FAKULTÄT FÜR !NFORMATIK
Faculty of Informatics

Applications of High Dynamic Range Imaging

Stereo Matching and Mesopic Rendering

PhD THESIS

submitted in partial fulfillment of the requirements for the degree of

Doctor of Technical Sciences

within the

Vienna PhD School of Informatics

by

Eng. Tara Akhavan
Registration Number 1128895

to the Faculty of Informatics

at the Vienna University of Technology

Advisor: Assoc. Prof. Mag. Dr. Hannes Kaufmann

External reviewers:
. Alan Chalmers, WMG, University of Warwick, United Kingdom.
. Kadi Bouatouch, ISTIC, University of Rennes I, France.

Vienna, 8th April, 2017

Tara Akhavan

Hannes Kaufmann

Declaration of Authorship

Eng. Tara Akhavan
Address

I hereby declare that I have written this Doctoral Thesis independently, that I have completely specified the utilized sources and resources and that I have definitely marked all parts of the work - including tables, maps and figures - which belong to other works or to the internet, literally or extracted, by referencing the source as borrowed.

Vienna, 8th April, 2017

Tara Akhavan

Acknowledgements

Firstly, I would like to express my deepest gratitude to my supervisor Hannes Kaufmann for his unlimited support, inspiration, motivation, belief, and guidance. My special thanks to Christian Breiteneder who was always there to help resolving hardest problems with wisdom and vision. I am very thankful to Alan Chalmers and Kadi Bouatouch who agreed to be my thesis external reviewers but most importantly planned valuable workshops through the course of my PhD and invited me to network with and learn from the experts in the field.

Part of my PhD research was conducted at Tandemlaunch Inc. and later at IRYStec. It was an incredible experience to be part of a motivated and driven research team. I am thankful to Helge Seetzen who provided me with this opportunity. As part of IRYStec I got the chance to collaborate with Rafal Mantiuk, Gregory Ward, James Clark and all researchers who brainstormed with me during my research.

Most importantly, I would like to thank my parents, my husband and my friends who kept pushing and supporting me through all the ups and downs. This thesis would not have been possible without my parents' loving support, my husband's never ending belief and my friends' encouraging spirit.

Last but not least, I am grateful to Vienna PhD School of Informatics, IMS research group, that funded this PhD research.

Abstract

The High Dynamic Range Imaging (HDRI) pipeline provides us with greater level of detail in color and brightness pixel values by introducing methods from image/video acquisition, storage and compression, to display and quality evaluation. HDRI techniques are considered to offer massive potential for other computer vision and graphics applications to leverage the higher range of information which have not been the main purpose of the HDRI pipeline so far. In this work, we focus on two important computer vision problems and introduce new techniques and algorithms to take HDRI into account to solve them.

Stereo Matching: State of the art stereo matching and 3D reconstruction approaches do not work as well in HDR scenes and low textured regions. The greater luminance information available in HDR content is taken into account to provide more informative disparity maps in tricky matching regions of stereo images. Our combinational approaches close the gap between available HDR methods and stereo matching. We introduce an HDR stereo data set as well as a ground truth HDR stereo matching algorithm as a reference for our tone mapped stereo matching benchmarking. HDR stereo images were captured, calibrated and rectified. A combinational tone mapping approach is proposed for the most effective disparity map computation in HDR scenes. Our combinational tone mapped stereo matching method reduces the stereo matching errors by a factor of three.

Mesopic Rendering: Human visual perception varies in different ambient conditions. Our color and contrast perception especially in darker environments is influenced by the combination of rods and cones photoreceptors. Available standard rendering methods do not take this important perceptual effect into account. HDR images cover the full luminance range of scenes from photopic and mesopic vision to schotopic vision. In the second part of our research we focus more on color and contrast perception in mesopic vision based on HDRI concepts. As another usage of tone mapped images, effective content rendering in dark environments such as Virtual Reality (VR) head mounted displays, night time driving and watching TV in the dark is proposed. Furthermore, subjective evaluations were performed to suggest an effective tone reproduction along with a color retargeting method to compensate for visual detail loss in dark environments on dimmed displays. In the subjective studies our perceptual results were ranked as the most preferred solution in more than 95% of the cases.

Contents

CHAPTER 1

Introduction

This thesis proposes novel solutions and algorithms for using High Dynamic Range (HDR) tone mapped images to generate more accurate results in computer vision applications. HDR capture and display has been studied for the last couple of decades. In this thesis we focus on using tone mapped HDR images to solve computer vision problems, more specifically disparity map computation, 3D reconstruction and optimized solutions for computational displays.

1.1 Motivation

The ultimate goal of computer vision is to use a computer (software and hardware) to understand a scene. The human visual system perceives the three dimensional structure of the world with apparent ease [Sze10]. Constructing a real world 3D scene from two or multiple 2D images is a key step to describe the world that we see. Therefore, we reconstruct the image's properties, such as shape, illumination, and color distributions [Sze10]. The image's properties are massively used in computer vision techniques such as object recognition, optical flow, stereo matching, and object tracking. Extracting the best image properties is substantial. In the first part of this thesis, we use HDR images and different tone mapped images to extract more informative properties such as color, contrast, and illumination for stereo matching.

High dynamic range imaging serves to represent a real world scene which contains a wide range of low to high luminances. The natural world contains a range of radiance values that is far greater than can be captured with any photographic sensor or film [Sze10]. Taking a set of bracketed exposures, one can calibrate the radiometric response function of a camera which is used to create a single image containing the over saturated and under saturated regions of a scene [DM97]. Tone Mapping (TM) operators are designed to compress the contrast of HDR images into a limited dynamic range of displays or printers.

Having the most informative image properties extracted, effectively applying them into different computer vision applications is the next challenge. Today's world is becoming more and more dynamic and we have less control over the environments in which the computer vision applications are being used in. TVs are not being watched in controlled environments but displays and screens are everywhere. On our smart watch, tablet, cell phone and even our car. Lighting has an important effect on the way we perceive images. The second part of this thesis focuses on perceptual color and contrast retargeting methods to keep the image quality in different lighting conditions.

The human visual system works in three different modes called photopic, mesopic and scotopic vision. Photopic vision refers to our vision at day light situations (high light levels) at which only cones are responsible for our vision. As the light level falls on to a luminance of 10 $cd/m2$ [KP04a], the visual system smoothly goes from photopic vision to mesopic vision, in which both cones and rods contribute to visual perception. In the so called scotopic situations, the light level is lower than the absolute threshold of cone photoreceptors and the human vision is only mediated by rods.

1.2 Problem Statement

We introduce and address two sets of problems respectively in the areas of HDR stereo matching and mesopic display rendering.

1.2.1 HDR Stereo Matching

During the last decades, stereo matching has become more and more important for many applications based on 3D reconstruction such as depth image based rendering, 3DTV, and multi-view autostereoscopic displays. The main goal of stereo matching is to reconstruct 3D real world information from a pair of images obtained from two different views. There are many stereo matching algorithms which perform well when there are no under-exposed or over-exposed areas in the scene. However, it is difficult to obtain a high quality disparity map using conventional photography, when a scene contains high dynamic range content. By combining HDR and tone mapping algorithms with stereo techniques, the disparity map of a high dynamic range scene can be obtained. In the first section of this thesis, we address the question in which way the HDR and stereo processing steps should be combined in order to maximize the quality of the disparity maps that can be obtained from a given set of multi-exposed stereo input images. The first two research questions explained below are addressing this problem.

- *HDR Stereo Matching: How to generate and use HDR stereo images for stereo matching?*

  Real world scenes that contain high dynamic range illumination present a special challenge for stereo matching algorithms due to a lack of texture in over- or under-exposed image regions. In Figure 1.1 some of the challenging areas for stereo

