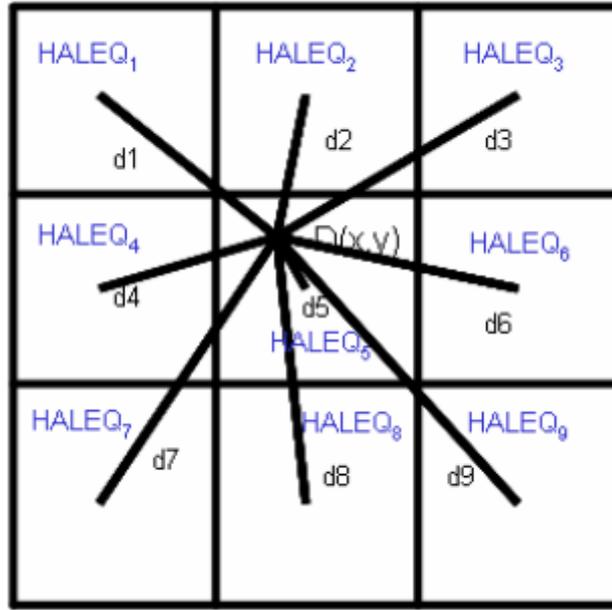


Figure 4.9 Distance weighting function approach is introduced to remove the boundary artefacts. Usually use 7*7 blocks.

Then, a distance weighting function is used for smoothing:

$$d(x, y) = \frac{\sum_{n=1}^{n=K} HALEQ_n[D(x,y)] * wd(n)}{\sum_{n=1}^{n=K} wd(n)} \quad (4.16)$$

where

$$wd(n) = e^{-\frac{d_n}{\sigma_d}} \quad (4.17)$$

$HALEQ_n$ is the $block_n$'s mapping function; $wd(n)$ is the distance weighting function.

Setting σ_d to bigger values, the image will be free from boundary artefact but shows less local sharpness. In the experiment, it shows good results if setting σ_d to 20.0. d_n is the Euclidean distance between current pixel position and the centres of each of the blocks.

Now the results are shown as Figure 4.10:



Figure 4.10: Local histogram adjustment tone mapping results

4.3.2 Other Tone Mapping Algorithms

This section will briefly introduce another four commonly used local tone mapping algorithms, Gradient Domain [70], Bilateral Filtering [17], and Tumblin et al's two algorithms [52].

4.3.2.1 Gradient Domain

Raanan Fattal et al. proposed a method to compress high dynamic range image [70]. They manipulate the field of the luminance image by attenuating the magnitudes of large gradients in the logarithmic space, because the logarithmic luminance corresponded to perceived brightness and the gradients in the logarithmic domain corresponded to the local contrast in the luminance domain. The basic idea is: first, in the logarithmic luminance domain, calculate the gradient and detect the contrast magnitude in the original luminance domain. Next, different scale factors are employed at each pixel to attenuate the larger gradient more than smaller gradient change, so that it compresses the larger contrast more than smaller contrast. At the end, convert attenuated gradient field back to the logarithmic luminance domain by

integrating the compressed derivatives, and then return to the luminance domain by exponentiation.

It sounds that this is a simple and straightforward idea. However, the attenuated gradient field of 2D image is not necessarily integrable, which causes the difficult. Authors then proposed to search a function whose gradient was the closest to attenuated gradient field from the space of all 2D potential functions. Results from the algorithm presents are considerable excellent. However this algorithm needs to adjust the parameters manually every time for good results.

4.3.2.2 Bilateral Filtering

This algorithm's basic idea is to reduce the contrast but without losing any detail [17]. To implement this, it firstly decomposes the image into a base layer and a detail layer. In the base layer, for preserving detail only its contrast will be reduced. It use bilateral filter (edge-preserving filter) to gain the base layer [6]. This is a non-linear filter with Gaussian filter in the spatial domain multiplied by an influence function in the intensity domain to calculate each pixel' weight and the weight of pixels with large intensity differences will be decreased.

This algorithm's results are quite excellent when HDR image with much details information because it will preserve them well.

4.3.2.3 Tumblin et al's two algorithms

Based on human visual system, Tumblin et al. [52] introduced two approaches to display high dynamic range images on conventional low dynamic range devices. The first approach called laying method is only applied for synthetic images which directly providing luminance layer, reflectance layer and transparency layer information. This approach compresses only the luminance layers to preserve the

details included in another two layers. The other approach called the foveal method actually is a revised brightness matching tone mapping operator from [51]. This approach adopted a foveal region to preserve the local contrast using a sigmoid function.

Chapter 5

Shutter Speed based HDR Radiance Maps Recovery and Photograph Fusion

5.1 Introduction

Real world's scenes often contain a broad range of light variations. When they are represented by images captured by conventional digital cameras, normally only 256 luminance levels per colour channel will be kept. Therefore, much of the scenes' detail information is lost during the capture process where the camera non-linearly converts and compresses the scenes irradiance intensities to finite bit-depth (normally 8 bits) pixel values. This is particularly true when bright and dark areas exist together in the same scene. To solve this problem, high dynamic range imaging (HDRI) was introduced in the late 1980s. A HDRI file (often called the radiance map) stores a scene's radiance values with floating point numbers and has much higher dynamic range than a traditional image file [15].

Although there have been a number of cameras that can directly capture wide dynamic range scenery and output 12-bit or 16-bit images, they are very expensive and still have the dynamic range limit problem. An alternative low-cost approach is to use a sequence of low dynamic range (LDR) images taken in the same scene under different exposure times. In high exposed photos, information in bright areas is lost, and information in dark regions is well kept. In contrast, low exposed photos

record more information in the bright regions than in the dark regions. Therefore, it is possible to increase the dynamic range by using multiple differently exposed LDR images.

An influential technique introduced by Debevec [67] has effectively become the *de facto* standard for recovering high dynamic range radiance map from photographs and is widely used in research and commercial systems. However, we have discovered an important defect in the original algorithm that makes this technique often fail to produce reasonable results in the extremely bright or dark regions of a scene. We have found that many software implementations by researchers and enthusiasts also carry this defect. We found one piece of work [74] that has also discovered such defect and implemented a simple fix. However the fix introduced in [74] did not completely resolve the problem. Fixing this defect will make the already excellent technique of [67] more complete.

In this chapter, we introduce a novel technique to correct this defect in the original algorithm of [67]. We also show that our new technique can be successfully used to directly fuse differently exposed photographs into a single low dynamic range image for display in conventional devices.

5.2 Recovering High Dynamic Range Radiance Map from Photographs Revisited

The widely used approach to recovering HDR radiance maps works as follows: (i) use multiple LDR images taken under different exposure time from the same static scene to calculate the imaging system's transfer function (the so-called response function) and (ii) use its inverse to estimate the radiance values and build an estimation of the irradiance values of the original scene (radiance maps). Many algorithms for recovering HDR radiance map have been developed including [5], [46], [67], [74], [75], [80] and [91]. The method developed in [67] is particularly

influential and the most widely used. In Chapter 3, we have introduced the technique by [67].

To fuse the multiple LDR images into a single radiance map, [67] just simply used each pixel's value in each frame to decide how heavy the weighting of a particular exposure should be as shown in equations (5.1).

$$\ln E_i = \frac{\sum_{j=1}^n w(Z_{ij})(g(Z_{ij}) - \ln \Delta t_j)}{\sum_{j=1}^n w(Z_{ij})} \quad (5.1)$$

where E is linearly related to scene radiance, Z is the low dynamic range image's value, i represents a spatial index, j represents images' index over exposure time Δt_j , n is the number of images, g is defined the natural logarithm of the camera response function's inverse: $g = \ln f^{-1}$

and, $w(Z)$ is defined as equations (5.2)

$$w(Z_{i,j}) = \begin{cases} Z - Z_{\min}, & \text{if } Z \leq 0.5(Z_{\min} + Z_{\max}) \\ Z_{\max} - Z, & \text{if } Z > 0.5(Z_{\min} + Z_{\max}) \end{cases} \quad (5.2)$$

Where $Z_{\min} = 0$ and $Z_{\max} = 255$

In the following particular cases, this approach will fail to produce sensible radiance values: In the extremely bright areas, in a spatial coordinate where the pixels across every frame are saturated, i.e., the pixel value is 255, all weightings for differently exposed pixels will be the same. In extremely dark areas, in a spatial coordinate where all pixels across every frame are either 0 or 1, all weightings for differently exposed pixels are either 0 or 1. This means that in these cases, we cannot determine a reasonable estimation of the radiance values.

In practice, when there is a need to use HDR imaging technology to record the real world scene, these cases (saturated or dimly illuminated) will always present. The author of [74] has found such cases and he fixed the problem through a post-processing procedure. After constructing the HDR radiance map based on [67],

saturated pixels are identified and the radiance values of the saturated pixels are set to a slightly higher value than the maximum radiance intensity of non-saturated areas. This process not only involves an extra procedure to identify and record the saturated pixels but also can only set all saturated pixels to one value that is surely not proper.

5.3 Camera Shutter Speed based Photograph Fusion

To fix the defect discussed in Section 2, we propose to explicitly exploit information contained in the shutter speeds used to take the photographs. Our method works as follows:

Step 1: calculate the “base shutter” value b as

$$b = \sqrt[n-1]{\Delta t_{max}/\Delta t_{min}} \quad (5.3)$$

where n is the frame number, and Δt is the shutter speed. For example, if the inputs' shutter speed is between $1/32$ and $1/512$, and there are 5 frames, then the base shutter value is

$$b = \sqrt[5-1]{(1/32)/(1/512)} = 2$$

Step 2: calculate “suitable shutter” value s

We consider the frame with an EV value of 0 as the reference frame. For each pixel, we estimate the shutter speed that will make it best exposed and we call this shutter speed “suitable shutter”. We estimate the suitable shutter as

$$s_i = \Delta t_{ref} * b^{\left(\frac{128-Z_{i,p}}{128} * p\right)} \quad (5.4)$$

where Δt_{ref} is the reference frame's shutter speed, p is the reference frame's index, $Z_{i,p}$ is the pixel value in spatial index i in the reference frame. Based on equation (5.4), we can obtain a suitable shutter speed for each pixel. When a pixel is close to

mid-grey (128), its suitable shutter speed will be close to that of the reference frame. When a pixel is close to 0, the suitable shutter will be close to Δt_{\max} . When a pixel value is close to saturate (255) then the suitable shutter will be close to Δt_{\min} .

A. HDR Radiance Maps Recovery

To use (5.1) to recover the HDR radiance map, we employ the suitable shutter value to obtain the weighting for a pixel. The pixel from the suitable shutter frame should make the largest contribution. Instead of (5.2) the weighting of a pixel is calculated as

$$w_s(Z_{i,j}) = \begin{cases} \frac{\Delta t_j}{s_i}, & \Delta t_j < s_i \\ \frac{s_i}{\Delta t_j}, & \Delta t_j \geq s_i \end{cases} \quad (5.5)$$

To use all frames to recovery the HDR radiance map

$$\ln(E_i) = \frac{\sum_{j=1}^n w_s(Z_{i,j})(g(Z_{i,j}) - \ln \Delta t_j)}{\sum_{j=1}^n w_s(Z_{i,j})} \quad (5.6)$$

where j is the image's index, Δt_j is each frame's shutter.

B. Exposure Fusion

This idea can also be applied to exposure fusion [79]. If the pixel is in a frame with shutter speed close to the suitable shutter, it should get a larger weighting for blending. In other words, well-exposed pixel should always get a larger weighting. Blending differently exposed photographs can be done as

$$F_i = \frac{\sum_{j=1}^n w_s(Z_{i,j})Z_{i,j}}{\sum_{j=1}^n w_s(Z_{i,j})} \quad (5.7)$$

where F_i is the fused pixel value and w_s 's are calculated as equation (5.5). Compare to previous methods such as [79], (5.7) is much simpler and we shall show in the experimental results that our new simple scheme outperforms [79]. In fact, (5.7) can be used with [79] together.

5.4 Experimental Results

In the first experiment, we compare the HDR radiance maps generated using our new technique using equations (5.3), (5.4), (5.5) and (5.6) and compare the results with the HDR file generated by the software HDR Shop developed based on [67].

Figures 5.1 and Figure 5.2 show two examples of recovering high dynamic range radiance maps from photographs. From the 3D plot of the radiance values in the very bright areas (the lamps), it can be seen that whilst [67] show unnatural transition and abrupt changes, our results are much smoother. One explanation for the different results is that whilst in [67] there are only 128 possible discrete weightings, our fix has continuous weightings which also give more weights to well exposed pixels (equation 5.5). Note that here we only take the logarithm of the HDR radiance values and directly map it to 8-bit LDR image for display the HDR radiance map.

We then perform photograph fusion experiment. Instead of going through the process of recovering HDR radiance map, this approach directly fuse differently exposed photographs. We compare our results with that of [79].

Figures 5.3, 5.4 and 5.5 show examples of fusing differently exposed photographs using our new method (equation 5.7) and [79] (We obtained the results by running the MATLAB code provided by the authors.). These results show that our technique is competitive. It is seen that our technique has always managed to give higher weightings to well exposed pixels. The method of [79] also weights each pixel by its value only just like [67]. The results of [79] show more details in texture regions because a Laplacian filtering has been used to boost high frequency features. However, in highly saturated areas, [79] has failed to capture the details whilst our technique has managed to do so success fully. It is easy to see that we can easily combine equation (5.7) and the procedures of [79] together.

5.5 Summary

The technique of [67] has become a de facto standard for recovering high dynamic range radiance maps. However, there is a defect in the original algorithm where it often failed to produce sensible results in the extremely dark and bright areas of the scene. We found that many software implementations by researchers and enthusiasts also carry this defect or have implemented simple fixes, eg. [74].

This chapter has presented a novel high dynamic range radiance maps recovery method based on [67]. This method uses each frame's shutter speed and pixel values together to calculate the weighting, which makes sure well-exposed pixels get larger weighting, and produce satisfactory results, particularly in bright and dark areas. We also demonstrated the new method can be successfully used for fusing photographs and is competitive to state of the art techniques.