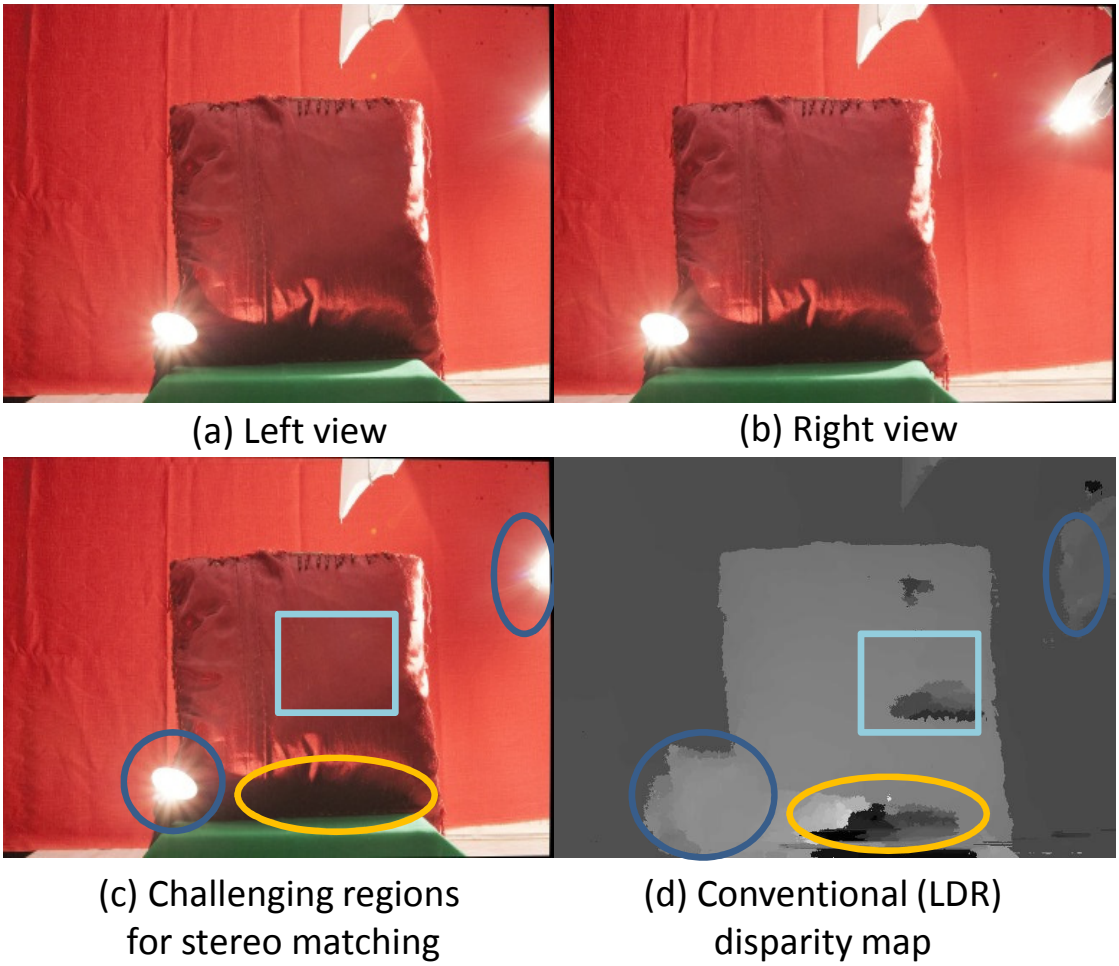


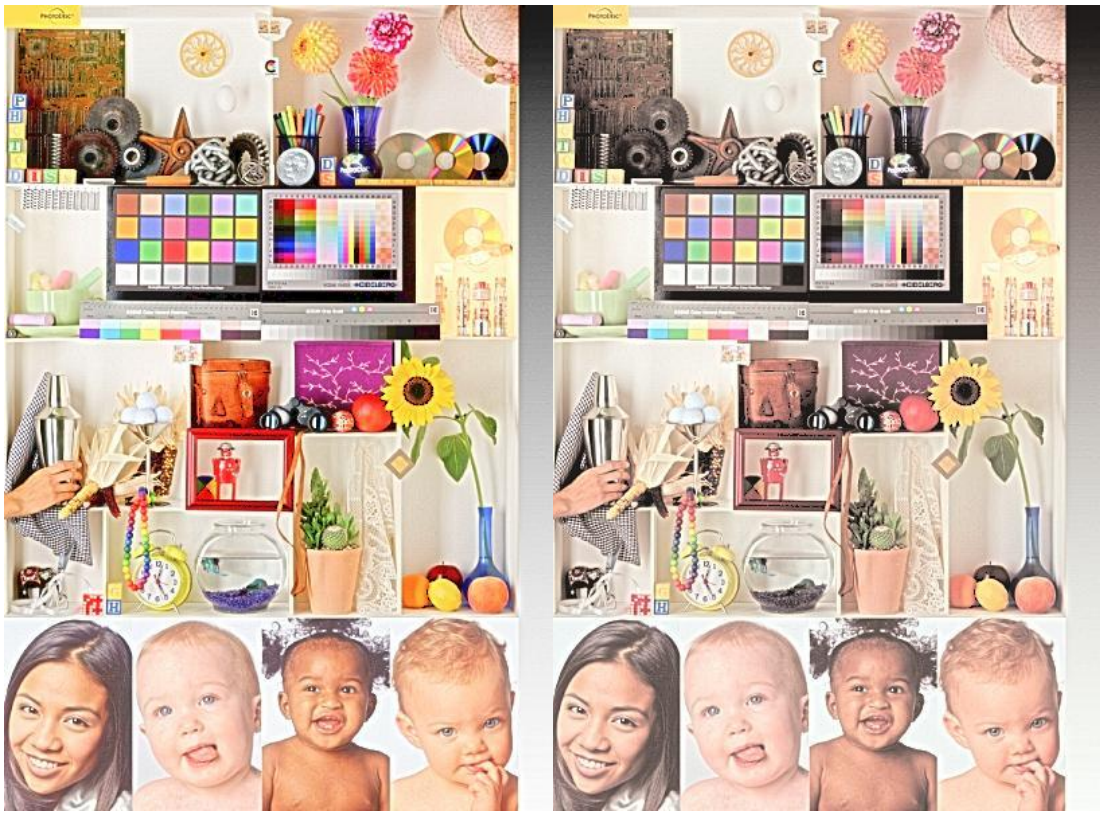
Figure 1.1: Example of an HDR scene containing challenging regions for stereo matching. The left and right views are shown in (a) and (b). In (c), regions specified in dark blue, yellow and light blue, respectively, highlight over-exposed, under-exposed and low-textured areas in the left image. The disparity map (d) is calculated using the cost-volume filtering method [HRB+13]. Baseline: 75 mm.

matching are highlighted on an example from our self-captured stereo data set. Conventional stereo matching methods are unlikely to perform well on an HDR scene as shown on an example in Figure 1.1 (d). In an HDR scene captured with conventional LDR methods like the one presented in Figure 1.1(c), there are over exposed and under exposed regions. Over exposed regions in this image are the areas around the lamp which are specified with dark blue circles. The under exposed area below the pillow is specified using a yellow circle. Another challenging region for traditional stereo matching is the low textured regions. An example of such area is specified using the light blue square in the central part of the pillow. What all these areas have in common is limited texture to be used effectively for the matching process. In addition to that, there is not enough color similarity for traditional stereo matchers to specify the matches only based on the color information.

Available stereo matching methods are basically finding the matching between the right and left pairs shown in Figure 1.1 (a) and (b) by computing the color and texture similarities between the two views. HDR scenes and low textured areas lack enough color similarity and texture to be matched. Figure 1.1 (d) shows the mismatched areas for each of the challenging regions specified in (c). The HDR stereo matching challenges are discussed in more details in Chapter 4. In this thesis, we discuss possibilities for combining state-of-the-art stereo matching algorithms with High Dynamic Range (HDR) imaging techniques, in order to exploit a set of multi-exposed input images of both the left and right stereo view for high-quality stereo reconstruction. We sketch the overall concept of our HDR stereo matching framework and demonstrate some first steps of its implementation, including the acquisition of HDR stereo test data, stereo matching experiments on tone-mapped images, and ideas for combining disparity maps derived from different exposures.

- *Tone Mapped Stereo Matching: Which tone mapping operator suits stereo matching the best?*

  Stereo matching under complex circumstances, such as low textured areas and high dynamic range (HDR) scenes, is an ill-posed problem. In this work, we introduce a stereo matching approach for real world HDR scenes which is backward compatible to conventional stereo matchers. For this purpose, (1) we compare and evaluate tone-mapped disparity maps to find the most suitable tone-mapping approach for the stereo matching purpose. Thereof, (2) we introduce a combining graph-cut based framework for effectively fusing the tone-mapped disparity maps obtained from different tone-mapped input image pairs. And finally, (3) we generate reference ground truth disparity maps for our evaluation using the original HDR images and a customized stereo matching method for HDR inputs. Our experiments show that, combining the most effective features of tone-mapped disparity maps, an improved version of the disparity is achieved. Our results not only reduce the low dynamic range (LDR), conventional disparity errors by the factor of 3, but also outperform the other well-known tone-mapped disparities by providing the closest results to the original HDR disparity maps. More details will be discussed in Chapter 5.



(a) Perceived color and contrast on a standard display with the Luminance of 200 cd/m^2

(b) Perceived color and contrast on a dimmed display with the Luminance of 2 cd/m^2

Figure 1.2: Perception of the same content being displayed on a full brightness display (a) versus a dimmed display in a dark environment (b).

1.2.2 Mesopic Display Rendering

The main goal in display manufacturing is to produce perceptual displays that create natural images for viewers [MWDG13]. To achieve this goal, visual system mechanisms such as contrast, luminance and color perception have to be taken into account in display rendering units [FPSG96a, Dee05]. To have perceptual displays, it is vital to know human color perception mechanisms and be able to model them thoroughly. The model should be comprehensive enough to take into account all aspects of human color vision in all visual conditions such as different light levels [KWK09]. The third research question mentioned below focuses on keeping the perceived quality of an image in different ambient conditions, especially on dimmed displays in a dark environment such as head mounted displays used for VR.

- *Mesopic Display Rendering: How to display images on a dimmed display in a dark*

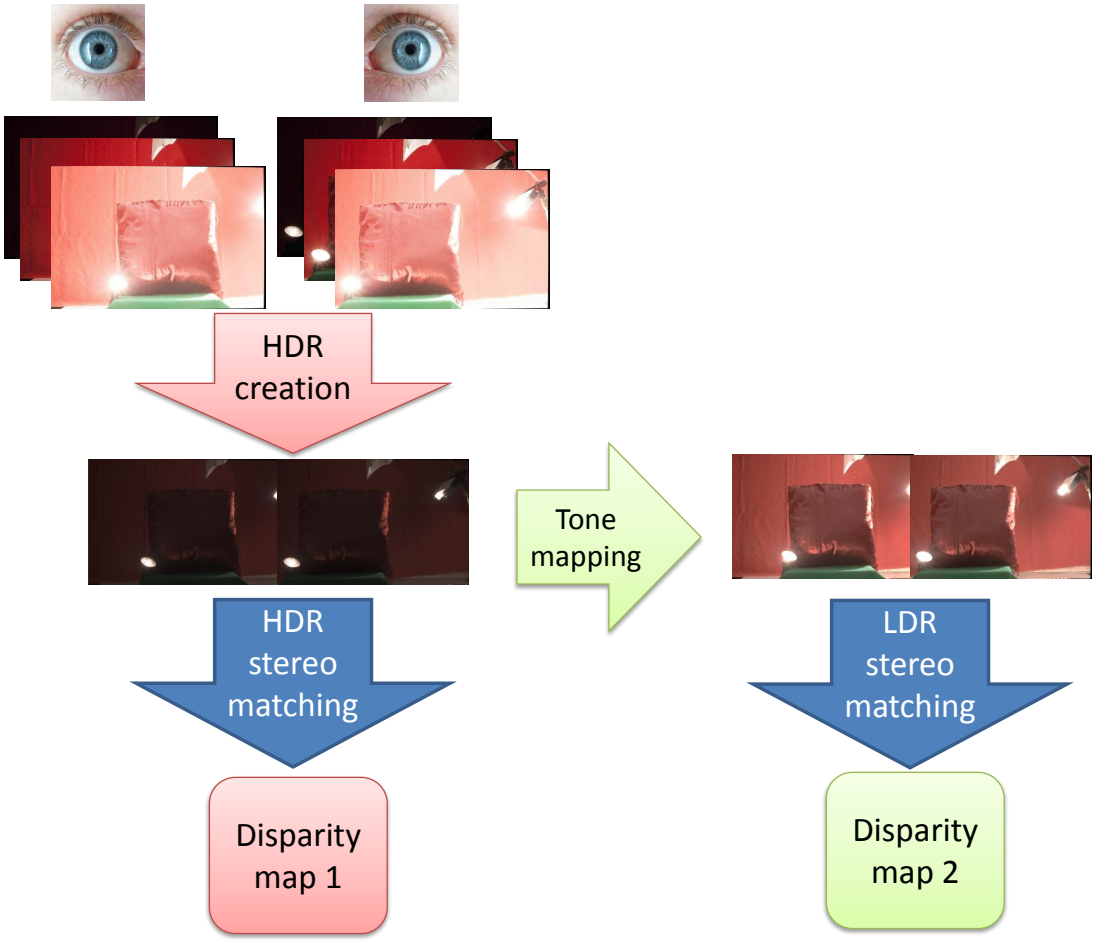


Figure 1.3: Two general approaches for solving the multi-exposed stereo matching problem in HDR scenes.

environment while keeping the perceived quality of the image as it is seen in a room light condition?