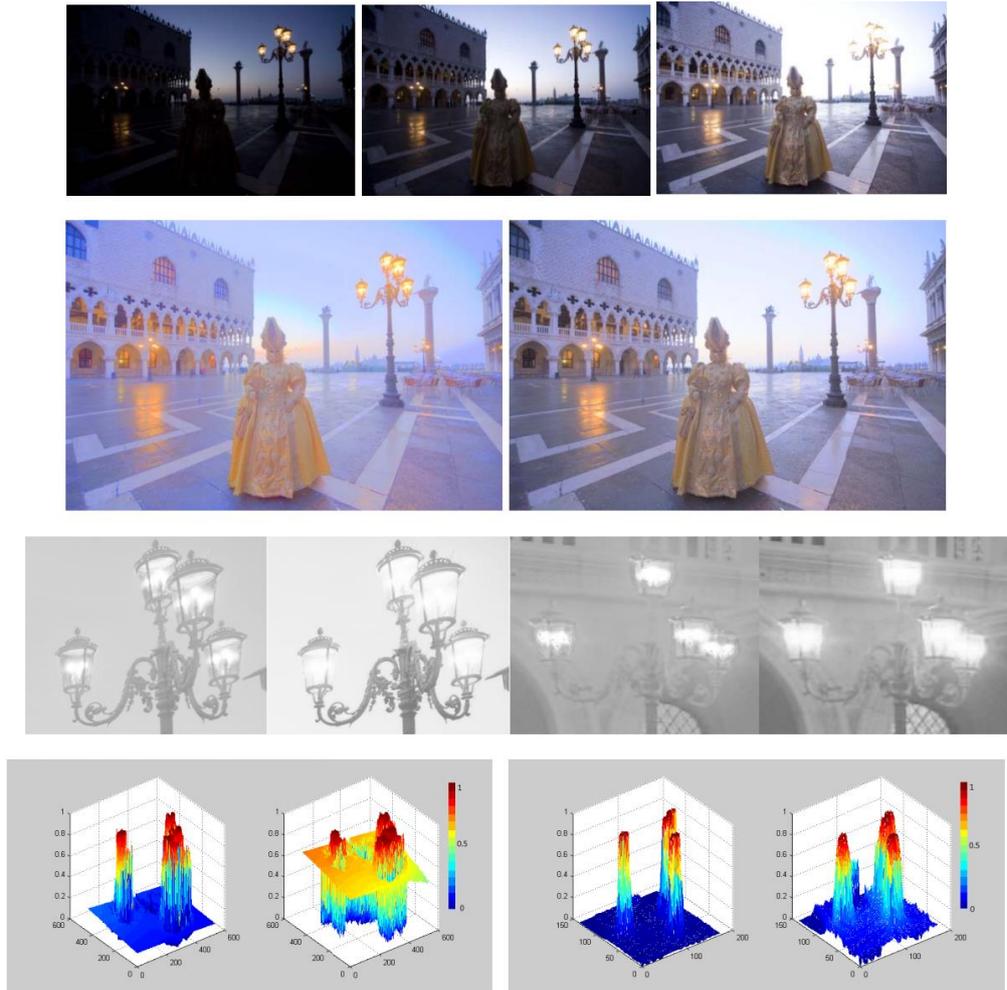


Figure 5.1: 1st row: Three differently exposed photographs. 2nd row: Logarithmically mapped LDR images of HDR radiance maps, result of [67] (left) and our new result (right). Note that these images are for informal visualization purposes. 3rd row: two sub-images in grey scale from the logarithmically mapped HDR radiance maps, [67] (left) and ours (right). 4th row: 3D plots of the HDR radiance values of the sub-images in the 3rd row, [67] (left) and our result (right), the Z-axis is the radiance values. It is seen that [67] contains more abrupt and unnatural transitions.

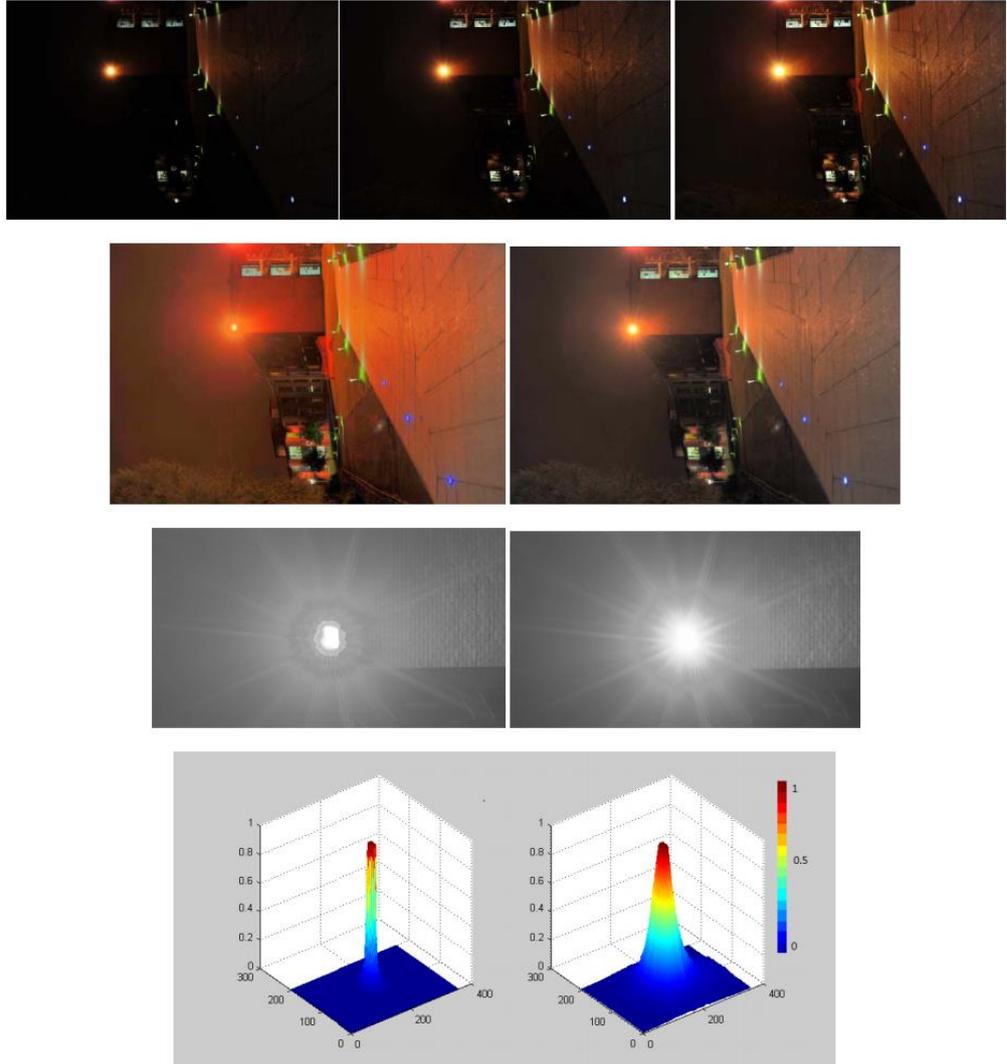


Figure 5.2: 1st row: Three differently exposed photographs. 2nd row: Logarithmically mapped LDR images of HDR radiance maps, result of [67] (left) and our new result (right). Note that these images are for informal visualization purposes. 3rd row: a sub-image in grayscale from the logarithmically mapped HDR radiance maps, [67] (left) and ours (right). 4th row: 3D plots of the HDR radiance values of the sub-images in the 3rd row, [67] (left) and our result (right), the Z-axis is the radiance values. It is seen that [67] contains more abrupt and unnatural transitions.



Figure 5.3: Top row: At the extreme right are 3 differently exposed photographs. In the middle are fusion result of our new technique and on the left is the fusion result of [79]. Bottom row: Enlarged regions highlighted on the top row, on the left is the result of [79] and on the right is our result. It is seen that in these brightly lit areas, our technique has managed to keep much more details. The Laplacian filter of [MLV09] has made their result appear sharper and contains more details [79]. Furthermore, our technique can be easily combined with procedures such as [79] to maintain the advantages of both techniques



Figure 5.4: Top row: At the extreme right are 3 differently exposed photographs. In the middle are fusion result of our new technique and on the left is the fusion result of [79]. Bottom row: Enlarged regions highlighted on the top row, for each pair, the one on the left is the result of [79] and the one on the right is our result. It is seen again that in these brightly lit areas, our technique has managed to keep much more details. The Laplacian filter of [79] has made their result appear sharper and contains more details [79]. Furthermore, our technique can be easily combined with procedures such as [79] to maintain the advantages of both techniques.



Figure 5.5: More photograph fusion results. At the extreme right are 3 differently exposed photographs. In the middle are fusion result of our new technique and on the left is the fusion result of [79]. These results demonstrate that our new technique is competitive to more complicated procedures such as [79]. Furthermore, our technique can be easily combined with procedures such as [79] to maintain the advantages of both techniques.

Chapter 6

A Comprehensive Optimization-Based Tone Mapping Algorithm

6.1 Introduction

Tone mapping is an important component in the digital imaging pipeline. The objective of tone mapping is to manipulate the image data such that a visually pleasing image can be reproduced either on an electronic display (e.g. an LCD panel) or hardcopy prints. Often, the original image maybe too dark, various details may not be apparent, or the overall contrast may not match that of the reproduction media. In these cases, it is necessary to adjust the image data, i.e. tone map the image, so that in the reproduction (display or printing) the details are visible, the colors are correct, the contrasts are appropriate and the reproduction conveys a visual appearance that reflects as truthfully as possible the world scene. Sometimes, tone mapping takes the form of image enhancement, the purpose is to make details visible or to enhance the contrasts or improve image sharpness or visibility. In the case of color image reproduction, gamut mapping is often implemented to make the color ranges in the image data fall within that of the reproduction devices. In the case of reproducing high dynamic range (HDR) images in conventional low-dynamic range (LDR) media, tone mapping is used to compress the dynamic range of the image data such that it can be fitted within the range of the reproduction

media [14], [15], [22], [25], [28], [42] and [43].

Whether to enhance details, to map color gamut, or to compress the dynamic range, the original data will be altered in some way. An important purpose of reproduction is to produce an output that will truthfully depict the original world scene which is recorded in the original image data (we assume that the original data is a faithful recording of the world scene, especially in the case of high dynamic range imaging). Manipulating the recorded data in various ways will inevitably distort the original recording making the data a less truthful representation of the world scene. If unchecked, the distortion can be so large that the display will be a poor representation of the original world scene.

In this chapter, we introduce an optimization based general tone mapping framework. The new framework is based on the premise that tone mapping should enhance the image in some way to make the details in the image visible and at the same time it should produce a correct rendition of the real world scene. Based on this principle, tone mapping can be formulated as an optimization problem where an objective function that measures the difference between the tone-mapped image and the original image and that between the tone-mapped image and a visually enhanced image are optimized. The relative weightings of the two terms in the cost function not only offers an insightful and simple mechanism to control the appearance of the tone-mapped image but also enables the introduction of spatially varying or uniform weighting functions thus unifying local and global tone mapping in a single framework.

6.2 Related works

High dynamic range imaging is a promising digital imaging technology that has attracted very extensive interest in recent years [15]. This imaging paradigm breaks away from the traditional limited bit-depth barrier by introducing an image format

that records the true dynamic range of the real world scene using as many bits per pixel as necessary. Such an image sometimes referred to as a high dynamic radiance map can be obtained by taking multiple photographs of a scene with different shutter speeds or special hardware [15].

One of the problems of high dynamic range imaging is the display of high dynamic range radiance maps on conventional reproduction media such as LCD panels. The difficulty stems from that fact that these reproduction media have limited dynamic range and it is not possible to directly feed the high bit depth high dynamic range radiance map to these devices for display. One solution to this problem is to compress the dynamic range of the radiance maps such that the mapped image can be fitted into the dynamic range of the display devices. This mapping is called tone mapping. Several tone mapping methods have appeared in the literature in recent years, including those in [14], [43], [22], [25], [28] and [43].

These methods can be divided into two broad categories. The global tone mapping techniques use a single appropriately designed spatially invariant mapping function for all pixels in the image; the local mapping techniques adapt the mapping functions to local pixel statistics and local pixel contexts. Global tone mapping is simpler to implement but tend to lose details. Local tone mapping is much more computationally intensive and harder to get it right since there are often a number of parameters in the algorithms which have to be set empirically. Given the nature of the problem, it is not possible to have one method fits all, i.e., it may not be possible to have one method that will solve the problem once and for all. What are needed are multiple methods and depending on the particular requirements of the users, one method will be better suited than others, or a combination of methods will be necessary.

6.2.1 Global Tone Mapping Curve Using Optimization

In [42], [43], [22], [25], they have described several optimization based global tone

mapping curve techniques. These techniques are based on the premise that a tone-mapped image should convey the maximum information and at the same time should be faithful to the original observation. To adhere to this principle, [43] treated tone mapping as a pixel quantization problem where, an L_2 norm was used to design the codebook which would minimize the reconstruction errors, thus ensuring the mapped image is faithful to the original data; a frequency sensitive competitive learning mechanism was used to ensure that every code word would be used by similar number of pixels, thus ensuring the mapped image has maximum contrast. In [42], [22], [25], tone mapping was based on manipulating the histograms of the input image. Heuristic and fast methods were first developed [42] [22] to manipulate the histogram where a single parameter controlled the mapping between linear mapping and histogram equalization mapping. A simpler and more elegant closed form solution was presented in [25] where a two term cost function, one favored linearly mapping and the other favored histogram equalization mapping, was solved in closed form. A single weighting parameter determined the relative importance of the two terms and controls the mapping from linearly mapping to histogram equalization mapping.

In this chapter, we introduce a general tone mapping technique based on optimization. Instead of restricting to global tone mapping curve operation, we show that the new method can be developed to include spatially adaptive tone mapping.

6.3. General tone mapping using optimization

Suppose that we have an input high dynamic range image $I(x, y)$, where the pixel values are normalized to $[0, 1]$. Global tone mapping (or dynamic range compression) of high dynamic range image can be regarded as the task of dividing the input value range into N range intervals, $[c_{i-1}, c_i]$, $i = 1, 2, \dots, N$, where N is the total number of displayable levels in the reproduction device, often we set $N = 256$

([42], [43], [22], [25]).

The mapping is done as

$$d(x, y) = i \text{ if } c_i \leq I(x, y) < c_{i+1} \quad (6.1)$$

where $d(x, y)$ is the display level of pixel $I(x, y)$. The crucial task in this scheme is finding $[c_{i-1}, c_i]$. A principle established in [22], [25], [42], and [43] which has been shown to work very well is that the mapping should be between linear scaling and histogram equalization. Let $hist(v)$ be the normalized histogram of the high dynamic range image, the mapping intervals $[c_{i-1}, c_i]$ can be obtained by minimizing the following objective function

$$E = \sum_{i=1}^N \left(c_i - \frac{i}{N} \right)^2 + \lambda \sum_{i=1}^N \left(\int_0^{c_i} hist(v) dv - \frac{i}{N} \right)^2 \quad (6.2)$$

It is clear that the first term of (6.2) favors the mapping to be linear scaling and the second term favors the mapping to be histogram equalization mapping. The weighting value λ determines the relative weightings of the two terms. A larger λ favors the mapped image to have higher contrast and vice versa. Solving the optimization problem can be achieved through heuristic [42] [22], closed form solution [25] and through learning [43].

The drawback of the above method is that it is restricted to global tone mapping curve operators which can sometimes lead to the lost of details and local contrasts due to the fact that only one mapping curve is used throughout the entire image. However, the optimization based tone mapping principle can be extended include more general cases.