Displays are being dimmed automatically using the auto-brightness feature to prevent eye strain for users in a dark environment. Reducing the backlight in LED displays or the current in OLED display panels is a common solution. Figure 1.2 shows the auto-brightness dimmed display for a 2 $cd/m2$ environment on the right. Although adjusting the display brightness is inevitable in the dark, but dimming the display results in losing the original content details as shown on the left side of 1.2. In this work, different tone and color retargeting approaches for visual compensation in mesopic vision are analyzed, compared and evaluated on different content by naive observers in a dark room. We reviewed the tone reproduction approaches for mesopic vision as well as the color retargeting methods and show in the subjective evaluation that a suitable combination of tone and color retargeting

methods is the most preferred result by observers in the dark. More discussion on this topic is provided in 6.

1.3 Contributions

As discussed in Section 1.2, this thesis focuses on two main sets of problems. We will discuss our contributions for solving each of the problems separately in this section.

1.3.1 HDR Stereo Matching

Working in HDR space provides more informative disparity maps. This is especially true in challenging lighting conditions and low-textured areas. However, two important challenges when switching calculations from LDR (Low Dynamic Range) to HDR domain are *lack of data* and *backward compatibility*. To generate a backward compatible solution to conventional stereo matchers, we use tone mapping operations (TMO) to compress the dynamic range into conventional range while preserving details of an HDR image. The schematic solution towards stereo matching on multi-exposed HDR stereo input is illustrated in Figure 1.3.

The main contributions of this PhD thesis towards solving the HDR Stereo Matching problem can be categorized as follows:

- Capturing and preparing a HDR stereo data set of twenty two stereo image pairs in eight different exposures
- Customizing one of the top state of the art stereo matching approaches for HDR input (used for generating ground truth disparity maps)
- Evaluating and comparing 10 tone mapped stereo image pairs
- Introducing a fusion framework for backward compatible tone mapped stereo matching

1.3.2 Mesopic Display Rendering

The main contributions of this PhD thesis towards solving the Mesopic Display Rendering problem can be categorized as follows:

- Introducing a color compensation method for mesopic vision
- Performing a subjective evaluation to compare different color compensation solutions for mesopic vision scenarios
- Performing a subjective evaluation on combination of color and contrast (tone) compensation for mesopic vision and introducing the most effective combination

Scientific Publications and Patents. The results of this PhD research are published in the following publications:

1. Tara Akhavan, Hyunjin Yoo, Margrit Gelautz, A framework for HDR stereo matching using multi-exposed images.*In HDRi2013 - First International Conference and SME Workshop on HDR imaging*, pp. 1–4, 2013.

2. Tara Akhavan, Christian Kappeler, JiHo Cho, Margrit Gelautz, Stereo HDR disparity map computation using structured light, *In HDRi2014 - Second International Conference and SME Workshop on HDR Imaging*, pp. 1-4, 2014.

3. Tara Akhavan, Hyunjin Yoo, Margrit Gelautz,Evaluation of LDR, tone mapped and HDR stereo matching using cost-volume filtering approach,*In 22th European Signal Processing Conference (EUSIPCO 2014)*, pp. 1–6, 2014.

4. Tara akhavan, Hannes Kaufmann, Backward Compatible HDR Stereo Matching: A Hybrid Tone-Mapping Based Framework, *EURASIP Journal on Image and Video Processing*, vol. 36, pp. 1-18, 2015.

5. Mehdi Rezagholizadeh, Tara Akhavan, Afsoon Soudi, Hannes Kaufmann, and James J. Clark, A Retargeting Approach for Mesopic Vision: Simulation and Compensation, *Journal of Imaging Science & Technology*, vol. 60, pp. 1-12, 2015.

6. Tara Akhavan, Mehdi Rezagholizadeh, Afsoon Soudi, James J. Clark, Hannes Kaufmann, and Rafal Mantiuk, Tone Reproduction and Color Retargeting for Mesopic Vision: A Comparison & Subjective Evaluation, *submitted to Journal of the Society for Information Display 2016.*

7. Afsoon Soudi, Mehdi Rezagholizade, Tara Akhavan, Irystec DriveSafe, Ambient Adaptive Software, Makes Driving Safer, *Invited Paper, SID Symposium Digest Of Technical Papers*, Volume 47, Issue 1, 2016.

8. Greg Ward, Hyunjin Yoo, Afsoon Soudi, and Tara Akhavan, Exploiting Wide-Gamut Displays, *Color and Imaging Conference, Society of Imaging, Science and Technology*, 2016.

9. Tara Akhavan, Mehdi Rezagholizadeh, Afsoon Soudi, System and Method for Color Retargeting,*PCT patent, available publicly, WO/2016/183681, International Application No.:PCT/CA2016/050565*, Publication Date:24.11.2016, International Filing Date:19.05.2016.

10. Greg Ward, Hyunjin Yoo, Afsoon Soudi, and Tara Akhavan, System and Method for Age-based Gamut Mapping, Provisional patent, *Application No.: 62/418.361*, Filing Date:7.10.2016

1.4 Thesis Outline

The thesis continues as follows. Theoretical background and related work are discussed in Chapter 2. Chapters 3, 4 and 5 are focused on our first problem statement and provide details on required aspects of our proposed solutions. Chapter 3 explains our HDR stereo image capturing and lessons learned. We then continue with discussing HDR stereo matching and customization to provide ground truth. A Combinational graph cut based approach proposed for combining different disparity maps is explained in Chapter 5. In this Chapter different disparity maps are generated from different tone mapping operators applied on the HDR stereo pair. Our second problem statement and proposed solutions on mesopic rendering is discussed in detail in Chapter 6. In each Chapter the results and evaluation are provided as well. Finally Chapter 7 concludes the thesis and talks about the future work.

CHAPTER 2

Theoretical Background and Related Work

In this Chapter, we discuss three different fields of research in computer vision and graphics: HDRI, stereo matching and mesopic rendering. The discussion is categorized to follow the dissertation's flow of topics. In a top down approach, we start from the computer vision goal and available toolboxes continued by an overview of HDRI techniques in image acquisition and display. In order to elaborate on our contributions in leveraging HDRI techniques for stereo matching and mesopic rendering, two Sections are dedicated to discuss state-of-the-art research in each of the two. The review is followed by an analysis of few combinational solutions available for closing the gap between HDRI and stereo matching/mesopic rendering. This Chapter will provide the necessary background and benchmark analysis for the rest of the dissertation.

2.1 Real-world Vision: Brightness, Contrast, 3D and Color

The goal of Computer Vision is to simulate human vision. Image processing tools are aimed at capturing, transmitting and presenting content as closely as possible to what we see in real world. Real world vision contains an enormous amount of information in categories such as brightness levels, contrast range, 3D and millions of shades of colors. Each of the mentioned categories are tackled in the field separately.

Our main goal in this work is to close the gaps and bring the separated research fields closer to each other and leverage one on top of the other. Our second main goal is to get some distance from pure image processing which focuses on precision improvement and image enhancement. We intent to do so by making sure that solutions are based on human perception and suit the visual system the most. We make this happen by

leveraging subjective studies and come up with more perceptual solutions which are based on direct perceptual measurements and not only image enhancement metrics.

Paying specific attention to perceptual models of the human visual system can massively help with simplifying algorithms and solutions as well as acheiving more natural results. Perceived brightness and dynamic range [War01], color appearance models [CPSZ08a, SKCP11, Fai13a], contrast sensitivities in different brightness conditions [Kul76], Hunt effect [Hun94a], Purkinje shift [Ans02] and many other physiological research provide us with a great base to work on perceptual aspects of algorithms. Each of the mentioned perceptual models and the way we used them in our algorithms are explained in detail throughout this thesis.

In this Chapter, we discuss several separate categories of computer vision and graphics problems such as HDR and Tone Mapping, 3D, Mesopic vision and Color Appearance Models and provide theoretical and perceptual background on each. We also discuss some overlap among the mentioned fields and start the discussion around leveraging their combination in improving existing solutions.

2.2 HDR Imaging

High Dynamic Range Imaging (HDRI) provides pixels which contain a greater range of color and brightness levels compared to the existing standard dynamic range images and videos. This improvement will result in images to look more natural and realistic to human eyes since in real world we see a huge dynamic range of details. The visible range of colors are much larger than the ones achievable by cameras or displays. HDRI provides a new set of techniques and algorithms to capture and represent pixel values. HDRI methods manipulate and store all colors and brightness levels that we can perceive with our naked eyes in contrast to limiting the techniques to the range of colors on a specific device [MS15].

2.2.1 HDR Image Acquisition and Display

Capturing and displaying HDR images has been widely explored during the last two decades [DM97, DMAC03, SHS+04, SDBRC13, WT14]. However, HDR images and techniques are not broadly used in image processing and computer vision applications such as stereo matching and 3D reconstruction, segmentation, alpha matting and face recognition.

The most obvious goal of the HDRI pipeline is to provide and standardize the methods from image/video acquisition, then storage and compression, to display and quality evaluation [MS15]. Figure 2.1 shows an overview of the HDRI pipeline and available techniques [MS15]. To create an HDR image using a digital camera, a camera response curve should be reconstructed from multiple varying exposed images. To show the HDR image on a conventional display with low dynamic range, a tone-mapping process is needed.