Let Θ be an image enhancement operator, which can be one of many methods in the existing literature, for example, histogram equalization, gamma correction, frequency domain methods etc. [31], or retinex [92]. We seek a tone-mapped image $d(x, y)$ that should depict details about this original input $I(x, y)$ to make all features visible and enhance visual contrasts to ensure the output is visually pleasing

(evidence has shown that human subjects will prefer well contrast images), at the same time it should convey the correct overall visual appearance about the original input. Making the details visible and enhancing contrasts can be achieved through various image enhancement operations, however, one of the important reasons that many image enhancement methods failed is because the objectives were solely enhancement. For example, histogram equalization often fails because the objective is to achieve maximum contrast (equalized histogram), what is missing in the enhancement criteria is that the output image should stay faithful to the original input as well. The reason that methods such as (6.2) work well is because a constraint term has ensured that the enhancement is not overdone so as to cause the output to completely loose the visual faithfulness to the original scene.

In fact, in the image restoration literature, see e.g., [10], image restoration can often be done by minimizing a two term cost function where one term measures the smoothness and other measure the faithfulness of the output. Our new tone mapping cost function can be defined as follows

$$E = \arg \min_{\forall d} (\|d - I\|^2 + \lambda \|d - \Theta(I)\|^2) \quad (6.3)$$

Where λ is a positive weight value which determines the relative importance of the two terms. It is not hard to see that minimizing E with respect to d means that we want d to resemble the enhanced version of the image achieved through an image enhancement operator Θ , and at the same time it should be close (similar) to the original image. In (6.3) and in this chapter, we assume that the enhancement is done by an independent operator (however, it is possible to put d inside the operator Θ and formulate the problem similar to regularization functional [10] and solve for the tone-mapped d as part of Θ , we will present such solution in the future). Solving for d from (6.3) is easy, and we have

$$d(x, y) = \frac{I(x, y) + \lambda \Theta(I(x, y))}{1 + \lambda} \quad (6.4)$$

Depending on the enhancement operator Θ , the mapping can be local or global. In (6.4), the weighting factor λ is constant across the image. Sometimes, it make sense to vary the weighting factor according to the local statistics, we can therefore introduce a location adaptive version as

$$d(x, y) = \frac{I(x,y) + \lambda(x,y)\Theta(I(x,y))}{1 + \lambda(x,y)} \quad (6.5)$$

where $\lambda(x,y)$ is the weighting factor at location (x, y) . Using (6.5), even if the operator Θ is global, e.g., histogram mapping, the tone mapping can be local since $\lambda(x,y)$ varies locally. $\lambda(x,y)$ will have to be found according to Θ and we will present one method to calculate this in the next section.

6.4. Experimental results

We have performed experiments to verify the effectiveness of the proposed method. For the results presented in this chapter, the enhancement operator Θ is either global or local histogram equalization or histogram adjustment [42], [22], [25]. In the case of local histogram processing, we divided the images into $(W/32) \times (H/32)$ blocks and applied histogram processing to each block and then smooth the block boundaries. For the location adaptive weight factors, we first compute the local variance to form an intensity image $\sigma(x,y)$, where $\sigma(x,y)$ is the variance in a window centered around (x, y) . We then smooth $\sigma(x, y)$ and set $\lambda(x, y) = \mathbf{Smooth}(\sigma(x, y))$ and scale it to within the range $[0, \lambda_{max}]$, where λ_{max} is normally set interactively to achieve the right look. The effect of such location varying weighting function is that for regions of high details, the enhancement operator should play a bigger role whilst in smooth flat areas the tone-mapped image should be more like linearly scaled values.

Fig. 6.1 shows the well-known memorial church high dynamic range radiance map rendered through linearly scaling, global histogram equalization mapping and local

histogram equalization mapping. Note that the linearly mapped image lacks details, the global histogram equalization mapped image has unnatural contrast and the local histogram equalization mapped image has noticeable visual artifacts.



Figure 6.1: From left to right: linearly mapped image; histogram equalization mapped image; local histogram equalization mapped image; and spatially varied weighting function. Note the various artifacts in the first three images.

Tone-mapped images of various combinations are shown in Fig. 6.2. It is seen that the various artifacts in Fig. 6.1 have disappeared in these images. Informal comparison with other state of the art techniques shows that our results are very competitive. Here we show Reinhard’s photographic tone mapping result for comparison convenience. Our method can be easily applied to tone map low dynamic range image as well, and an example is shown in Fig. 6.3

Note that the complexity of our current implementation is only slightly higher than histogram equalization. The value of λ_{max} will have to be set empirically. In fact, we believe that it will not be possible to set this parameter automatically. Many existing algorithms in the literature have great difficulty in setting the correct parameters. The best solution is to build parameters like this into the software to enable users to adjust them interactively. It should be pointed out that the advantage of our current technique is that this parameter has a clear physical interpretation and we know a

larger value will results in more detail and contrasts. The simplistic of our algorithms also make it possible to change the values interactively to show the effects.

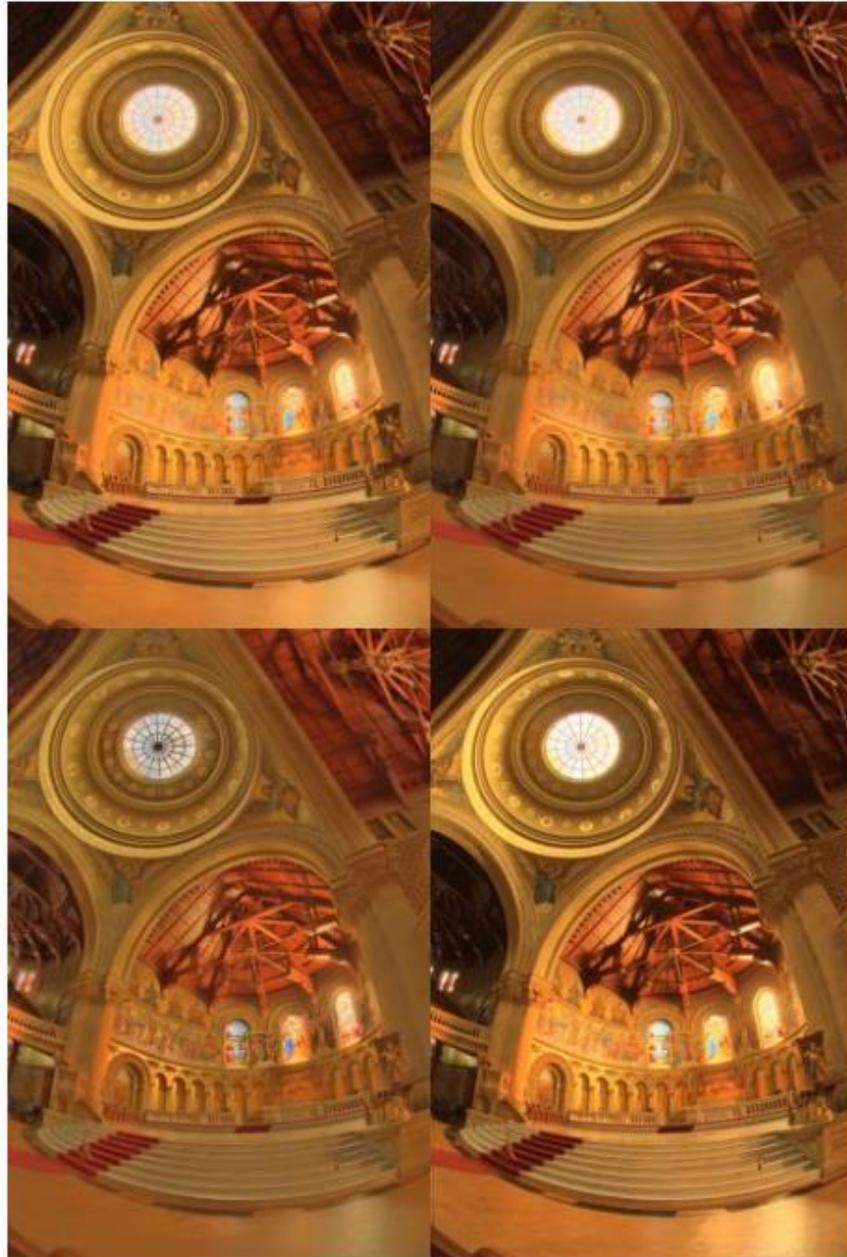


Figure 6.2: Top left: Output of the new method, Θ = global histogram equalization, and $\lambda(x,y) = \text{constant} = 1$. Top right: Θ = global histogram equalization, and $\lambda(x,y) = \text{Smooth}(\sigma(x,y))$ (see bottom right of Fig. 6.1). Note the improvement of using locally adaptive weighting values. Bottom left: Θ = local histogram equalization,

and $\lambda(x,y) = \mathbf{Smooth}(\sigma(x,y))$. Bottom right: Output of Erik Reinhard's photographic tone mapping [14]. It is seen our results is at least as good as or than Reinhard's.



Figure 6.3: Top left: A low quality photograph. Top right: Histogram equalization enhancement. Bottom left: A mapped version using equation (6.5) with $\Theta =$ global histogram equalization. Bottom right: Retinex result [92]. It is seen our result is an improvement over histogram equalization and is comparable to retinex. Note that our method is significantly simpler.

Finally, we remark that it is actually difficult to evaluate the relative performances of one TM operator against the other. We have also built a website (not accessible now, but source website code is available on request, see Chapter 8) where we intend to use Web 2.0 technology to collect user inputs from anywhere at any time under any viewing conditions. Based on evaluations by thousands maybe even millions of Internet users, we will then use mathematical modeling to rank the TM operators and to compare their performances.

More results are shown in Figure 6.4-11:

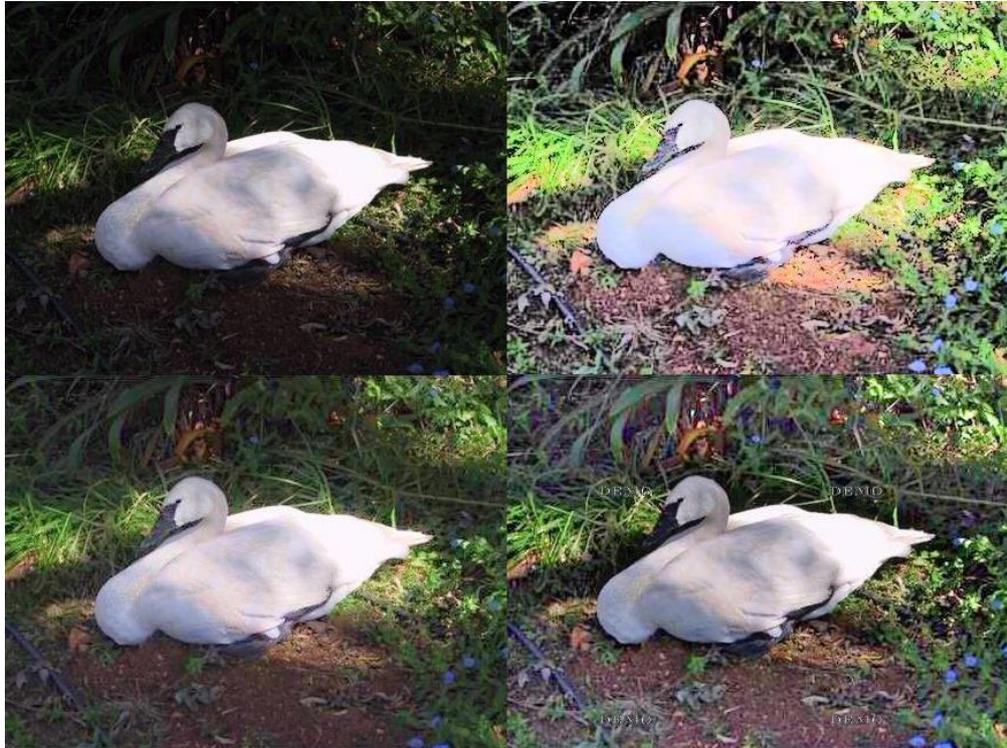


Figure 6.4: Top left: A low quality photograph. Top right: Histogram equalization enhancement. Bottom left: A mapped version using equation (6.5) with Θ = global histogram equalization. Bottom right: Retinex result [92].



Figure 6.5: Top left: A low quality photograph. Top right: Histogram equalization enhancement. Bottom left: A mapped version using equation (6.5) with Θ = global histogram equalization. Bottom right: Retinex result [92].



Figure 6.6: Top left: A low quality photograph. Top right: Histogram equalization enhancement. Bottom left: A mapped version using equation (6.5) with Θ = global histogram equalization. Bottom right: Retinex result [92].



Figure 6.7: Top left: A low quality photograph. Top right: Histogram equalization

enhancement. Bottom left: A mapped version using equation (6.5) with Θ = global histogram equalization. Bottom right: Retinex result [92].



Figure 6.8: Top left: A low quality photograph. Top right: Histogram equalization enhancement. Bottom left: A mapped version using equation (6.5) with Θ = global histogram equalization. Bottom right: Retinex result [92].



Figure 6.9: Left: Histogram equalization result. Middle: Output of the new method, Θ = global histogram equalization, and $\lambda(x,y) = \text{constant} = 1$. Right: Θ = global histogram equalization, and $\lambda(x,y) = \text{Smooth}(\sigma(x,y))$



Figure 6.10: Left: Histogram equalization result. Middle: Output of the new method, $\Theta =$ global histogram equalization, and $\lambda(x,y) = \text{constant} = 1$. Right: $\Theta =$ global histogram equalization, and $\lambda(x,y) = \text{Smooth}(\sigma(x,y))$



Figure 6.11: Left: Histogram equalization result. Middle: Output of the new method, $\Theta =$ global histogram equalization, and $\lambda(x,y) = \text{constant} = 1$. Right: $\Theta =$ global histogram equalization, and $\lambda(x,y) = \text{Smooth}(\sigma(x,y))$

6.5 Summary

In this chapter, we have presented a general tone mapping framework where we cast tone mapping as an optimization problem in which a two term cost function, one

favors the output to have enhanced details and contrasts and the other favors the output to stay faithful to the original input, is optimized. We show that this solution is only slightly more complicated than the enhancement operator. We have presented experimental results which shown that our new method works well and is competitive to state of the art.

Chapter 7

Saliency Modulated High Dynamic Range Image Tone Mapping

7.1 Introduction

Tone mapping results are what the viewers will see. It is obviously one of the most crucial steps in the imaging pipeline. Not surprisingly, such an important step has attracted significant research interest. Despite many fine efforts including our own presented in the previous chapters, tone mapping is still an open problem. The major difficulty can be attributed to the fact that tone mapping results are ultimately judged by the human observers and this is invariably subjectively [47], and as such it is very difficult to have an automatic algorithm that will work well on any given image content under any viewing environments.