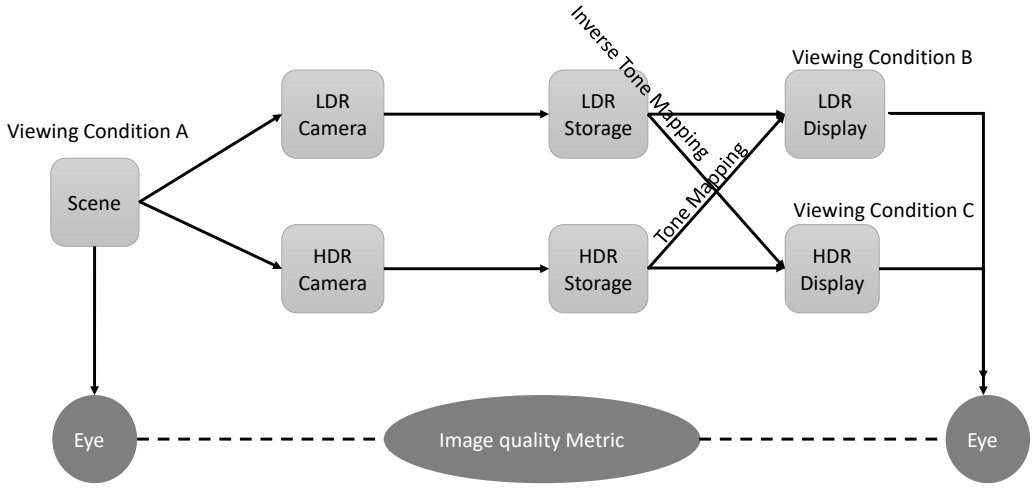


Figure 2.1: High Dynamic Range Imaging Pipeline. [MS15].

In this work, for stereo matching of a scene containing HDR illumination, we used not only multi-view images, but also multi-exposure images. We follow HDRI acquisition methods to do so [DM97]. Our dataset contains eight or nine differently-exposed images for each view. More details about the dataset is discussed in Chapter 3.

2.2.2 HDR Imaging Applications

We believe there is a huge potential in leveraging the larger amount of color and brightness details captured in HDR content to help with image processing and computer vision applications. In such applications, one might not necessarily need to display the HDR content but to replace the standard low dynamic range data in the pipeline with HDR or even tone mapped images. This has not been the essential purpose of HDRI.

In the rest of this work, we talk in more detail about how to leverage HDR content for disparity map calculation as well as mesopic vision compensation on displays. Even though we do not always use the HDR images as is, but based on the application, we leverage HDRI methods and techniques. Figure 2.2 illustrates our proposed high level architecture of HDRI applications, more specifically in stereo matching and mesopic rendering.

On one hand the advantages of high dynamic range imaging are widely being considered and recognized in many related fields to photography (i.e. computer graphics, computer vision and high quality photography); and on the other hand during the last decades stereo matching has become an important requirement for so many applications such as 3D reconstruction. In Chapter 4 and Chapter 5 of this thesis we try to get a more

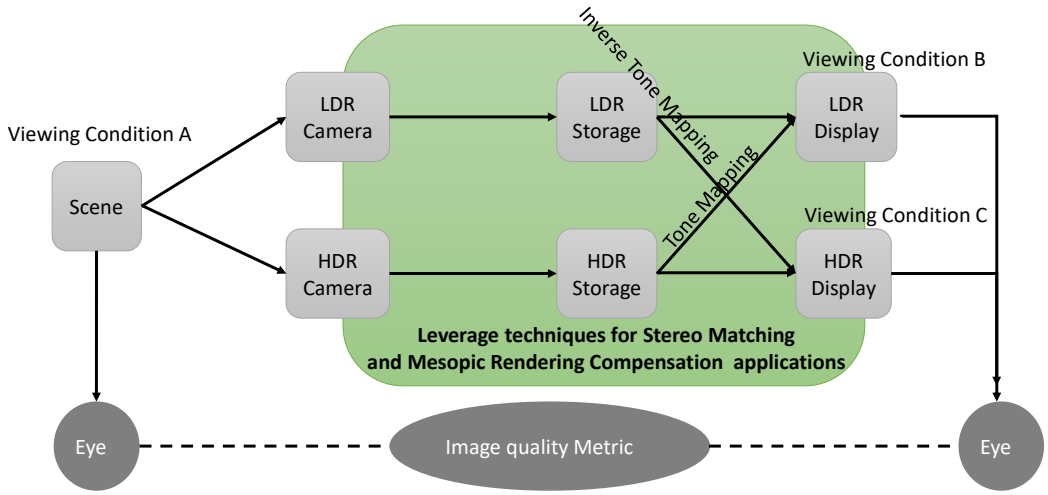


Figure 2.2: Leveraging HDRI techniques in other computer vision applications.

accurate disparity map for stereo matching with the help of high dynamic range images and relative concepts.

The biggest issues in the stereo matching field of research are occlusions, fine structures, boundaries as well as illumination and exposure differences. Hence we came up with the idea of using multi-exposed images for each view of the stereo matching methods. Our dataset contains eight or nine different images for the right view with different exposures (under-exposed, normal-exposed, over-exposed) and eight or nine images for the left view with the exact same exposures as the right view images.

Working in HDR space can lead to better results by using more detailed brightness information. It is not hard to predict that working in radiance (HDR) space provides more informative disparities. This is especially true in challenging lighting conditions and low-textured areas. However, two important challenges when switching calculations from LDR to HDR domain are *lack of data* and *backward compatibility*. Acquisition of stereo HDR content is not as challenging as leveraging that content by conventional stereo matching algorithms.

The state of the art stereo matchers work with conventional eight bit data that result in lack of depth information in low-textured and challenging lighting areas. Changing the input to HDR will require appropriate customization in the stereo matching method [AYG14]. Our study provides a framework to use any conventional stereo matching method to achieve more discriminative disparity maps in HDR scenes. We use tone-mapped HDR stereo pairs with no change in the stereo matcher (backward compatible) which reduces the disparity map Root Mean Square Error (RMSE) of the conventional

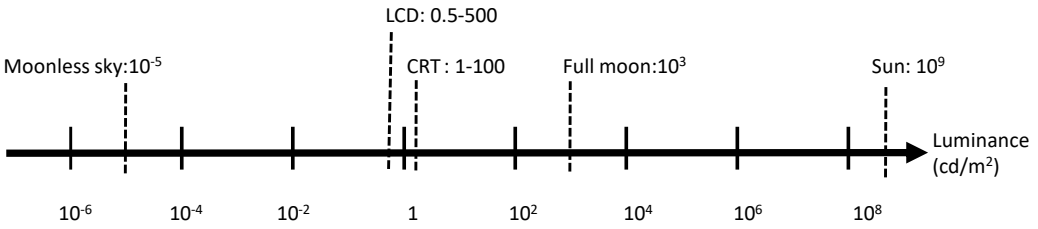


Figure 2.3: Visible luminance levels to human eye. [MS15].

approaches by the factor of three.

As mentioned before, HDRI methods are focused on capturing, storing and representing a huge range of luminance levels. There is a massive amount of literature on each of those aspects of HDRI. However our main discussion is around the best ways to leverage the existing HDRI techniques and HDR content in other computer vision problems. In this thesis we will introduce two combinational fields which will directly and indirectly leverage HDRI methods in different ways:

1. HDR and stereo matching: How can we leverage HDRI to generate more accurate disparity maps?

   It has been proven in recent years that taking into account depth information helps with capturing more accurate HDR images [SMW10, BLV+12, TKS06]. However, only limited research has been done on the contribution of HDRI techniques in calculating more informative depth maps. HDR stereo images or tone mapped stereo images with more available detail for matching, could potentially enter the standard stereo matching pipeline and help us calculating more accurate disparity maps with less mismatched pixels. We will discuss the state of the art in detail in Section 2.3. We perceive HDR content in 3D in the real world, so the next step for computer vision solutions is also to get closer to the combination of the two.

2. HDR and mesopic vision: How can we leverage HDRI to compensate for information loss on dimmed displays in mesopic vision?

   Not only the visible color gamut depends on the luminance level, but also our visual system works differently in different luminance conditions. Visible luminance levels in different situations are shown in Figure 2.3. The human eye is capable of perceiving a huge range of brightness but only a small part of the range is covered with current cameras and display panels. HDR displays [SHS+04] provide a greater range closer to the range perceived in real world. However, the fact that

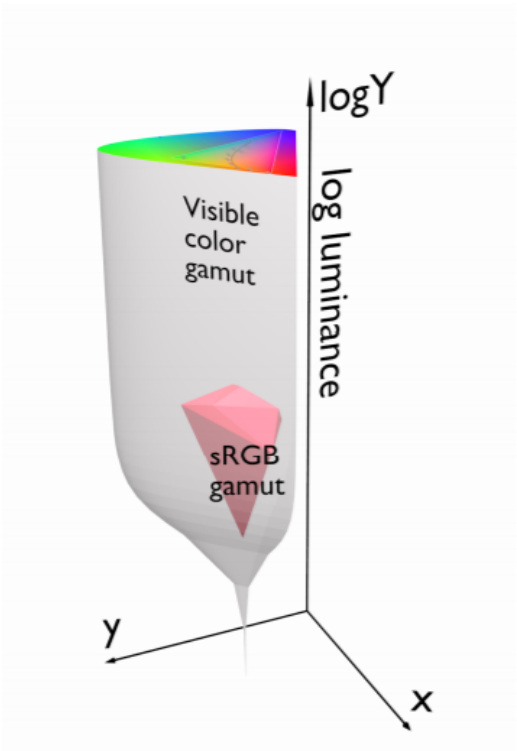


Figure 2.4: Visible color gamut to human eye compared with sRGB gamut. Please note that the color perception degrades in lower luminance levels. [MS15].

our visual color perception also changes based on the luminance level has not been taken into account [CPSZ08a]. As shown in Figure2.4, in lower luminance levels, known as mesopic and schotopic vision, visual perception differs. The fact that rod photoreceptors contribute to our vision in mesopic, as well as the change in our contrast sensitivity result in the need to compensate the content and displays to match our vision in mesopic. HDRI knowledge, tone reproduction methods and color retargeting approaches are combined to achieve this goal in Chapter 6.

High dynamic range imaging introduces great pixel colorimetric precision, which enables representing all colors found in the real world that can be perceived by the human eye. HDRI concepts can be adapted to a huge range of perceptual cues that are not achievable with traditional imaging to mimic the human visual system. With HDR images we can cover the full luminance range of scenes from photopic and mesopic vision to schotopic vision [MS15].

In the second part of our research we focus more on color and contrast perception in mesopic vision based on HDRI concepts. Due to the Hunt's effect [Hun91] we tend to regard objects more colorful when they are brightly illuminated. In Chapter 6 we

provide solutions to render enhanced colorfulness properly based on the information preservation about the actual level of luminance of the original scene, which is not possible in the case of traditional imaging.