It is well known that the amount of incoming information to the visual system is much greater than that can be fully processed, only part of this information is processed in full detail while the remainder is left relatively unprocessed [12]. The mechanisms of selective visual attention play an important role in biological vision. Certain regions of the visual field are more important than others and more perceptual resources are directed to the more important or salient regions. Two major attention mechanisms, bottom-up and top-down are known to control this selection process. There is now clear evidence indicating that attention can be

captured by highly salient features [26]. A number of saliency algorithms have also been developed to identify salient image regions that are highly likely to attract visual attention, e.g., [61], [62], [54]. Although visual saliency mechanism has been exploited in rendering computer graphics [36] to save computational effort, in volume visualization [86] to highlight saliency regions and in image retargeting [82] to protect important objects and regions, there has been surprising little research into the use of saliency in tone mapping HDR photographic images.

In this chapter, we present a saliency modulated tone mapping (SMTM) framework in which the bottom-up saliency map of the HDR image is directly exploited to modulate the tone mapping operator so that perceptually important features in the visually salient (important) regions are better protected in the tone-mapped low dynamic range (LDR) image. We present an implementation of the SMTM framework in which we adaptively adjust the tone mapping curves of local regions according to its saliency. We present experimental results to show the effectiveness of this new technique in rendering visually pleasing images and in protecting visual conspicuous features in the LDR images.

7.2. Related work

Most models of bottom-up attention are based on the centre surround mechanism that is ubiquitous in the early Stage of the biological vision [8]. One of the earlier models that are perhaps the most often cited is [54] where the input image is decomposed into different channels: intensity, colour, and orientation. A centre surround operation is implemented to produce feature maps of different scales which are then combined to create conspicuity maps for each channel. The conspicuous regions of these maps are then normalized and linearly combined to form the overall saliency map. This model has been shown to be successful in predicting human fixations and to be useful in object detection. Bruce and Torsos [61] calculate the

saliency of a location as the self-information of the location's visual features in the context of its surrounding pixels, Ago and Vasconcelos formulate saliency in a decision theoretic setting [7], Zhang et al [57] formulate a Bayesian framework to compute saliency using natural statistics. The authors of [62] introduced a fast implementation to [57]. Although other saliency models exist, a comprehensive survey is however beyond the scope of this part, references within above papers contain more information on various other models of visual saliency.

Visual saliency has attracted extensive interest in the computer vision and cognitive vision communities in recent years, see e.g., [7] [26]. Saliency has also been employed to compute a spatiotemporal error tolerance map to accelerate rendering of computer graphics [36]. For volume visualization, [86] uses saliency to highlight important areas. In image retargeting, saliency has been used to protect important foreground objects in the target images [82].

7.3 Saliency modulated tone mapping

Salient regions of an image will attract more attention, it is therefore important to protect such regions when compressing a HDR image to a LDR image for display. As the HDR is not directly visible, whatever that is displayed in an LDR form would have been processed and there will be a loss of information. One of the objectives in tone mapping is to ensure that the LDR image will convey as faithfully as possible the visual features of the HDR data. As information loss is inevitable, there is the question of what should be protected. It is known that local details are important and should be preserved. Methods such as bilateral filtering based tone mapping [17] protects the details by first removing them from the HDR data, the "detail free" base layer HDR data is then compressed, and the protected details are then added back to the compressed base layer unaltered. It is also known that edges are important visual features. In gradient domain based methods [70], the edges orientations are protected while attenuating their magnitudes. In quantization based methods, such as

[24], the mapping optimizes a cost function that favors the LDR image that will make full use of all available display levels and at the same time will stay faithful to the original HDR data. Although the ways in which different methods protect different kinds of visual features are different, the principle remains the same: first, identifying the important visual properties, such as details, edges, fidelity to original data, etc.; second, developing appropriate computational methods to compress the HDR to an LDR while ensuring these important visual properties are well preserved.

Although it is known that visual attention mechanism is an important cognitive vision attribute and that salient image regions will attract visual attention, there is surprisingly little research to exploit this important visual phenomenon in HDR tone mapping. In this section, we present a method that directly employs visual saliency map of the HDR image to modulate a tone mapping operator such that the salient image regions will stay salient in the LDR image and will attract visual attention.

A. High Dynamic Range Image Saliency Map

Although the HDR data is not directly visible, its visual attributes such as saliency map can be computed from the HDR radiance map. In fact, because the HDR data faithfully records the full dynamic range of the scene, its saliency map may identify salient regions of the scene more accurately - corresponding better to the saliency map of the real scene. A number of saliency computational models have been proposed in the literature, e.g., [61], [7], [54], [57], many are computationally complicated and not suitable for tone mapping application which requires a reasonably fast implementation. We adapt the fast saliency map computational method of [62] to compute the saliency map of HDR image and the modified algorithm is listed as pseudo code in Algorithm1. The luminance signal of the HDR data is computed and used to compute the saliency map (we do not use colour in the HDR saliency map calculation). Difference of box filtering (DOB) [64] is used to approximate the difference of Gaussian filtering and a probability distribution with a

unit variance is used to model the output of the DOB. It has been shown that this fast algorithm can be thousands of times faster than conventional methods without sacrificing performances. Some saliency map examples are shown in Figures 7.1, 7.2 and 7.3.

Algorithm 1 Fast HDR Image Saliency Map Algorithm

Initialization

- 1) $L \leftarrow$ Luminance of HDR radiance map
- 2) $S \leftarrow 5$ {Parameter: # of scales}
- 3) $Min\sigma \leftarrow 1$ {Parameter: Smallest Box Filter Radius $\in [1, \infty)$ }
- 4) **for** $i = 1$ to S **do**
- 5) $\sigma[i + 1] \leftarrow 2\sigma[i]$
- 6) **end for**

Start Saliency Map Calculation

- 1) $SaliencyMap \leftarrow \vec{0}$
 - 2) $Im \leftarrow DownSample(L)$
 - 3) $BF[1] \leftarrow$ Filter Im with box-filter, $width = 2\sigma[1] + 1$
 - 4) **for** $i = 1$ to S **do**
 - 5) $BF[i+1] \leftarrow$ Filter Im with box-filter, $width = 2\sigma[i + 1] + 1$
 - 6) $DoB[i] \leftarrow BF[i] - BF[i + 1]$
 - 7) $SaliencyMap \leftarrow SaliencyMap + abs(DoB[i])$
 - 8) **end for**
 - 9) **Return** $UpSample(SaliencyMap)$
-

B. Saliency Modulated Tone Mapping

The saliency map has identified regions that will attract attention and therefore should be treated as more important when mapping the HDR radiance map to a LDR image for display. In other words, the same regions that attract visual attention in the HDR radiance map (or equivalently the real world scene of the image) should have high saliency in the LDR image such that these regions will attract visual attention of the observers. It is known that high contrast regions with busy details attract

attention, it is therefore reasonable to ensure high saliency regions are mapped with higher contrasts and more details. We here present a saliency modulated tone mapping method which adjust local tone mapping curves adaptively according to the saliency values of the regions.

HDR tone mapping can be performed through tone mapping curves adapted to the distributions of the pixels. The mapping curve of [22] can vary continuously from linear mapping to equal distribution (maximum contrast) mapping. Initially developed as a global mapping operator in [22] was later extended to performing local tone mapping [44]. Essentially, the algorithm finds two reference sets of cut points in the HDR value range. The first set of N points $\{l_i\}$ divides the HDR range into N segments of equal length. The second set of N points $\{e_i\}$ divides the HDR range into Segments each contains equal number of pixels. Here N is the number of display levels of the LDR devices and $N = 256$ in all our implementation. The method then use these two sets of cut points to form a new set of cuts points $\{le_i\}$ controlled by a parameter β as

$$le_i = l_i + \beta (e_n - l_n) \quad (7.1)$$

where $0 \leq \beta \leq 1$ controls the amount of contrasts in the mapped LDR image. All HDR pixels falling in the interval between le_i and le_{i-1} are mapped to the same LDR display level d_i . As with many other methods, only the luminance channel is compressed and the colour information is unchanged.

If $\beta = 0$, the mapping becomes linear scaling and if $\beta = 1$, the mapping becomes histogram equalization mapping. The authors suggested that a good mapping is achieved with $0 \leq \beta \leq 1$. With $\beta = 0.5$ often gave reasonably good results.

Selecting the controlling parameter is an art and is data dependant. In Duan et al [44] an algorithm based on the pixel distributions has been used to select the parameters automatically:

$$\beta = (1 - \beta_0)(1 - e^{-(20-\mu)}) \quad (7.2)$$

where β_0 is the default value and set to 0.6 and μ is calculated as

$$\mu = \frac{\sum_{i=0}^1 |H(i) - M|}{\#Bins} \quad (7.3)$$

Where H is the pixel histogram and $H(i)$ is the normalized pixel population in i^{th} bin and M is the average of $H(i)$. Essentially μ measures the uniformity of the image (patch).

We propose to use the saliency map of **Algorithm 1** to control the parameter β and the new saliency modulated tone Mapping (**SMTM**) is described in **Algorithm 2**.

Algorithm 2 SMTM: Saliency Modulated Tone Mapping

Initialization: Compute HDR Saliency Map using Algorithm 1

- 1) Partitioning the image into blocks of $n \times n$ pixels $\{n = 32$ is set as default}
 - 2) **for** $i = 1$ to $\{\#$ of blocks} **do**
 - 3) $S_i \leftarrow$ Average Saliency of i^{th} block
 - 4) $\beta_i = \beta_{min} + \left(\frac{S_i}{S_{max}}\right)^p (1 - \beta_{min})$ $\{p$ is positive; $p = 1.5$ and $\beta_{min} = 0.2$ are set by default}
 - 5) Tone map block i with the tone mapping operator of (1) and set $\beta = \beta_i$
 - 6) **end for**
 - 7) Post processing to remove block boundaries
-

In the new saliency modulated tone mapping, a region with a high saliency will have a larger β value which means that these regions will be mapped with high contrasts. Although the method of equation (7.2) will assign a larger β values to more incongruence regions, it is purely based on local variance which is in contrast to the saliency map which identifies regions that will attract visual attention based on

The centre surrounds mechanism. As we will show later that the saliency modulated algorithm is better able to highlight important areas in the LDR images.

7.4 Experimental results

We have implemented the saliency modulated tone mapping (SMTM) in software mentioned in Chapter 3. We have made β_{min} in the SMTM adjustable with a simple sliding bar to enable users to easily control the amount of details in the final LDR image. Some examples of SMTM results are shown in Figure 7.1, Figure 7.2 and Figure 7.3. More examples can be seen in Figure 7.6-15. In general, the technique works very well and is competitive to state of the art techniques.



Figure 7.1 From top to bottom, left to right: SMTM result, saliency map, result of [44], amplified sections of SMTM and [44] respectively. Note here that light beams coming down the ceiling and shining on the curtain become clearly visible in the SMTM result as clearly shown in the amplified sections. From the saliency map, it is clear these regions have high saliency and should attract attention.

As can be seen from these results, SMTM generally rendered salient regions very well. For example, in Figure 7.1, the light beams coming down from the ceiling shining on the curtains are areas with relatively high saliency, whilst using the method of equation (7.2) has failed to bring out these features; SMTM has

successfully displayed this important information. Another example is shown in Figure 7.2, in this case regions around the windows areas have high visual saliency and are in general attract more visual attention, it is clearly seen that SMTM has successfully preserved finer details in these areas whilst the method of (7.2) has not been able to do so.



Figure 7.2: From top to bottom, left to right: SMTM result, saliency map, result of [44], amplified sections of SMTM and [44] respectively. Note here also that the SMTM results rendered the areas outside the windows with much more details and these areas have high saliency as shown in the saliency map.

Figure 7.3 shows a comparison with the gradient domain tone mapping result [70]. In this example, areas under the ceiling light, e.g., the black notice board areas and regions outside the glass door are highly salient. It is seen that both SMTM and the gradient domain techniques have rendered outdoor features very well while the

SMTM has managed to make the notice board areas under the light beam with more details and more easily visible. Note all results presented here were obtained using the default state of the algorithm.



Figure 7.3: Top left: Saliency map, Top Middle: Result of SMTM, Top right: Gradient domain method of [70]. Bottom Left: Section of SMTM result, Bottom right: Section of gradient domain method. Note here that the SMTM method renders the regions outside the glass door equally as well as the gradient method but has made the notice board much more conspicuous than the gradient domain result as this area of the image is also highly salient.

In general, evaluating tone mapping operators is very difficult and costly since it is heavily subjective. Recent advances in internet technology has seen researchers exploring web2.0 and cloud-sourcing as an inexpensive and convenient way to perform large scale subjective HDR tone mapping operator assessment [89]. There has also been effort to develop objective quality metrics for assessing HDR tone

mapping operators, for example [37] attempts to assess the structure similarities of the HDR and tone-mapped LDR image. Saliency has also been found to be useful in developing objective quality measures [84] [90]. In the SMTM operator, the saliency is explicitly incorporated in making saliency regions remain salient.

To verify that the SMTM can indeed preserve the saliency features of the HDR image in the compressed LDR image, we have designed and conducted a subjective study. Instead of the usual approach of asking observers to rate the subjective qualities of the images, we set out to find out the relations between the saliency computed from the HDR data and regions that attract viewer's attention in the LDR images. Ten images as shown in Figure 7.4 rendered by the new SMTM and the method of [44] were sent to 10 observers.

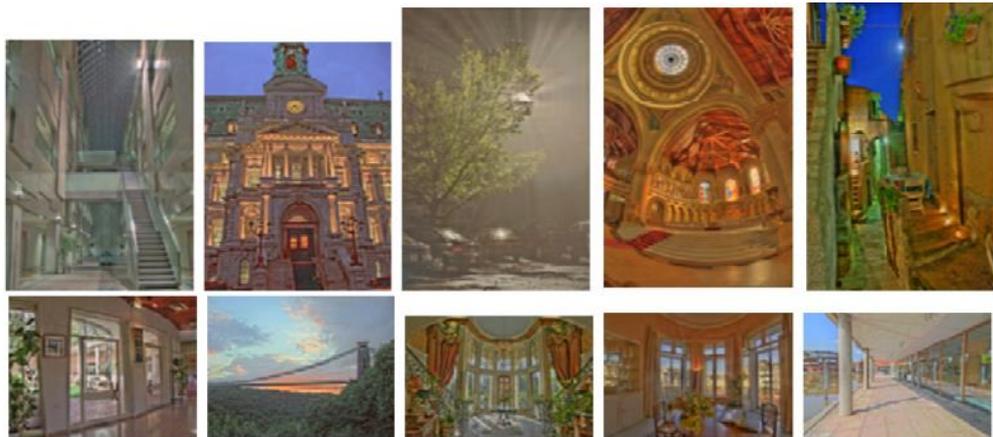


Figure 7.4: 10 HDR images used in the subjective study.

We asked the subjects to mark regions that they think the SMTM mapped images stand out and attract their attention. We then divided each image into $8 \times 8 = 64$ patches and calculate the average saliency of each patch. The relation between the patches' average saliency and the chance of it being selected as standing out and attracting attention is shown in Figure 7.5.

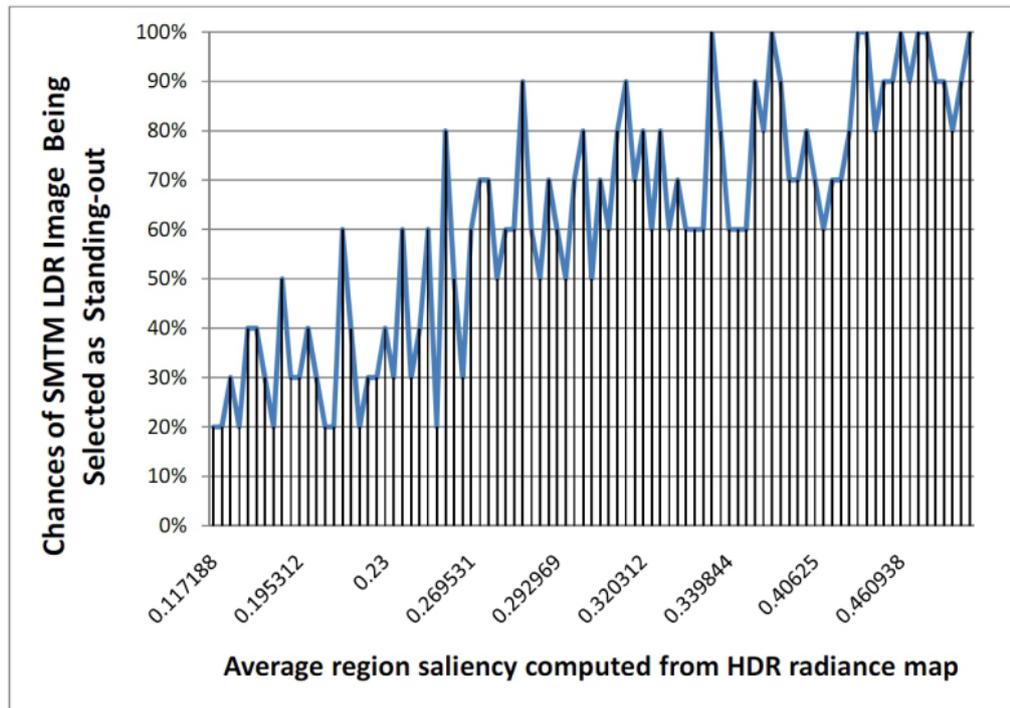
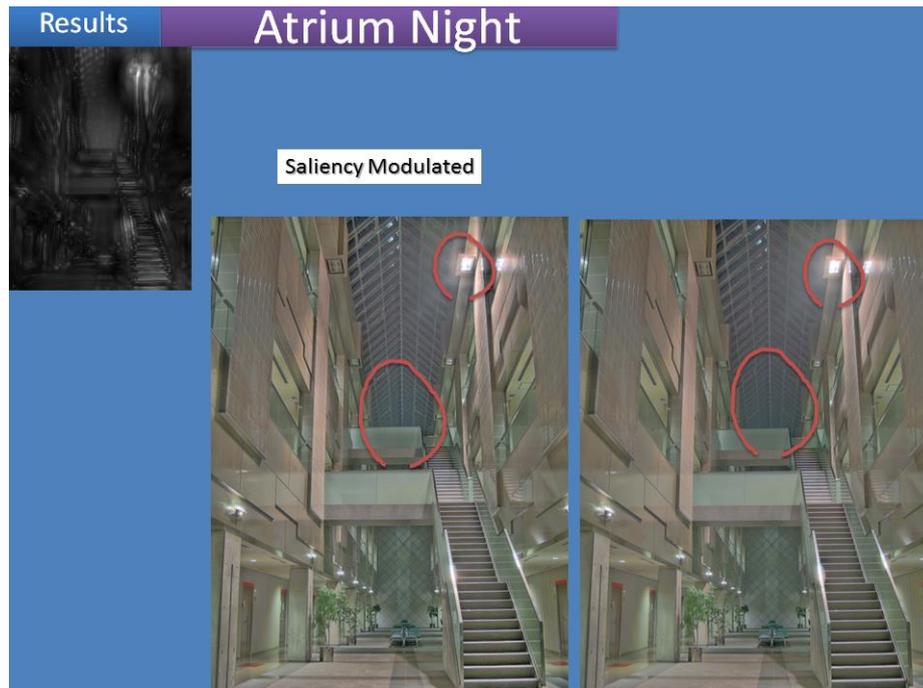
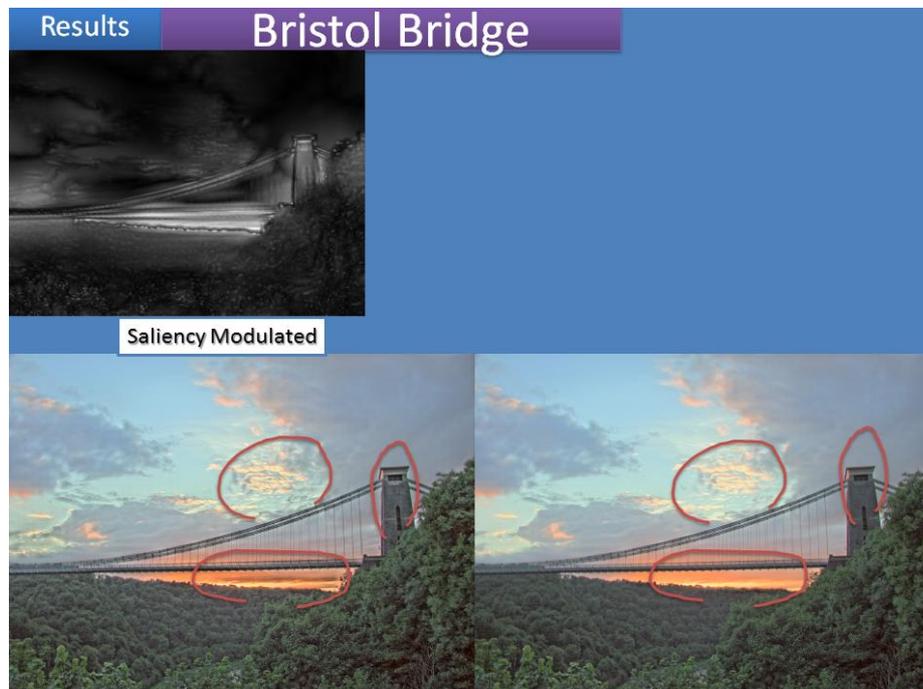


Figure 7.5: Relationship between region saliency and chances of the SMTM LDR image being selected by observers as standing out and attract their attentions

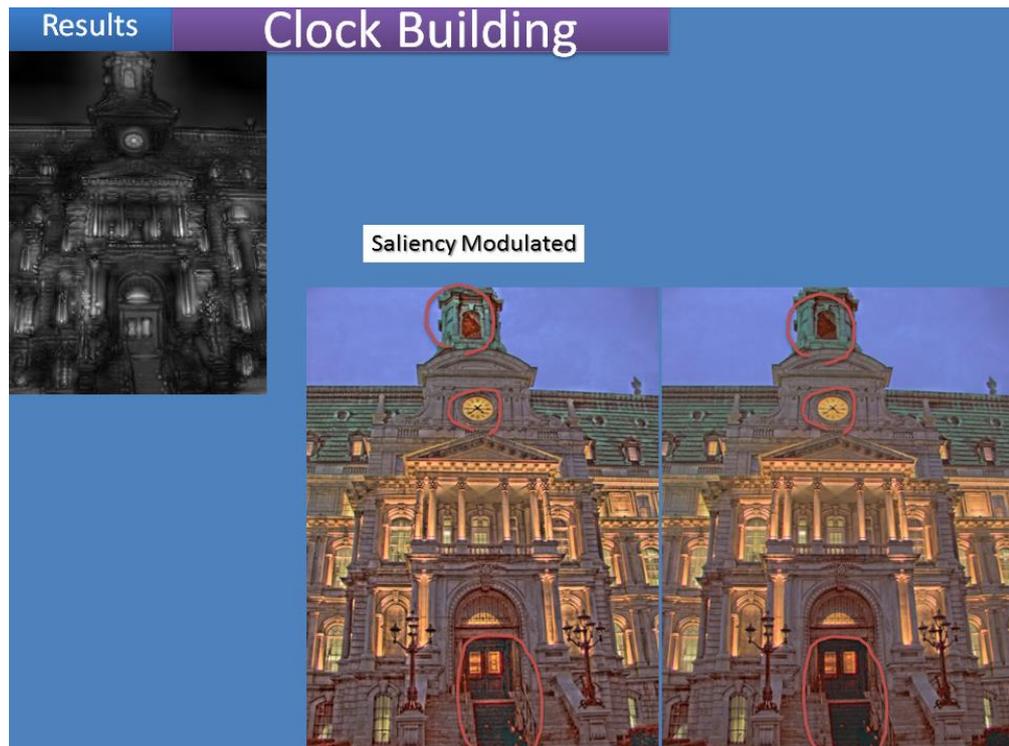
It is seen that although the relation is not monotonic, the general trend is that the higher the saliency, the more likely the region is recognized as standing out. It is interesting to observe that no patch with an average saliency of 0.1 was picked. Note that the saliency is computed based on HDR radiance data not the tone-mapped LDR images. Examples of observers marked regions are shown in the LDR images in Figures 7.6, 7.7, 7.8, 7.9 and 7.10.



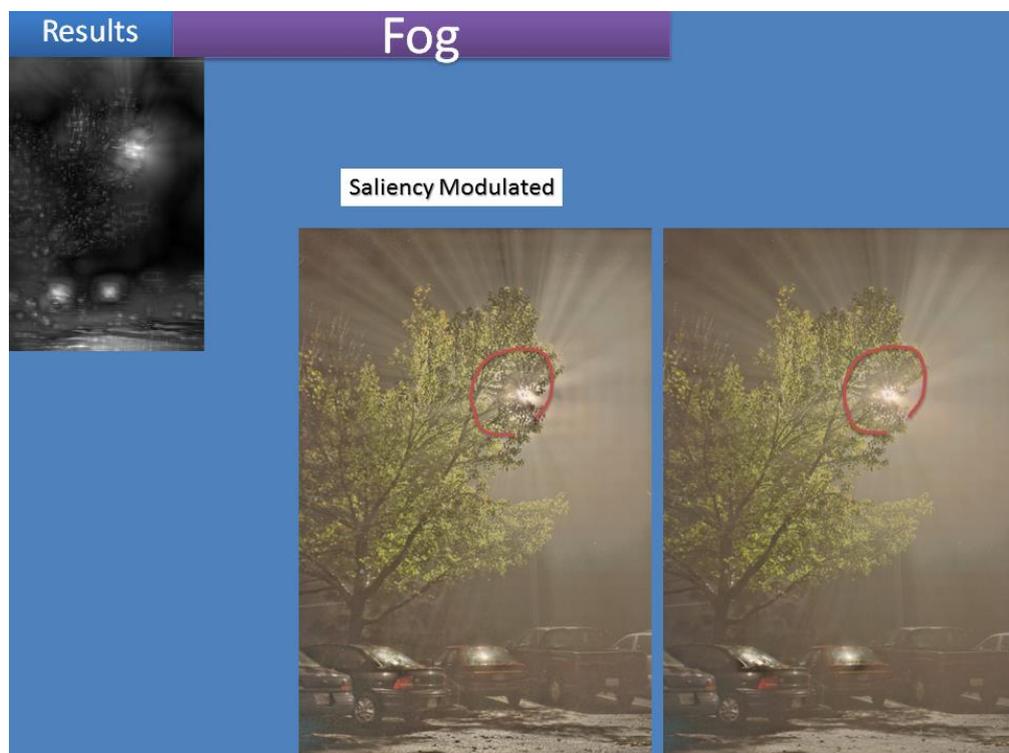
Figures 7.6: Experimental result. Left top: Saliency map. Left result image: SMTM mapped image. Right result image: gradient domain tone mapping result [70]. Regions that observers think the SMTM mapped images stand out and attract their attention are marked as red circle.



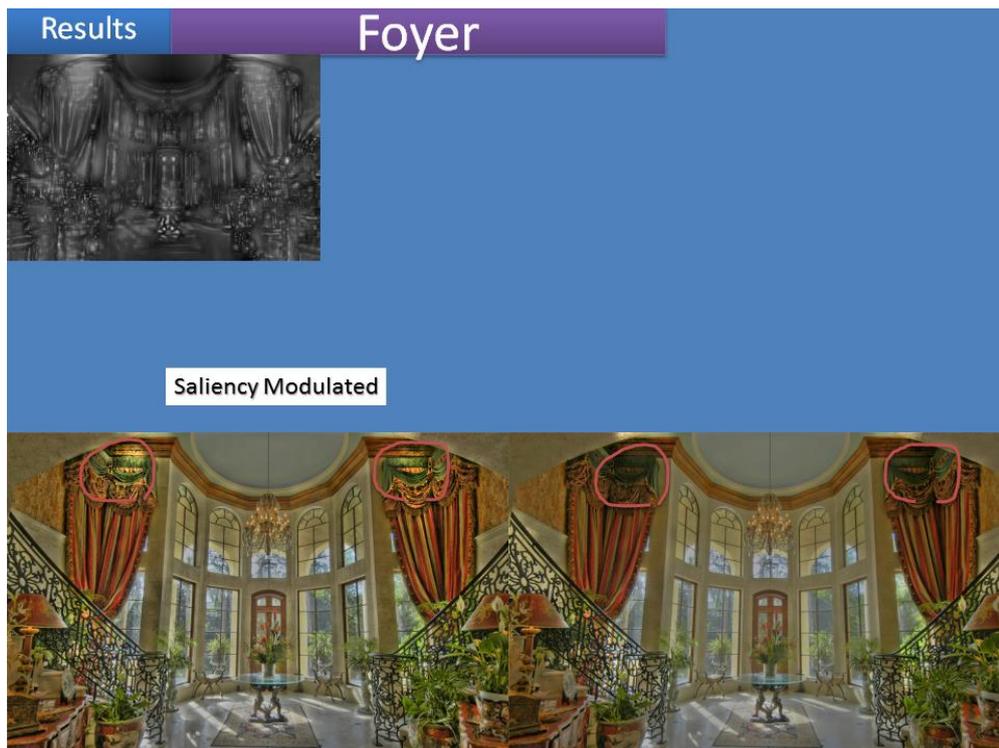
Figures 7.7: Experimental result. Same as Figure 7.6.