Capturing and displaying are correspondingly the first and last steps of the imaging pipeline. In Chapter 6, we focus on display rendering in mesopic vision, aiming to optimize the user's display visual experience on head mounted displays used in VR glasses, smart phone screens and car displays for night time driving. It is very common for the smart devices (e.g. mobile phones, tablets, car displays, laptops) to automatically match the screen brightness with the environment in the dark to reduce eye fatigue [eye]. This feature is called auto brightness. Although, dimming the display reduces the user's eye fatigue in the dark, it also reduces the perceived color and contrast by dropping the luminance level of the screen. The effect of color retargeting approaches and Color Appearance Models (CAM) on compensating the color perception in mesopic (on a dimmed display) has been recently reviewed in [RAS+15]. Most of the color retargeting solutions are based on the physiological studies [KJF07a], [SMYS04a], [Fai13b], [McC07], [Hun95]. A good number of tone reproduction operators, also resemble parts of human neurophysiology, and in particular the behavior of photoreceptors [Rei11], [KO11], [WHM+05], [MSTK09]. Therefore, we believe, color retargeting and tone reproduction may seem to be very different fields of research and communities, but their aims overlap and one without the other will not deliver the complete solution.

Both of the mentioned combined areas have been touched in the literature before. However, after discussing what is available in the state of the art in this Chapter, we will take the discussions and experiments one level further in Chapters 4, 5, and 6. The rest of this Chapter is focused on first, HDRI and it's available methods, stereo matching and 3D reconstruction and a discussion around HDRI combined with stereo matching. Finally mesopic vision compensation will be discussed followed by HDRI's contribution to the field.

2.3 Stereo Matching and 3D Reconstruction

Stereo matching, that is the identification of corresponding points between the left and right image of a stereo pair, has been an active area of computer vision research for many years. There are some well-known challenges in stereo matching such as low-textured regions, repeating patterns, saturated areas and illumination changes between the two stereo views.

The importance of stereo-matching methods in computer vision and image processing during the last decade is evident. There are hundreds of stereo-matching approaches available to estimate the depth of a scene from left and right images. Lots of effort has been made on the algorithms but not so much attention was paid to the scene characteristics. On scenes with Higher Dynamic Range than the controlled indoor scenes

(Low Dynamic Range or LDR), the conventional stereo matching methods are unlikely to perform well. Having 256 values per color channel is inadequate for representing many scenes [RWPD05].

Capturing a huge range of illumination in HDR scenes using only 256 values per color channel - as in conventional LDR images - typically causes over- and/or under-exposed regions in the image. Many of the available state-of-the-art stereo matching methods perform well in absence of the mentioned regions. However, it is not possible to avoid these challenging areas in real world imaging.

Szeliski, Zabih and Scharstein have introduced a complete benchmark to categorize and compare different stereo matching algorithms performances [SS02b, SS03b, SS02a, SS03a, Sch07, HS07]. The well-known Middlebury data set is a result of their great research in the field. The main focus in stereo matching is on two frame methods but it is easy to generalize them to multi frames for multi view matching.

It is very common to base computer vision approaches on few assumptions. There are three main assumption in stereo matching [SS02b]:

1. The first common assumption in stereo matching is that all scene surfaces are Lambertian surfaces, surfaces whose appearance does not vary with viewpoint.
2. The second assumption is the smoothness assumption assuming that the physical world consists of piecewise-smooth surfaces. Without the smoothness assumption the correspondence problem would be an ill-pose problem.
3. The third assumption is regarding the camera calibration and epipolar geometry. Based on this assumption, we should make sure that the stereo input images are rectified.

Most stereo matching algorithms compute a disparity function $d(x, y)$ with respect to a reference image, which could be one of the input images [SS02b]. Throughout this work we calculate the disparity with respect to the left image. In computer vision, disparity is often known as an inverse depth [BBM87]. In this thesis, all of our images are taken on a linear path with the optical axis perpendicular to the camera displacement, and the classical inverse-depth interpretation will suffice. The (x, y) coordinates of the disparity space are taken to be coincident with the pixel coordinates of our reference image, the left image, from our input data set. The correspondence between a pixel (x, y) in reference image r and a pixel (x', y') in matching image m is then given by Equation 2.1.

$$x' = x + (r - m)d(x, y), y' = y \tag{2.1}$$

From a disparity space, one can calculate a disparity space image or DSI presented as (x, y, d). The DSI usually represents the confidence or log likelihood (i.e., cost) of a particular match implied by $d(x, y)$. The goal of a stereo matching method is to produce a

univalued function in disparity space $d(x, y)$ that best describes the shape of the surfaces in the scene [SS02b].

In general, most of the stereo matching algorithms are based on four main steps [SS98].

1. Matching cost computation
2. Cost aggregation
3. Disparity computation / optimization
4. Disparity refinement

In recent years, many innovative ways of defining and computing cost functions have been proposed. One of the well-performing ones is the cost volume filtering introduced by Hosni et al. [HRB+13]. This approach is following the four main steps mentioned above. The main computational steps of the cost volume filtering can be summarized below:

1. Constructing the cost-volume
2. Filtering the cost volume using guided filter [HST10]
3. Label selection (winner takes all, optimization problem)
4. Occlusion detection and filling
5. Post-processing

In this work, we modify the fast cost-volume filtering method introduced in [HRB+13] to work in HDR space as well as with different tone mapped stereo images. Hosni et al. follow a discrete label-based approach by calculating costs for choosing a label, which corresponds to a disparity. These costs are determined using pixel correlations in the left and right image by calculating absolute differences in the intensity space. Throughout the rest of this work, we will use Hosni's stereo matching as our benchmark to compare and evaluate different content characteristics.

2.3.1 Combining Multiple Depth Maps

Combining several depth maps achieved from different view points is studied by Schuon et al. [STDT09]. Combining range information to generate a more accurate result is a well-known approach in the 3D society [DT05]. Another successful example was introduced by Izadi et al. [IKH+11] to fuse a sequence of depth maps generated by a kinect camera. We use information from several tone-mapped disparity maps and combine them to provide backward compatible stereo matching results for HDR scenes.

Many different methods can be used for combining results. Some of the simple ways are to calculate the average, weighted average or median of the candidate results. In

most applications simple combination methods, without taking into account any prior or statistical knowledge of candidate results, do not achieve the best outcome. Akhavan et al. used a machine learning based approach to combine results of color constancy in [AM10]. The Markov Random Field (MRF) based method was described and used by [DT05] to combine range information in 2005. In [AM11], a fuzzy integral method was introduced as a combination method which considers the dependencies between the candidate results. In Chapter 5, we use Markov Random Field and more specifically graph-cuts to combine the disparity maps. Using MRF has proved to be effective in labeling problems such as stereo matching [Li09], since the output values of the matching algorithms are *cost* values that are suitable targets for energy minimization algorithms.

2.3.2 HDR Stereo Matching

Even though there is a considerable amount of literature on the state of the art in each of the HDR and stereo matching fields, not much work has been done on joining them. Sun et al. [SMW10] constructed an HDR image by using two stereo images with two different exposures. They first calculated an initial disparity map to obtain the camera response function. After refining the disparity map, an HDR image is constructed. Since they only used two images in the stereo matching process, if one of them is significantly over-exposed or under-exposed relative to the other, it may be difficult to calculate a reliable disparity map. Troccoli et al. [TKS06] also adopted multi-view stereo with different exposures to obtain both the depth map and HDR image.

Rufenacht [Rüf11] created stereoscopic video content using different exposure times for the right and left cameras. He captured video frames, estimated the camera response function, aligned the image and estimated the HDR radiance map. In contrast to what we aim to achieve, his main focus was on the quality of HDR radiance maps and not the quality of disparity maps. Bonnard et al. [BLV+12] leveraged an eight view video camera to acquire HDR video content for autostereoscopic displays. They did not use any stereo matching, but built their pixel match list with the help of the camera properties with eight aligned, equally distributed objectives. Their main focus was on creating HDR content as well.

Recently, Selmanovic et al. [SDBRC12] introduced a number of compression methods for Stereoscopic High Dynamic Range (SHDR). An HDR image requires floating point value storage per color channel per pixel, so SHDR needs twice of the storage. None of the aforementioned papers provides a systematic overview and evaluation of different strategies for obtaining disparity maps from multiexposed stereo input data, which is the goal of our work in Chapter 4.

2.3.3 Tone Mapping Operators and Tone-mapped Stereo Matching

Several approaches have been presented for constructing an HDR image from two differently exposed LDR stereo images by calculating the depth information of the scene [TKS06, LC09, SMW10, Rüf11, LJDE11, RLMA13]. More recently, Batz et al.

[BRG+14] and Orozco et al. [ROLMA15] proposed interesting approaches for HDR video reconstruction using depth information. Orozco et al. introduced a patch match based method to generate 3D HDR video sequences using available hardware. The main goal in the mentioned articles is to generate better quality HDR image/video while our main focus is to use the available HDR content to achieve more informative disparity maps.

Few approaches have been introduced for subjectively comparing tone-mapped stereo images with the focus on stereoscopic data generation [MDNW12] or disparity map calculation [AYG14]. Recently Aydin et al. [ASC+14] evaluated some of the TMOs based on edge-aware filters for HDR video tone-mapping taking into account the visual artifacts and temporal coherency in the tone-mapped video. The authors also introduced a faster and more efficient filter for high motion scenes. The key contribution of this filter is to achieve temporal stability without ghosting on high motion videos. Our work focuses on still stereo image pairs and does not contain motion.

Although, our first objective was to report *the best* TMO regarding stereo matching, the outcome of our experiments is consistent with the result of many evaluations that have been done on TMOs [YBMS05, LCTS05, ČWNA06]; there is no single TMO which performs the best in all conditions. Therefore, we combined the disparity maps from the tone-mapped images to maximize the quality of the disparity map taking into account the strong points of each TMO. Our combinational approach is discussed in details in Chapter 5.

Combining Multiple TMOs

Combining results achieved from different methods is a common approach in image processing and computer vision research [PDKK03, AM11]. More specifically, fusing several TMOs to achieve better quality images is addressed in [YYBR14]. The idea behind this fusion is that each TMO works better in a special image region and under some specific conditions the best output can be calculated by taking into account a suitable TMO for each image region. Mai, et.al, [MDNW12] explored that HDR tone-mapping can significantly enhance perceptual quality of 3D images. We go one step further and use tone-mapped stereo pairs to obtain better disparity information.

In Chapter 5 we propose a new effective approach to combine different disparity maps calculated with different tone mapped stereo images of an HDR scene. Our combinational results show great improvements in reducing mismatched points for stereo matching.

2.4 Ambient Intelligent Rendering: Mesopic Vision Compensation

In recent years, mesopic vision and perception in the dark is getting a greater traction from both academia and industry [App, EVH+05, Fai13b]. One of the most important reasons for this traction is the advances in Virtual Reality (VR) [Occ] as well as the growth in using smart devices before bed time and the effect of late stimulation on

vision and circadian rhythms [con]. The automiotive industry and night time driving is another important industrial problem which is driving many research in the mesopic field [EVH+05].