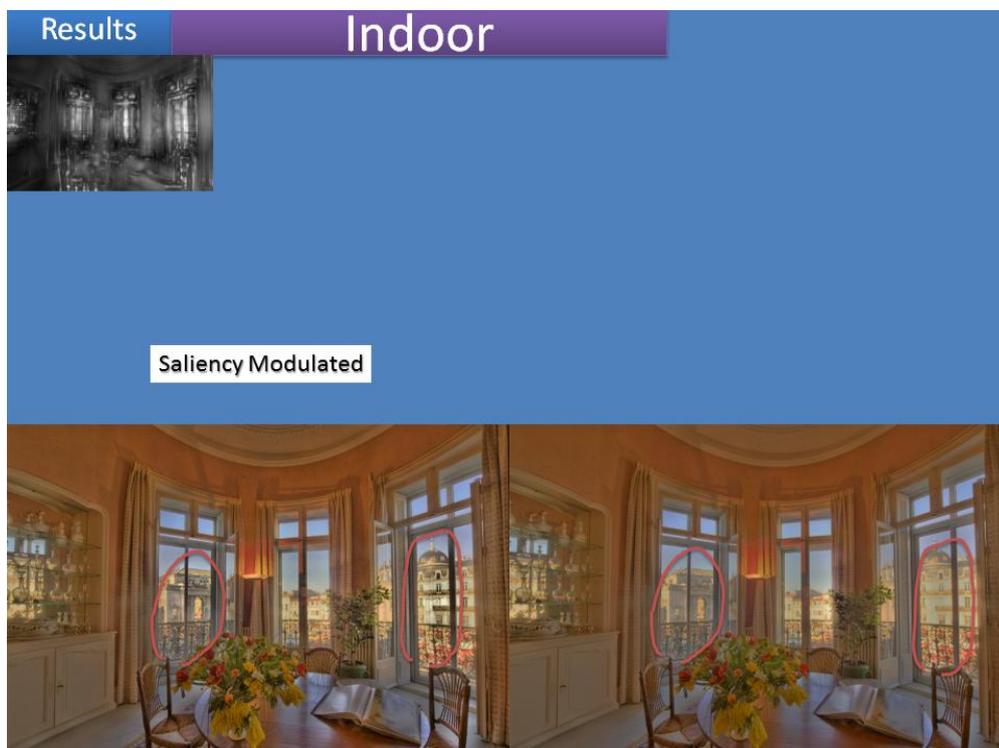
Figures 7.8: Experimental result. Same as Figure 7.6.



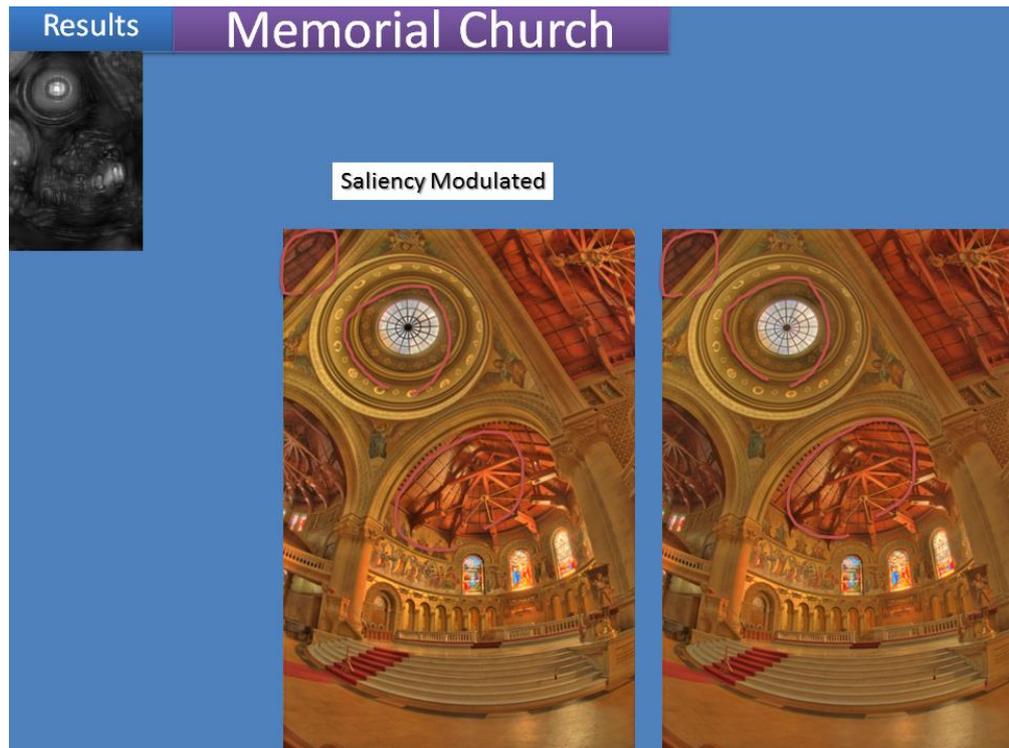
Figures 7.9: Experimental result. Same as Figure 7.6.



Figures 7.10: Experimental result. Same as Figure 7.6.



Figures 7.11: Experimental result. Same as Figure 7.6.



Figures 7.12: Experimental result. Same as Figure 7.6.



Figures 7.13: Experimental result. Same as Figure 7.6.



Figures 7.14: Experimental result. Same as Figure 7.6.



Figures 7.15: Experimental result. Same as Figure 7.6.

Since the HDR data is not directly viewable, it is not known exactly what the scene should look like. The introduction of the saliency map at least gives some guidance as to which part of the image is more conspicuous and more likely to attract attention and is therefore more important. The SMTM operator directly uses the saliency map to modulate the local tone mapping curve to ensure that regions of high saliency are rendered with higher contrast such that the saliency of the regions are preserved when the dynamic range of the full image is compressed to fit into the low dynamic range display. Although there may be different ways to exploit the saliency map, the SMTM techniques we introduce in the previous section has shown promises in rendering a high quality LDR display that also preserves the correct region saliency as demonstrated by the results presented in this section.

7.5 Summary

In this chapter, we have presented a new HDR tone mapping technique - saliency modulated tone mapping (SMTM), that employs visual saliency map to identify regions of the HDR image that are more likely to attract visual attention and directly use the saliency map to modulate the tone mapping of local regions. For highly salient regions, higher contrast and more details are rendered. We have presented experimental results which demonstrated that the SMTM technique is competitive with state of the art techniques in rendering visually pleasing LDR displays of HDR images. We have also demonstrated that SMTM can better preserve the visual saliency in the LDR images through conducting a user study where subjects marked out regions that make SMTM rendered image standing out as attracting visual attention. As the HDR data is not directly viewable, the saliency map may provide an important cue to guide the mapping of HDR image for display in LDR devices.

Chapter 8

Pair Comparison Evaluating System for Tone Mapping Algorithms

8.1 Introduction

High dynamic range (HDR) photography technology promises to bring new visual experiences to photograph consumers and to provide new means and capabilities for photographers to take more dramatic photographs.

An important part of the HDR photography workflow is the reproduction of the HDR photographs in conventional LCD monitors or photography prints. These reproduction media normally have a lower dynamic range than the input HDR images. In order to reproduce the HDR photographs in these conventional media, it is necessary to apply a tone mapping operator to render the HDR photos to a low dynamic range (normally 8 bits per pixel) image.

However, evaluating the relative merit of a tone mapping operator is a challenging task. Most authors use some form of limited subjective judgments to evaluate the results of a tone mapping operator (hence HDR photos). Comprehensive and systematic evaluation of tone mapping operators (hence HDR photographs that can be viewed by most consumers) is lacking.

In this chapter, we exploit Web 2.0 technology to evaluate the HDR photographs. With Web 2.0 technology, Internet users from anywhere in the world can access the website at any time thus can make a contribution to the evaluation process. Since typical consumers will be using all sorts of equipment and will be viewing the photographs under a variety of viewing conditions, evaluation results collected in this way may better reflect user preferences than evaluation results obtained under laboratory settings.

8.2 Related work

8.2.1 High dynamic range image display

The bit depth of high dynamic range (HDR) images is normally 32 bits per pixel or even higher. One of the biggest technical challenges in HDR imaging is how to display HDR images on conventional LCD displays or print them on printing papers. Since these reproduction devices have limited dynamic range, it is necessary to use a technique generally known as tone mapping to compress the dynamic range. We studied this in previous two chapters along with many others in the literature ([13], [14], [16], [17], [28], [50], [70] and [71]). However, it is known that it is extremely difficult to evaluate a tone mapping operators, and comparing one operator to the others is even more difficult. There has so far no serious attempt to systematically evaluate HDR tone mapping operators.

8.2.2 Image quality metrics

Visual quality of an image is a subjective concept. An objective quantitative image quality measure has widespread applications in image coding and image rendering. There are two types objective image quality measures, referenced and non-referenced. In a referenced quality measure, there is usually an original image,

the simplest quality measure is the mean square error (MSE) of an image and the reference (original image in the case of coding/compression). As has been recognized long ago, MSE is not a good visual quality measure because it does not correlate monotonically with visual quality in the sense that an image with a larger MSE is not necessarily visually inferior to an image with a smaller MSE, or vice versa. This phenomenon has motivated researchers to develop various other non-absolute error based quality measure such as the structured similarity index (SSIM) [93]. A crucial piece of information for building the image quality metrics is the so-called “ground truth” data which is normally obtained in a laboratory environment through subjects manually scoring the image quality [63], [59], [94], [93]. Such ground truth has a number of limitations. Firstly, it is not easy to collect large enough sample sizes because there are limited numbers of participants. Secondly, in the real world, the viewing environments and devices will be very different from those in the laboratory settings and this would mean that image quality metrics built using such ground truth will be biased towards laboratory environments.

8.2.3 Ranking with partial order and rank aggregation

The need to meaningfully combine sets of rankings is found in many applications including www search, information retrieval and many other real world applications [2], [20], [38], [78] and [88]. In our current application, we will use pair wise comparison; that is, for any given pair of images, users rank them according to their perceived visual quality. Based on these pair comparison data, we want to rank all images according to their perceived visual quality. This is a problem of ranking with partial order (because for any given image, we only know another one that is better or worse), this is a problem with long history and can be traced back to Condorcet [38].

8.2.4 Crowd sourcing, web2.0, social media, games with a purpose

This work employs Web2.0 technology and is related to crowd sourcing [39], [69] which is an emerging technology that employs the Internet to distribute tasks to internet users and to harvest group intelligence for solving difficult scientific tasks such as protein structure recognition [73] and image labeling [58]. Related current research includes games with a purpose (GWAP) where multiplayer games are used to collect human intelligence for solving specific problems [56].

8.2.5 Paired comparison and psychometrics

The method of paired comparison is a straightforward way of presenting items for comparative judgment. Items to be judged are presented in pairs to one or more judges who select the item that best satisfies the specified judgment criterion. The method of paired comparison is used in the scientific study of preferences, attitudes, voting systems, social choice, public choice, and multiagent AI systems. It has been shown that the method can yield an interval-scale ordering of items along a dimension such as preference or importance [33]. We use paired comparison in this study as the basic experimental protocol.

8.3 High dynamic range photo evaluation using web2.0

8.3.1 Technology Infrastructure and experiment design

We adopt a paired comparison strategy and a simple experimental protocol design. A web server with a simple front page as shown in Figure 8.1 has been designed and implemented to collect user input (not online at the moment due to system

restriction, but source website code is available on request). We have adopted a simple pair comparison strategy.

During the experiment period, the web server was accessible from anywhere at any time via the Internet and Internet users can participate in the voting. It is our intention to keep this server running to continuously collect viewer inputs. It also enables tone mapping developers to upload their results onto the server. We hope that such facility will facilitate comparison of different techniques and ultimately become a useful resource of evaluating tone mapping operators for the research community.



Figure 8.1: A screen shot of the web server interface. The viewer is presented with two versions of the same image and is asked to either click on the one that he/she thinks is better or click on the “no difference” button. The user input is recorded on the server and put into a database for analysis.

8.3.2 Data collection and processing

Thirteen HDR radiance maps each rendered by 6 to 10 tone mapping operators are placed on the web server. The tone mapping operators are (users and tone mapping researchers can upload more images):

- Adaptive logarithmic mapping for displaying high contrast scenes [16]
- Contrast reduction using LCIS [50]
- Display Adaptive Tone Mapping [71]
- Dynamic Range Reduction Inspired by Photoreceptor Physiology [13]
- Fast Bilateral Filtering [17]
- Gradient Domain High Dynamic Range Compression [70]
- Hierarchical Tone Mapping [23]
- Local Histogram Adjustment [44]
- Photographic Tone Reproduction for Digital Images [14]
- Tone mapping operator of Ward Larson et al. [28]

The paired comparison page (please refer to Figure 8.1 and the website) chooses two tone-mapped images of the same file. To ensure that every image will receive similar number of votes, we have developed a simple rule of image selection to ensure that the images that have received smaller number of votes will have a higher probability to be chosen, we also take care to ensure that each HDR maps will receive similar number of votes.

We then calculate the rankings for different tone-mapped versions of the same HDR file. The image quality ranking index (IQRI) of a tone-mapped image is calculated based the number of votes that image received. Let n_w be the number of times an image wins the comparison and n_d is the number of time it draws the comparison, the IQRI is calculated as

$$IQRI = \frac{\text{Total Number of Votes}}{n_w + \frac{n_d}{2}}$$

A lower IQRI value corresponds to a better image quality. Note that IQRI is a relative index, the smallest IQRI in a group of images will take the value of 1 (rank 1), the 2nd smallest IQRI will take the value of 2 (rank 2), etc.

8.4 Results and analysis

The website has been collecting data for a period of 6 months for data collection. Users can access it to view the images and their ranking and also participate in the voting. It collected around 15,000 votes from more than 100 users. The website is now off-line due to logistical reason (but source website code is available on request).

Figure 8.2 and 8.3 show two examples of the rankings of two HDR files by several different tone mapping algorithms. We can make the following observations about the IQRI ranking of the tone-mapped images.



Figure 8.2: Tone-mapped image ordered according to their IQRI from left to right and top to bottom.



Figure 8.3: Tone-mapped image ordered according to their IQRI from left to right and top to bottom.

The one that is ranked at the top seems to always stand out as having a better overall visual quality than the rest. However, for some images the difference between the top ranked and the ones ranked in the 2nd and 3rd do not have obvious difference. Those middle-ranked images seem to not differ from each other very much, but do seem to be poor than the top ranked and better than those ranked behind. Those ranked last always have inferior quality. The results suggest that pair comparison and our Web2.0 based experimental platform is a valid method for ranking tone mapping operators. We have only implemented the most straightforward method to process the paired comparison data to obtain the full ranking of the images, other methods exist and in a companion paper we have presented a method based on Dykstra's extension of Bradley-Terry method [21].

To understand how the IQRI relate to other quantitative measures of the tone-mapped images, Figures 8.4, 8.5 and 8.6 plot the relation between the IQRI values of tone-mapped images and their average brightness, average edge magnitude and average local variances. From Figure 8.4, it is seen that the average brightness differences between the rank-one image and others increases with IQRI, meaning that the larger IQRI is, the larger is the average brightness difference from the rank-one image. Although the relation is not exactly monotonic, the overall trend seems to be very clear. This suggests that there may be an "optimal" average brightness for an image that will result in it being ranked at the top. The optimal average brightness seems to be image dependent. Although determining the optimal average brightness value needs further study and is beyond the scope of this chapter, this finding nevertheless has provided evidence to guide tone mapping operators that they should ensure the tone-mapped image to have the appropriate average brightness.