Our visual color and contrast perception changes in the dark. In the mesopic range, rods and cones both contribute to color vision [PC10]. As a result, some color retargeting or reproduction is needed to match the current standards -which are mostly based on cones response [MFH+02]- to take into account the mesopic vision characteristics [RAS+15]. Contrast retargeting or tone reproduction to compensate for the contrast in the dark environment has been also widely investigated [WM14b], [KML03] in the last decade.

Our perceived color and contrast varies significantly depending on the brightness level of the ambient environment. Based on our cones and rods photoreceptors' contribution to our vision, scientists have defined three main visual modes called photopic, mesopic and scotopic. In daylight situation or photopic vision our cones are responsible for our visual response. As the luminance level falls down to 10 cd/m^2 [KP04b] and less, our visual system moves from photopic to mesopic vision in which both cones and rods are active. In less than 0.1 cd/m^2, when only rods are active and cones cannot distinguish much color, our visual system moves toward scotopic vision. Photopic vision has attracted the most focus of researchers since most of the camera and display applications in the real world are in a photopic and daylight environment. Mesopic and scotopic vision, are correspondinlgy attracting more attention in the industry [EVH+05], [Fai13b], [CHVW10].

In this work, we review some of the most highly-ranked approaches for both color retargeting and contrast reproduction with focus on mesopic vision as well as conducting a subjective evaluation on the effectiveness of each of the color and contrast retargeting methods as well as the combination of both color and contrast. Our result is very consistent with some of the prior art in this field which insists on the importance of taking advantage of both color and contrast corrections to provide the end user with an optimized visual experience [WM14b], [Rei11].

Tone or contrast retargeting is required for mesopic rendering as well as HDR rendering on LDR displays. Displaying HDR content on available LDR displays lead to the tone mapping research field. In this work, we leverage the tone mapping state of the art techniques, even though our goal is to reproduce and tune LDR and HDR images on dimmed displays for mesopic vision.

HDRI techniques include a massive insight into different tone reproduction approaches. Color correction is also required after any contrast retargeting. Figure 2.5 illustrates the proposed flowchart to take both contrast and color retargeting into account for rendering purposes. We follow this model and discuss both contrast retargeting and color retargeting methods and the best combination of the two with focus on mesopic rendering. It must be mentioned that this generic flowchart is valid for all three photopic, mesopic and schotopic vision conditions.

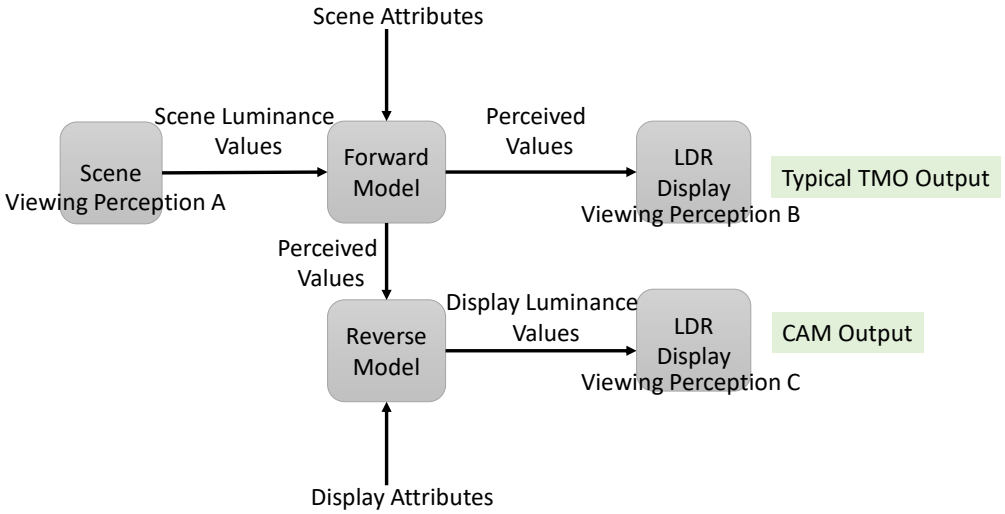


Figure 2.5: Outlined flowchart for tone and color retargeting. The input scene is passed through the forward model using scene reffered parameters. Most of tone retargeting approaches directly display the output. Color Appearance Models (CAM) use display reffered parameters in reverse to correct the tone reproduction and render displayable luminance values. The flowchart is based on [Rei11] and [WM14c].

2.4.1 Contrast Retargeting Approaches

In this section, we focus on contrast reproduction (sometimes known as tone reproduction) operators which are appropriate for mesopic range. Pattanaik et al. developed a model of adaptation and spatial vision based on a multiscale representation of the human visual system, color processing, as well as luminance [PFFG98]. This model accounts for a wide range of changes, such as visual acuity, colorfulness, and apprarent contrast, which varies with illumination.

Ferwerda et al. proposed a model for visual adaptation in which different human visual system phenomena such as threshold visibility, visual acuity, temporal adaptation, and color correction are involved [FPSG96b]. Durand and Dorsey extended the Ferwerda tone mapping operator by adding a blue shift mechanism to address the mesopic color appearance of night scenes [DD00]. This blue shift operator is built on the Hunt data showing that white preference changes in very dark conditions toward the normalized $RGB = [1.05, 0.97, 1.27]$ in Rec. 709 color space. However, this model over simplifies the complex mesopic vision mechanisms. Mikamo et al. decoupled the luminance component from the chromatic content of the image and then discount the red content of the image in the Lab color space depending on the average luminance level of the image [MSTK09].

Krawczyk et al. introduced a local contrast compression technique in which they included some perceptual phenomena related to mesopic vision such as change of visual acuity and rod contributions to mesopic vision [KMS05].

Recently Blakenback et al. proposed a tone mapping algorithm for automotive Human Machine Interface images where only the bright light condition is taken into account [BSKZ14]. It is very common to apply a simple color correction during the tone reproduction process, however, using a perceptual color retargeting method is proven to be very effective in providing a complete simulation of our night vision [Rei11]. A thourough review of tone reproduction operators can be found in [Dev02].

As mentioned above, even though there are several available approaches for contrast reproduction but some of them are not proposed with a focus on mesopic vision and some others are not complete in a sense they do not work perfectly with color retargeting. Our subjective and objective evaluations show that contrast and color retargeting methods for mesopic should come hand in hand to be most effective and generate natural viewing experience in the dark. In Chapter 6 we perform detailed evaluation and ranking techniques to choose the contrast retargeting method which is the most perceptually close to human mesopic vision.

2.4.2 Color Retargeting Approaches

There is an abundant number of color appearance models and/or tone mapping algorithms that could be applied to photopic conditions (i.e., high light levels). There are some well-known color models which intend to cover wide luminance levels, but often show limited performance for mesopic and scotopic visions. For instance, CIECAM02 [MFH$^+$02], one of the most widely applied color appearance models, is shown to work the best for the photopic vision [WM14b]. However, the outcome of an ideal color appearance model should resemble human color perception in all conditions such as different adaptation, light levels, and viewing conditions. The HDRI pipeline and it's perceptual insights were taken into account later on by Kuang et al in [KJF07b]. They refined the image appearance model to work for HDR rendering.

One of the most extensive and complex models was developed by Hunt in 1991, which predicts the appearance of both unrelated and related colors at a wide range of luminance from scotopic to bleaching levels and other viewing conditions [Hun91]. An improvement of colorfulness predictor was made to the Hunt 91 model later on to generate more consistent predictions under various viewing conditions [Hun94b]. Kwak et al. [KML03] developed a lightness predictor which included rods' contribution to the achromatic signal for mesopic conditions.

Recent findings show that rod cells contribute to off-bipolar cells during the scotopic condition by forming chemical synapses [RR98]. Based on this theory, Khan and Pattanaik [KP04b] hypothesized that these synapses were only established between the rods and cones to explain the *blue shift* at scotopic light levels. However, their proposed algorithm did not produce natural output. Clark and Skaff [CS09] proposed a *spectral theory of*

color perception, which was later extended for mesopic vision by Rezagholizadeh and Clark [RC13]. However, this model suffers from high computation load, which makes it hard to use in real-time applications.

Cao et al [CPSZ08b] investigated the rods' contribution to color perception as a function of rods' contrast at mesopic vision. Their results showed a linear relationship between rods' contrast and rods' contributions to the parvocellular (PC), koniocellular (KC), and magnocellular (MC) pathways. However, other studies show that this model is inadequate to reproduce colors for both color patches [RC13] and complex stimuli [WM14b] at low light levels. Later on, Wanat and Mantiuk [WM14b] proposed a color retargeting method in Mesopic based on Cao's physiological studies.

Shin et al. proposed a modified version of the Boynton two-stage model with fitting parameters to account for the rod intrusion in mesopic vision [SMYS04a]. The goal of the model is to find the matching colors in photopic range for the input colors in mesopic range. The parameters of the model are obtained as a function of luminance based on the asymmetric color matching experimental data. In their experiment, the observer is presented with a Munsell color chip under the mesopic condition and is asked to match the appearance of that patch with the simulated image reproduced by this model in the CRT display under photopic condition. In Chapter 6 based on the Shin Mesopic model, we derived the inverse-Shin model and developed a color retargeting method for rendering images at Mesopic light levels compensating for color appearance changes of images viewed on dimmed displays in dark environments.

CHAPTER 3

HDR Stereo Data Acquisition

One of the most important challenges in using HDR images for computer vision applications is lack of data. To persue our goal of using extra information in HDR images for generating better disparity maps, we had to generate a stereo data set of HDR scenes containing the challenging regions for stereo matching. Unfortunately, the well-known Middlebury stereo data set [Sch07, HS07] does not contain HDR scenes with a great spectrum of luminance, even though it contains multi-exposure views. Most of the scenes captured in Middlebury could be categorized as normal LDR indoor scenes of 10^2 cd/m^2. The multi-exposure images do not cover a sufficient range of brightness to be considered HDR data, as their exposure times are relatively close to each other. Therefore, we generated our data set of HDR scenes including highly exposed regions and low-textured areas for our experimental process.

3.1 HDR Stereo Capturing

A comparison of different lighting conditions is provided by [Wan95]. The luminance level of different lighting environments vary massively. As it can be seen in Table 3.1, indoor controlled environments which we consider as LDR environments only contain scenes with the luminance level of up to 10^2 cd/m^2. In contrast, an outdoor scene in a sunny day can contain a huge range of luminance levels up to 10^5 cd/m^2. Ideally we should have captured sunlit HDR scenes in different exposures for our left and right stereo pairs. However, because of our setup limitations and the mechanical controls which move our camera for the stereo capturing purpose we were limited to our laboratory settings.

To capture a higher range of luminance levels we used various lighting sources in our setup to generate over exposed as well as under exposed regions in each scene. As shown in Table 3.1, the average brightness of our scenes are 10^3 cd/m^2, even though HDR scenes can contain higher luminance levels. As we will discuss in the future chapters, luminance level 10^3 cd/m^2 adds a huge challenge to current stereo matching approaches to deal

Condition	**Illumination** (in cd/m^2)
Starlight	10^{-3}
Moonlight	10^{-1}
Indoor lighting (controlled)	10^{2}
Sunlight	10^{5}
Our data set lighting	10^{3}
Max intensity of common monitors	10^{2}

Table 3.1: Our data set luminance level in comparison to some common lighting environments [Wan95].

Figure 3.1: Scene capturing process. Top left: experimental setup, top right: HDR image capturing.

with. Our proposed methods all through the thesis can be applied to higher dynamic range images.

In order to create HDR stereo images using the bracketed technique proposed in [DM97], we capture normal (LDR) images in eight different exposures for each view. Our scenes are high dynamic range scenes of 10^3 cd/m^2. Fig. 3.2 shows all the diferently exposed LDR images for both left and right views for a scene containing low-textured regions captured under a non-uniform luminance in the range of 1 to 8000 cd/m^2. Our data set comprises two different baselines of 75 mm and 150 mm between the two stereo views. For both baselines the images were separately rectified. The HDR images were directly generated from raw image files of the eight LDR exposures following the approach of Debevec et al. [DM97]. A sample from our stereo LDR-HDR data set is shown in Fig. 3.2 which represents one complete LDR data set in eight different exposures for the two rectified views.

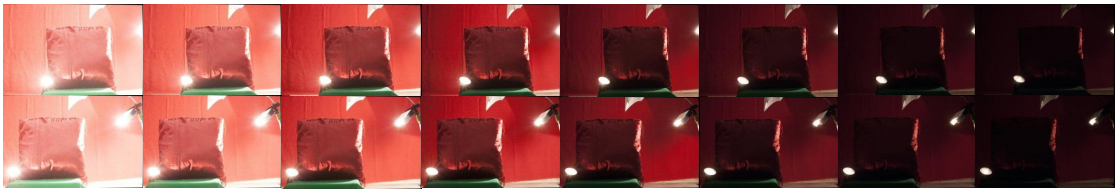


Figure 3.2: A complete LDR sample from the multi-exposed stereo data set. First row: left view, second row: right view. Exposure times from left to right: 1/15s, 1/30s, 1/60s, 1/125s, 1/250s, 1/500s, 1/1000s and 1/2000s. Baseline: 150 mm.

Capturing a scene in multiple exposures, when the camera and the scene are both still, is very common [14] and does not require so much effort or time since most of the cameras can be configured for this purpose. Our data set consists of:

1. Raw LDR stereo images in eight different exposures of: 1/15s, 1/30s, 1/60s, 1/125s, 1/250s, 1/500s, 1/1000s and 1/2000s for the left and right views.
2. Eight pairs of rectified left and right images for each of the eight exposure using [Bra00].
3. HDR stereo images for each view in .hdr format.
4. Log-HDR stereo images. These images are used in Chapter 4 for HDR stereo matching computations.
5. Various tone mapped stereo images using the HDR stereo pairs. These tone mapped stereo pairs are used in Chapter 5 for tone mapped stereo matching computations. More details about the tone mapping operators are explained in future chapters.

All of the stereo pairs such as the LDR, HDR or tone mapped ones are rectified to achieve epipolar geometry and simplify the search from two dimensions into one dimension. Rectification is considered as a preprocessing step for stereo matching to proceed the matching function from a two dimensional search to a one dimensional one. Regardless of the type of input data we feed to our stereo matching approach, we will always make sure throughout this thesis that the stereo pairs are rectified.

Capturing Setup: Our laboratory setup is shown in Fig. 3.1. The scenes are arranged on an optical table. For image capturing, we use a Canon EOS 1D Mark III DSLR with a 28.1 x 18.7 mm sensor capable of delivering RAW images of 3609x2699 pixel size. The images were shot with a focal length between 45 mm and 60 mm. The camera is mounted on a 150 mm motorized linear stage placed on one end of the table. We capture both left and right views with the same camera using a controller software. This configuration ensures that the intrinsic camera parameters of our vision system remain identical in all captures. We use several powerful direct light sources to generate highly saturated regions in our images as shown in the top right image of Fig. 3.1.

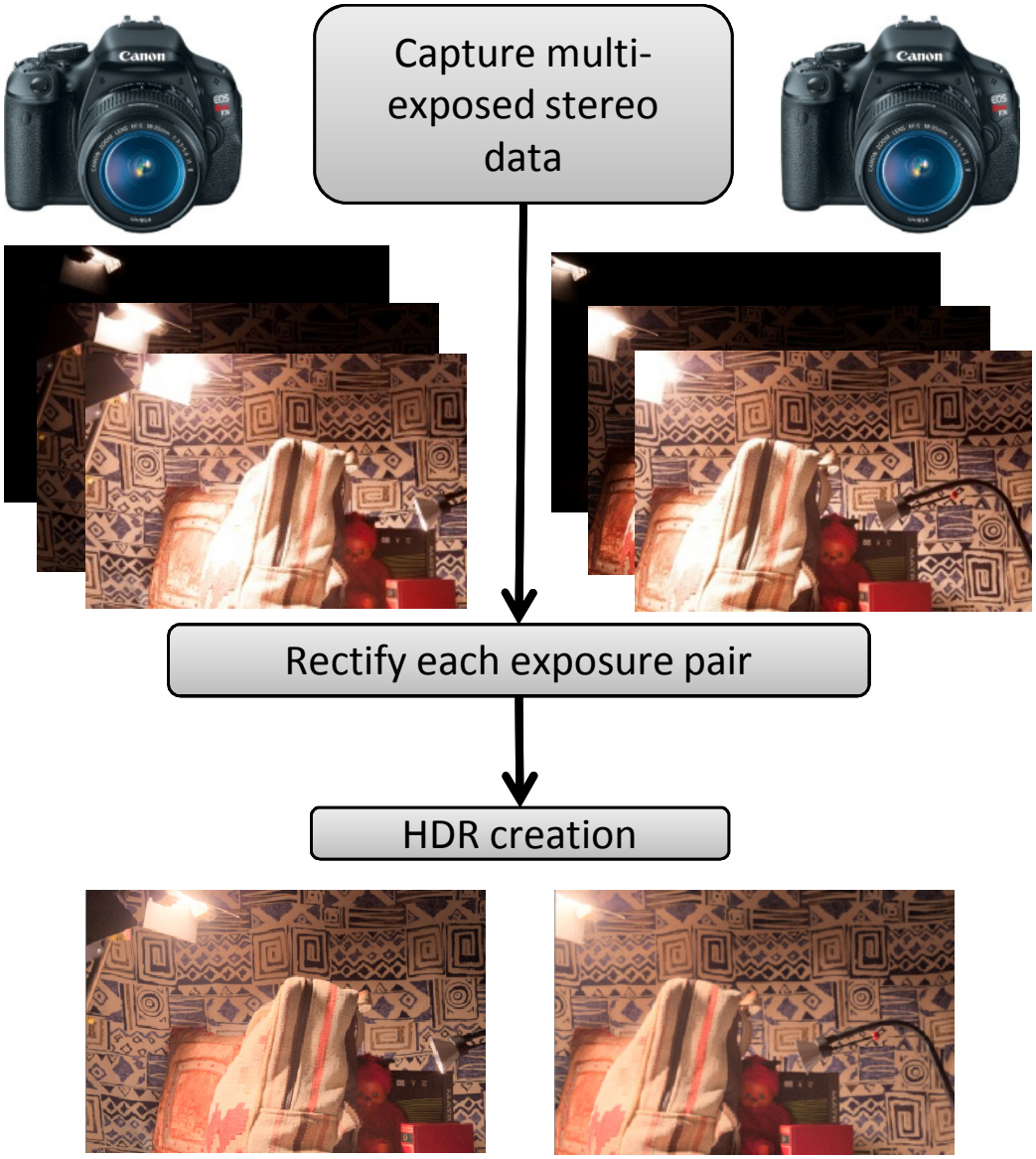


Figure 3.3: Stereo HDR capturing process. Note that the HDR pair are tone mapped using Drago TMO [DMAC03] for better display.

In view of HDR stereo matching, we seek to capture scenes which are interesting from both the HDR and stereo matching points of view. As mentioned before, we use an optical table with a fixed 150 mm motorized rail to move the camera. With this configuration we cannot achieve window scenes or sky HDR scenes in our current lab environment. To generate a variety of indoor HDR scenes, we use different light sources and objects with different sizes, material properties, and textures. Placing some objects close to strong

Figure 3.4: Stereo HDR data set example. First row: left and right views of an HDR pair (tone mapped using Drago TMO [DMAC03] for better display) with a baseline of 150 mm and dynamic range of 8140:1. Second row: left view in the highest, middle and lowest exposures (1/15s, 1/250s, 1/2000s). Third row: same as the second row for the right view.

light source causes under-exposed areas behind them.

Capturing the light source as well as under-exposed regions in one image enables us to reach a high dynamic range of luminance. From the stereo matching point of view, we place highly textured as well as low textured objects at different distances from the camera to achieve variations in disparities. Some slanted objects such as the pillow in Fig. 3.4 will make the stereo matching process more challenging, since many stereo matching algorithms favor front-to-parallel surfaces [BRR11].