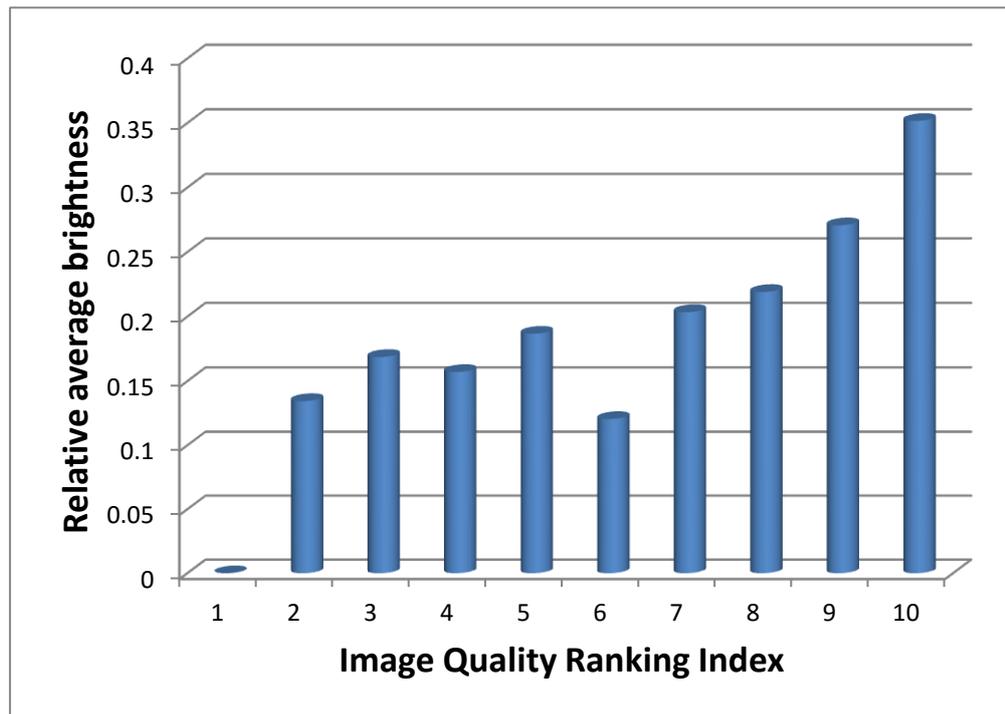


Figure 8.4: The relation between the average brightness of tone-mapped images and their IQRI's. We take the average brightness values of all images with the same rank and subtract the average brightness of the rank 1 images. The relative average brightness is defined as the difference between the average brightness of rank 1 images and the average brightness of images in the other ranks. Here it is seen that as the rank gets higher (quality gets worse), the relative average brightness gets larger (differ more from the top ranked image). Although the relation is not monotonic there does seem to have an overall trend.

In Figure 8.5, we plot the relation between IQRI and the average local variances. Again, the relation is not monotonic but there does seem to be an overall trend which is that a lower IQRI corresponds to a higher average local variance. A larger local variance corresponds to more image details and this is in agreement with findings in other contexts. Figure 8.6 plots the relation between the edge magnitude of an image and its IQRI ranking. Although the trend is again but no means

monotonic for every image, the overall trend for this set of image is monotonic and a lower IQRI corresponds to a higher average edge magnitude. This is again consistent with the literature that a high quality image should have as much detail as possible (edges correspond to details).

As can be expected, trying to correlate the quantitative measurements of a tone-mapped image with the qualitative measurement of IQRI is not easy and there is no straightforward simple relationships, nevertheless, the current study have demonstrated that the Internet and the Web2.0 technology can be exploited to effectively rank HDR tone mapping operators. Potentially, it will enable the collection of large samples which may enable future studies to systematically related quantitative measure with qualitative indices.

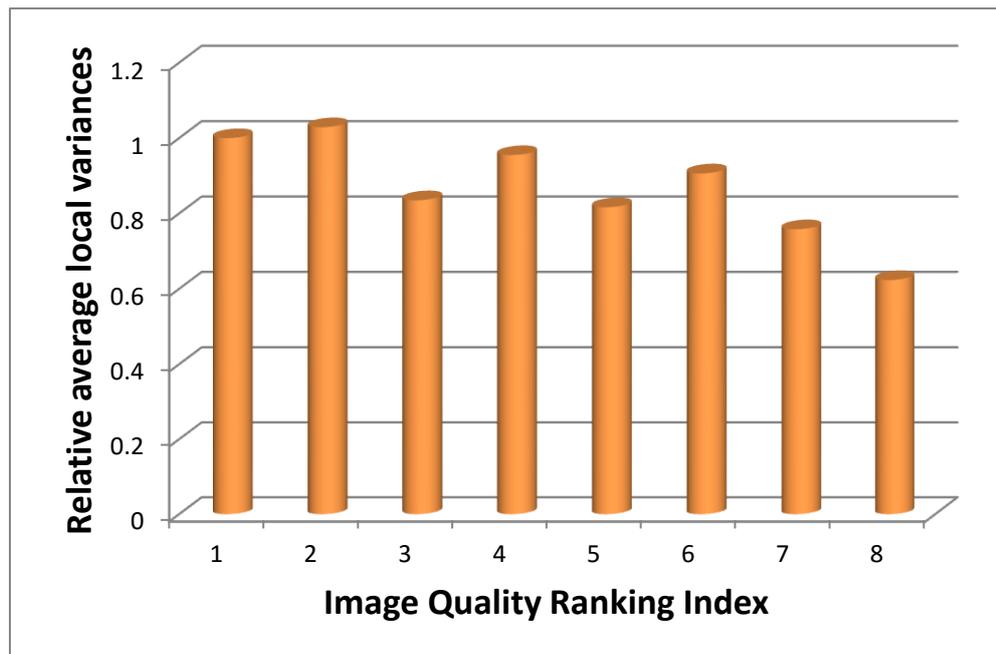


Figure 8.5: The relation between the average local variance of tone-mapped images and their IQRI's. We take the average local variance values of all images with the same rank and normalized them by the average variance of the rank 1 images. The relative average local variance is defined as the ratio of the average variance of images in the other ranks over the average variance of rank 1 images. Here it is seen that as the rank gets higher (quality gets worse), the relative average variance gets

smaller (differ more from the top ranked image). Although the relation is not monotonic there does seem to have an overall trend.

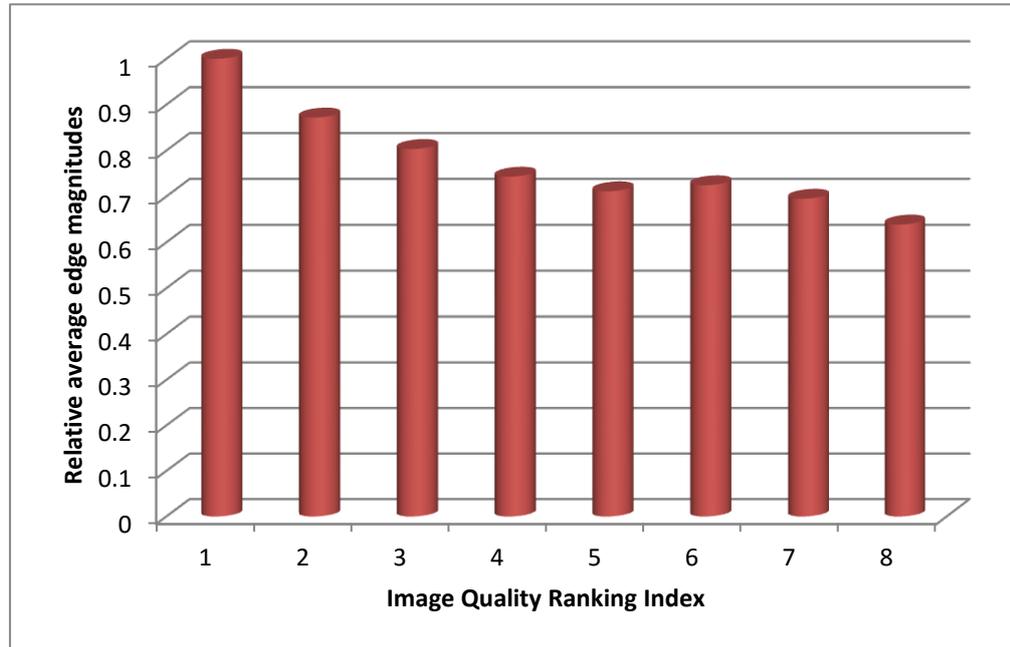


Figure 8.6: The relation between the average edge magnitudes of tone-mapped images and their IQRI's. We take the average edge magnitude values of all images with the same rank and normalized them by the average edge magnitude of the rank 1 images. The relative average edge magnitude is defined as the ratio of the average variance of images in the other ranks over that of rank 1 images. Here it is seen that as the rank gets higher (quality gets worse), the relative average edge magnitude gets smaller (differ more from the top ranked image). Again, although the relation is not monotonic there does seem to have an overall trend.

8.5 Summary

In this chapter, we have reported our attempt to use the Internet technology to evaluate high dynamic range image tone mapping algorithms. We have implemented a Web2.0 style web server to enable Internet users from anywhere at any time to perform paired comparison of HDR tone-mapped images. We have presented some

preliminary results which demonstrated the potential of using the Internet to develop new paradigms to harvest the brain power of the mass to tackle the long standing issue of image quality assessment.

Chapter 9

Conclusions and Future Work

9.1 Contributions

In This thesis, I have made the following original contributions to the field of high dynamic range imaging:

In Chapter 5, a defect in the excellent technique introduced by Debevec and Malik for recovering high dynamic range radiance maps from photographs is discovered. And a novel technique to correct this defect is proposed. This technique creatively uses each frame's shutter speed and pixel values together to calculate the weighting, which makes sure well-exposed pixels get larger weighting, and always produce satisfactory results, particularly in bright and dark areas.

Chapter 6 presented a novel tone mapping framework. In this framework, firstly we introduce a tone mapping fidelity principle which explicitly stipulates that tone-mapped image data should not only be visually enhanced but should also stay faithful to the original image. Second, this principle naturally translates tone mapping into a constrained optimization problem where a two-term cost function, one measures the difference between the tone mapped image and a visually enhanced version of the image, and the other measures the difference between the tone mapped image and the original image, is optimized. The relative weightings of

the two terms in the cost function not only offers an insightful and simple mechanism to control the appearance of the tone mapped image but also enables the introduction of spatially varying or uniform weighting functions thus unifying local and global tone mapping in a single framework.

Chapter 7 presented a new HDR tone mapping technique - saliency modulated tone mapping (SMTM), that creatively employs visual saliency map to identify regions of the HDR image that are more likely to attract visual attention and directly use the saliency map to modulate the tone mapping of local regions. We have presented experimental results which demonstrated that the SMTM technique is competitive with state of the art techniques in rendering visually pleasing LDR displays of HDR images. We have also demonstrated that SMTM can better preserve the visual saliency in the LDR images through conducting a user study where subjects marked out regions that make SMTM rendered image standing out as attracting visual attention.

Chapter 8 introduced a novel approach - pair comparison using Web 2.0 technology for evaluating tone mapping algorithms. In this approach, a Web2.0 style system is implemented for any users to take part in the evaluation from anywhere and at any time. The web server can collect users input data and analyse them by a rank aggregation algorithm.

In addition, a full-functioning HDR software system - HDR Photographer is implemented. It could control Nikon SLR camera to capture HDR image directly, and it also provides tone mapping and radiance map recovery functions. It was announced in IMA seminar on 22/05/2012.

9.2 Future work and closing comment

HDR imaging is a relatively new and great research topic. Although it has attracted

more and more attention in recent years, it mostly remains in the laboratories as a research field. To make the technology into daily use, there are still many unsolved problems.

Currently the most popular technique for acquiring the HDR radiance map is based on taking multiple photographs of a scene. However, this will only work when the scene is static or with relatively small movements. To make the technique into everyday use, new HDR image capture methods are needed.

As current display technologies cannot cope with high dynamic range image directly, tone mapping is a must step in the HDR imaging pipeline. As what is presented to the user in the end is tone mapped product, this is a crucial step. Despite many published tone mapping techniques, including the two new methods presented in this thesis, this is still an unsolved problem. Firstly, no single method will be able to always render any given scene correctly. Secondly, all methods have adjustable parameters which are always very difficult to set to the right values systematically. Therefore, developing better tone mapping solutions that will work for any scenes and can set the correct parameters automatically is very important and remains a challenge for future research.

One of the main applications of HDR imaging is consumer photography. To automatically measure the quality of the HDR image will be very useful but remains a very challenging problem. In this thesis, we have made a start to employ the latest development in Internet technologies and crowdsourcing to collect data to develop robust image quality metrics. However, there are still many unanswered questions, including, how to ensure the quality of the data collection, once we have high quality data, how to use it to development quality metrics need much research.

Bibliography

- [1] Adobe. Tiff 6.0 specification, 1992, Available at <http://partners.adobe.com/asn/tech/tiff/specification.jsp>.
- [2] A. Klementiev, D. Roth and K. Small, A Framework for Unsupervised Rank Aggregation. Proc. of the SIGIR Workshop on Learning to Rank for Information Retrieval (2008) pp. 32-39
- [3] AMP HDR. Available at: <http://amphdr.com/>
- [4] B. A. Wandell. Foundations of vision. Sinauer Associates, 1995
- [5] C. Pal, R. Szeliski, M. Uyttendaele and N. Jojic: Probability models for high dynamic range imaging. In Proc. IEEE Computer Society Conference on Computer Vision and Pattern Recognition CVPR 2004, volume 2, pages 173-180, June 2004.
- [6] C. Tomasi and R. Manduchi. 1998. Bilateral filtering for gray and color images. In Proc. IEEE Int. Conf. on Computer Vision, 836–846
- [7] D. Gao, V. Mahadevan, and N. Vasconcelos, “The discriminant center-surround hypothesis for bottom-up saliency,” Advances in Neural Information Processing Systems, vol. 20, pp. 1–8, 2007.
- [8] D. H. Hubel and T. N. Wiesel. “Receptive fields and functional architecture in two no striate visual areas (18 and 19) of the cat,” Journal of neurophysiology, vol. 28, pp. 229–89, Mar. 1965.
- [9] D. J. Jobson, Z. Rahman and G. A. Woodell, “A multiscale Retinex for dridging

- the gap between color images and the human observation of scenes”, IEEE Transactions on Image processing, vol. 6, pp. 965-976, 1997
- [10] D. Mumford and J. Shah, “Optimal approximations by piecewise smooth functions and associated variational problems”, Communications in Pure and Applied Mathematics, vol. 42, no. 4, 1989
- [11] E. H. Land and J. J. McCann, “Lightness and retinex theory”, Journal of the Optical society of America, vol. 61, pp. 1-11, 1971
- [12] E. Niebur and C. Koch, “Control of selective visual attention: modeling the where pathway,” in In D. S. Touretzky, M. C. Mozer, & M. E. Hasselmo (Eds.), Advances in neural information processing systems (vol. 8) (pp. 802-808). Cambridge, MA: MIT Press., 1996.
- [13] E. Reinhard and K. Devlin. 2005. Dynamic Range Reduction Inspired by Photoreceptor Physiology. IEEE Transactions on Visualization and Computer Graphics 11, 1 (January 2005), 13-24.
- [14] E. Reinhard, M. Stark, P. Shirley, J. Ferwerda, Photographic tone reproduction for digital images, SIGGRAPH2002, 267–276, 2002
- [15] E. Reinhard, W. Heidrich, S. Pattanaik, P. Debevec, G. Ward, and K. Myszkowski: High dynamic range imaging: acquisition, display, and image-based lighting Morgan Kaufmann, 2010
- [16] F. Drago, K. Myszkowski, T. Annen and N. Chiba. (2003), Adaptive Logarithmic Mapping For Displaying High Contrast Scenes. Computer Graphics Forum, 22: 419–426. doi: 10.1111/1467-8659.00689
- [17] F. Durand and J. Dorsey. “Fast Bilateral Filtering for the Display of High Dynamic Range Images”. Siggraph 2002.
- [18] F. Kainz, R. Bogart and D. Hess. The Openexr Image File Format SIGGRAPH

Technical Sketches, 2003

[19] F. X. Sillion and C. Puech. Radiosity and Global Illumination. San Francisco, CA: Morgan Kaufmann Publishers, Inc., 1994.

[20] G. Lebanon and Y. Mao, "Non-Parametric Modeling of Partially Ranked Data", Journal of Machine Learning Research 9 (2008) 2401-2429

[21] G. Qiu and A. Kheiri, "Social Image Quality", Proceedings of SPIE, vol. 7867, Conference on Image Quality and System Performance VIII, 23 – 27 January 2011, San Francisco, USA

[22] G. Qiu and J Duan, "An optimal tone reproduction curve operator for the display of high dynamic range images", ISCAS2005, Proc. of IEEE International Symposium on Circuits and Systems, vol. 6, pp. 6276- 6279, Kobe, Japan, 23 - 26 May 2005

[23] G. Qiu and J. Duan, "Hierarchical Tone Mapping for High Dynamic Range Image Visualization", Proceedings of VCIP2005, Visual Communication and Image Processing, 12 - 15 July 2005, Beijing, China

[24] G. Qiu, J. Duan, and G. D. Finlayson, "Learning to display high dynamic range images," Pattern Recogn., vol. 40, pp.2641–2655, October 2007.

[25] G. Qiu, J. Guan, J. Duan, M. Chen, "Tone Mapping for HDR Image using Optimization - A New Closed Form Solution", Proc. ICPR2006, vol. 1, pp.996-999, 2006

[26] G. Underwood, "Cognitive processes in eye guidance: Algorithms for attention in image processing," Cognitive Computation, vol. 1, pp. 64–76, 2009.

[27] G. W. Larson and R. A. Shakespeare. Rendering with Radiance. Morgan Kaufmann, San Francisco, 1998.

[28] G. W. Larson, H. Rushmeier, C. Piatko, A visibility matching tone reproduction

operator for high dynamic range scenes, *IEEE Trans.VCG*, pp. 291–306, 1997

[29] G. Ward (1991). “Real pixels”, In *Graphics Gems II*

[30] G. Ward and Anywhere Software. Available at

<http://www.anywhere.com/gward/hdrenc/>

[31] Gonzalez and Woods, *Digital Image Processing*, 3rd ed, Prentice-Hall, 2008

[32] Greg Ward. “A contrast-based scalefactor for luminance display”, In *Graphics Gems IV*, pp. 415–421, Academic Press, 1994.

[33] H. A., David, *The method of paired comparisons* (Second ed.), Oxford University Press, 1988

[34] H. Seetzen, W. Heidrich, W. Stuerzlinger, G.Ward, L. Whitehead, M. Trentacoste, A. Ghosh, A. Vorozcovs, "High Dynamic Range Display Systems", *ACM Transactions on Graphics (Siggraph 2004)*, 23(3): 760-768

[35] H. W. Jensen. *Realistic Image Synthesis Using Photon Mapping*. Natick, MA: AK Peters, 2001.

[36] H. Yee, S. Pattanaik, and D. P. Greenberg, “Spatiotemporal sensitivity and visual attention for efficient rendering of dynamic environments,” *ACM Trans. Graph.*, vol. 20, pp. 39–65, January 2001.

[37] H. Yeganeh and Z. Wang, “Objective assessment of tone mapping algorithms,” in *IEEE International Conference on Image Processing*, Hong Kong, China, Sept. 26-29, 2010.

[38] http://en.wikipedia.org/wiki/Condorcet_method

[39] <https://www.mturk.com/mturk/welcome>

[40] J. A. Ferwerda. Elements of early vision for computer graphics *Computer Graphics and Applications*, IEEE, IEEE, 2001, 21, 22-33

- [41] J. DiCarlo and B. Wandell, "Rendering high dynamic range images," Proc. SPIE, vol. vol.3965, p. 392 -401, 2001.
- [42] J. Duan and G. Qiu, "Fast Tone Mapping for High Dynamic Range Images", Proc. ICPR2004, pp. 847-850, August 2004
- [43] J. Duan, G. Qiu, G.D. Finlayson, Learning to display high dynamic range images, Pattern Recognition 40, 2641–2655, 2007
- [44] J. Duan, M. Bressan, C. Dance, and G. Qiu: Tone-mapping high dynamic range images by novel histogram adjustment. Pattern Recognition 43(5): 1847-1862 (2010)
- [45] J. Foley, A. van Dam, S. Feiner, and J. Hughes. Computer Graphics, Principles and Practice, 2nd ed., Addison-Wesley, Reading MA, 1990.
- [46] J. Herwig and J. Pauli. "Recovery of the Response Curve of a Digital Imaging Process by Data-centric Regularization". VISAPP 2009: 539-546
- [47] J. Kuang, H. Yamaguchi, C. Liu, G. M. Johnson, and M. D.Fairchild, "Evaluating hdr rendering algorithms," ACM Trans. Appl. Percept., vol. 4, July 2007.
- [48] J. Paul Getty Museum, G. L. G. Photographer, 2002
- [49] J. Shen, S. Fang, H. Zhao, X. Jin, and H. Sun, "Fast approximation of trilateral filter for tone mapping using a signal processing approach," Signal Processing, vol. 89, no. 5, pp. 901 – 907, 2009.
- [50] J. Tumblin and G. Turk. 1999. LCIS: a boundary hierarchy for detail-preserving contrast reduction. In Proceedings of the 26th annual conference on Computer graphics and interactive techniques (SIGGRAPH '99).
- [51] J. Tumblin and H. Rushmeier, "Tone reproduction for realistic images", IEEE Computer Graphics and Applications, vol. 13, pp. 42 – 48, 1993.

- [52] J. Tumblin, J. Hodgins, and B. Guenter, “Two methods for display of high contrast images”, *ACM Transactions on Graphics*, 18(3):56–94, 1999.
- [53] K. Perlin and E.M. Hoffert, “Hypertexture”, *Computer Graphics (Proceedings of ACM SIGGRAPH 89)*, ACM, 23, 253–262, 1989
- [54] L. Itti, C. Koch, and E. Niebur, “A model of saliency-based visual attention for rapid scene analysis,” *IEEE Transactions on pattern analysis and machine intelligence*, vol. 20, no. 11, pp. 1254–1259, 1998.
- [55] L. Meylan, D. Alleysson, and S. Su sstrunk, “Model of retinal local adaptation for the tone mapping of color filter array images,” *J. Opt. Soc. Am. A*, vol. 24, no. 9, pp. 2807–2816, Sep 2007.
- [56] L. V. Ahn, L. Dabbish, “Designing games with a purpose”, *Communications of the ACM*, vol. 51, no. 8, pp. 58 – 67, August 2008
- [57] L. Zhang and G. Cottrell, “A bayesian framework for dynamic visual saliency,” 2007.
- [58] LabelMe: <http://labelme.csail.mit.edu/>
- [59] LIVE Image Quality Assessment Database:
<http://live.ece.utexas.edu/research/quality/subjective.htm>
- [60] M. Ashikhmin and J. Goyal, “A reality check for tone- mapping operators,” *ACM Trans. Appl. Percept.*, vol. 3, pp.399–411, October 2006.
- [61] N. Bruce and J. Tsotsos. “Saliency based on information maximization,” *Advances in neural information processing systems*, vol. 18, p. 155, 2006.
- [62] N. Butko, L. Zhang, G. Cottrell, and J. Movellan, “Visual saliency model for robot cameras,” in *Robotics and Automation, 2008. ICRA 2008. IEEE International Conference on*, May 2008, pp. 2398 –2403.
- [63] N. Ponomarenko, et al, “Metrics Performance Comparison for Color Image

Database”, Fourth International Workshop on Video Processing and Quality Metrics for Consumer Electronics, Scottsdale, Arizona, USA

[64] P. A. Viola and M. J. Jones, “Rapid object detection using a boosted cascade of simple features,” in *Computer Vision and Pattern Recognition*, 2001, pp. 511–518.

[65] P. Dutré, P. Bekaert, and K. Bala. *Advanced Global Illumination*. Wellesley, MA: AK Peters, 2003.

[66] P. E. Debevec and J. Malik, “Recovering high dynamic range radiance maps from photographs,” in *ACM SIGGRAPH 2008 classes*, ser. SIGGRAPH '08. New York, NY, USA: ACM, 2008, pp. 31:1–31:10.

[67] P. E. Debevec and J. Malik. Recovering high dynamic range radiance maps from photographs, in *ACM SIGGRAPH '97*, 1997, 369-378

[68] P. Ferschin, I. Tastl, W. Purgathofer, "A comparison of techniques for the transformation of radiosity values to monitor colors", *Proc. IEEE Conference on Image Processing*, Austin Texas, 1994

[69] Q. Su, D. Pavlov, J. Chow, and W. C. Baker. 2007. Internet-scale collection of human-reviewed data. In *Proceedings of the 16th international conference on World Wide Web (WWW '07)*. ACM, New York, NY, USA, 231-240

[70] R. Fattal, D. Lischinski, and M. Werman. 2002. Gradient domain high dynamic range compression. *ACM Trans. Graph.* 21, 3 (July 2002), 249-256.

[71] R. Mantiuk, S. Daly, and L. Kerofsky. (2008) Display adaptive tone mapping. *ACM Transactions on Graphics*. Vol. 27, no. 3. Aug. 2008

[72] S. B. Kang, M. Uyttendale, S. Winder and R. Szeliski, “High dynamic range video”, *ACM Transactions on Graphics*, vol.22, no. 3, Pages: 319 – 325, July 2003

[73] S. Cooper, F. Khatib, A. Treuille, J. Barbero, J. Lee, M. Beenen, A. Leaver-Fay, D. Baker, Z. Popović and F. players. "Predicting protein structures with a

- multiplayer online game." *Nature*, volume 466 (05 August 2010), pp. 756-760
- [74] S. M. O'Malley. A simple, Effective System for Automated Capture of High Dynamic Range Images. In *Proc. IEEE international Conference on Computer Vision Systems ICVS '06*, pages 15-15, 2006
- [75] S. Mann and R. W. Picard. On Being 'undigital' With Digital Cameras: Extending Dynamic Range By Combining Differently Exposed Pictures *Proceedings of IS&T*, 1995, 442-448
- [76] S.S. Stevens and J.C. Stevens, "Brightness function: Parametric effects of adaptation and contrast", *Journal of the Optical Society of America*, 53, 1960.
- [77] SpheroCam HDR. Available at:
http://www.grafixgear.com/HTML/PDF/SpheroCam_HDR.pdf
- [78] T. Huang, C. Lin, and R. C. Weng. Ranking Individuals by group comparisons. *Journal of Machine Learning Research*, 9(2008), 2187-2216
- [79] T. Mertens, J. Kautz, and F. Van Reeth. "Exposure Fusion: A Simple and Practical Alternative to High Dynamic Range Photography". *Comput. Graph. Forum* 28(1): 161-171, 2009
- [80] T. Mitsunaga and S. K. Nayar, "High dynamic range imaging: Spatially varying pixel exposures", *Proc. CVPR'2000*, vol. 1, pp. 472-479, 2000
- [81] T.G. Stockham, "Image Processing in the Context of a Visual Model", *Proc. the IEEE*, 60, 7, 828-842, 1972
- [82] V. Setlur, T. Lechner, M. Nienhaus, and B. Gooch, "Retargeting images and video for preserving information saliency," *IEEE Computer Graphics and Applications*, vol. 27(5), 2007.
- [83] VIPER FILMSTREAM. Available at:
http://www.cinematography.net/Files/viper_brochure.pdf

- [84] X. Gu, G. Qiu, X. Feng, L. Debing, and C. Zhibo, "Region of interest weighted pooling strategy for video quality metric," *Telecommunication Systems*, pp. 1–11, 2010, 10.1007/s11235-010-9353-8.
- [85] Y. Bandoh, G. Qiu, M. Okudo, S. Daly, T. Aach and O. C. Au, "Recent advances in high dynamic range imaging technology", *ICIP2010*, 2010 IEEE International Conference on Image Processing, September 2010, Hong Kong
- [86] Y. Kim and A. Varshney, "Saliency-guided enhancement for volume visualization," *IEEE Transactions on Visualization and Computer Graphics*, vol. 12, no. 5, pp. 925–932, 2006.
- [87] Y. Li, L. Sharan, and E. H. Adelson. "Compressing and companding high dynamic range images with subband architectures," *ACM Trans. Graph.*, vol. 24, July 2005.
- [88] Y. Liu, T. Liu, T. Qin, Z. Ma, and H. Li. 2007. Supervised rank aggregation. In *Proceedings of the 16th international conference on World Wide Web (WWW '07)*. ACM, New York, NY, USA, 481-490
- [89] Y. Mei, G. Qiu, and J. Duan, "Evaluating hdr photos using web 2.0 technology," in *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, vol. 7867, 2011.
- [90] Y. Tong, H. Konik, F. A. Cheikh, and A. Tremeau, "Full reference image quality assessment based on saliency map analysis," *Journal of Imaging Science and Technology*, vol. 54, no. 3, p. 030503, 2010.
- [91] Y. Tsin, V. Ramesh and T. Kanade: *Statistical Calibration of CCD Imaging. Process*, *Proc. Int'l Conf. Computer Vision*, vol. I, pp. 480-487, 2001
- [92] Z. Rahman, D. J. Jobson, G. A. Woodell, *Retinex processing for automatic image enhancement*, *J. Electron. Imaging*, Vol. 13, 100 (2004). Images from: <http://www.truview.com/>

[93] Z. Wang and A. C. Bovik, "Mean squared error: love it or leave it? - A new look at signal fidelity measures," *IEEE Signal Processing Magazine*, vol. 26, no. 1, pp. 98-117, Jan. 2009

[94] Z. Wang, A. C. Bovik, H. R. Sheikh and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Transactions on Image Processing*, vol. 13, no. 4, pp. 600-612, Apr. 2004