LDR Multi exposed Stereo Pairs: The eight exposures for left and right views are captured using our camera setup explained above. Using our camera controller software, we configure the camera with the eight predefined exposures and a baseline value. The baseline value in our experiments was set to 75mm or 150mm. The camera takes the eight exposures for the left view and moves on the mechanical stand showed in Fig. 3.1 according to the baseline value set in the controller software and takes eight exposures of

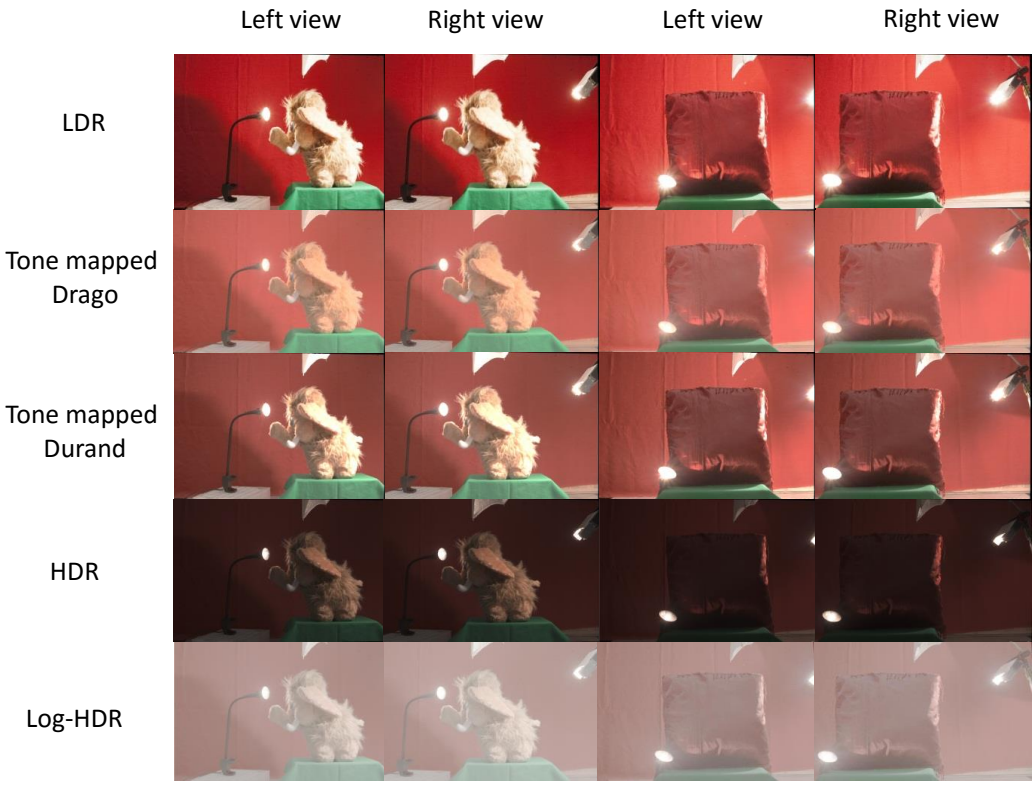


Figure 3.5: Two samples from of our data set. Samples of LDR mid-exposed, two different tone mapped (using Drago and Durand tone mapping operators), original HDR, and log-HDR stereo images. Baseline: 150 mm.

the right view. Fig. 3.2 shows one of our samples in every exposure.

HDR Stereo Pairs: We acquire our stereo HDR data following the steps illustrated in Fig. 3.3. The HDR image of each stereo view is computed using the corresponding multi-exposed rectified images for that view following the approach in [DM97]. Fig. 3.4 presents a stereo HDR example as well as three of the eight captured exposures for the left and right views.

Tone Mapped Stereo Pairs: Throughout this thesis, several well-known tone mapping operators are used, evaluated, analyzed and compared for the purpose of stereo matching of HDR scenes. Tone mapping operators are applied on the original HDR image pairs. In this chapter, for the purpose of introducing our data set, we show some of the examples of tone mapped HDR images for the left and right views. Figure 3.5 illustrates two samples of our HDR stereo data set. Each row shows a stereo pair in a different format, including the middle exposure of an LDR image, two different tone mapped images using Drago [DMAC03] and Durand [DD02] tone mapping operators applied on HDR images,

the original HDR pair computed from the multi exposed LDR images and the log-HDR images used for our computations. Please note that the HDR images appear very dark because they contain much higher range of values (0-8000) comparing to LDR images (0-255) and unless they are tone mapped and compressed to be shown on displays or printers, they look dark.

3.2 HDR Stereo Data Set

In this section, we introduce some of the successful and unsuccessful capturing samples of our data set. We captured twenty two multi exposed left and right images where only 12 of them were meeting our criteria to be considered as HDR stereo images. We are going to discuss some of our failed samples as well as the successful ones in this section. The successful samples from the data set are being used throughout this thesis for our experiments and the unsuccessful ones are only discussed in this chapter. Our analysis of the reasons for failed captures are discussed in a Capturing Lessons Learned section below.

3.2.1 Capturing Lessons Learned

Setting up the capturing process and defining the requirements for capturing HDR stereo images had its own challenges. We had few trial and errors based on the conventional, low dynamic range stereo capturing processes [Sch07, HS07]. Few important factors in the real world scenes which make stereo matching challenging are aimed to be captured in our data sets.

The most obvious requirement was for our scenes to contain a higher dynamic range of luminance levels. For this purpose the best would have been to capture outdoor scenes but unfortunately because of our laboratory setup limitations, we had to limit ourself to indoor bright scenes and environments.

An ill-posed problem in stereo matching is finding the matching pixel in low textured areas. Working in radiance space or HDR space can contribute to solve this problem. The second requirement we had for our data set was to contain low textured areas. Single color cloth worked the best according to our experience. We made sure to include some smooth and low textured regions in our scenes. More detail on results is available in Chapter 4.

Another challenge in stereo matching is matching scenes with bigger baselines or multi view matching where the lighting changes results in different color capturing between matching point and results in mismatching pixels in the disparity map. Since we are experimenting the effectiveness of stereo matching in HDR space, we made sure to capture scenes with bigger baselines.

The combination of the above challenges and the way they influence each other is key in capturing. For example having a smaller baseline can be very challenging in extreme

lighting conditions but if we capture in room lighting a larger baseline is required in order for us to show case the effectiveness of working in HDR space compared to conventional LDR space.

We started with capturing 75 mm baseline stereo images. After applying different stereo matching techniques on the image pairs, we realized 75 mm is not challenging enough for computing disparity maps in our laboratory's lighting conditions. In our first few scene setups we did not include light sources in neither of the left or right images but only tried to create bright conditions by having light sources close to the objects. As the first modification to our settings, we increased the baseline from 75 mm to 150 mm. That lead us into more suitable data sets but as our second attempt of settings modification, we also included the light sources in at least one of the left or right scenes. In addition, we started our process by capturing only 4 different exposures for each image which was not enough for the dynamic range we were trying to cover in our data set, especially having the light source captured. We gradually increased the number of exposures until we finalized the exposure numbers at 8 different exposures.

First we tried using available data sets or capturing methods such as Middlebury or Bumblebee stereo vision camera. We could not generate the challenging type of stereo data sets mentioned above with none of these techniques. Figures 3.6 to 3.8 show examples of our attempts with available conventional techniques to generate stereo HDR images.

We used the Bumblebee XB3 which is a 3 sensor two baseline stereo camera. It features 1.3 mega-pixel sensors and has two baselines available for stereo processing. The extended baseline and high resolution provide more precision at longer ranges, while the narrow baseline improves close range matching and minimum-range limitations.

In the rest of this section we share some of our unsuccessfully captured tries which resulted in data sets that were not satisfying the requirements of an HDR stereo set for different reasons.

1. **Lighting Conditions**

   As we mentioned before, even though Middlebury dataset scenes contain multi-exposed images, they are not considered HDR scenes. Therefore, a greater dynamic range of luminance values are required for our capturing. Figures 3.6 and 3.7 are examples from our experiments which did not qualify as HDR scenes. Figure 3.6 is the Doll scene and Figure 3.7 is the Bowling scene from Middlebury dataset [HS07]. We used three exposures and bracketed technique [DM97] to generate HDR images for left and right views. The scene is a LDR indoor scene with highly textured objects and background which is very suitable for stereo matching. Running the same stereo matching method [HRB$^+$13] on the middle exposed pair of images (which contain the most level of lighting detail among different exposures) as well as the HDR image pair results in an interesting observation.

   The Middlebury data set we used for our experiments was composed of 9 different data sets. Each dataset consists of 7 views taken under three different illuminations

and with three different exposures [HS07]. We used all three exposures and views 1 and 5 as our left and right images. Please note that ground truth disparity maps are only available for the mentioned views.

Another captured data set using the Bumblebee stereo camera is shown in Figure 3.8. Bumblebee was a limiting device for us compared to our professional Nikon camera. Even though Bumblebee provides us with eight different exposures but they are fixed exposures. We need to change the exposures based on the scene lighting and characteristics. As shown in our data set some of the scenes are brighter and our lowest exposure value is different in these cases compared to some of our other scenes which are darker.

2. **Number of Exposures**

   HDR scenes require more exposure variances in order to cover the greater luminance level compared to LDR scenes. We started our capturing process with three exposures as shown in Figure 3.9 and 3.10. We gradually increased the number of exposures to four (Figure 3.11 and Figure 3.12), six (Figure 3.13) and eventually decided that for the dynamic range of our scenes set up, eight exposures provide sufficient information for bracketing purposes.

   Figures 3.6, 3.7, 3.9 and 3.11 show a minor difference between the disparity map calculated from the middle exposed image pair- called LDR disparity map- compared to the disparity map calculated using the HDR stereo pair. The LDR and HDR disparity map results are shown in the third row of the Figures. In few of our captured scenes LDR and HDR results were very close to each other, this means working in HDR space could not add much value to the stereo matching process. These scenes are the ones that do not satisfy the challenges we introduced earlier in this Chapter such as the dynamic range, baseline or low-texture. In this work we focus on challenging scenes for stereo matching.

3. **Low Textured Regions**

   An important goal of this thesis is to show the effectiveness of working in HDR space on accurately calculated disparity maps for low textured regions in the scenes. Because of the limitations of our set up and capturing room, we tried various backgrounds and objects to generate low textured areas in our scenes. Figures 3.6, 3.9 and 3.11 consist of highly textured regions and objects but the rest of our data sets include few low textured regions and therefore satisfy this challenging criteria.

4. **Baselines**

   As indicated below each of our data sets, we experimented on different baselines to measured the relationship between baselines and effectiveness of working in HDR space. Middlebury data sets have the baseline of 160mm. Bumblebee XB3 provides 120mm and 227 mm baselines. In terms of baselines, the Bumblebee stereo camera could satisfy our needs, although it could not provide us with other challenges we were looking for in our capturing data set.

In our set up we captured stereo pairs with 75mm and 150mm baselines. We found the examples with 150mm baseline between left and right images to be more challenging taking into account the changs in the lighting conditions. Therefore, in our chosen data sets throughout this work we mostly use 150mm baseline.

5. **Processing Order**

   There are different formats of saving captured images and different orders one could apply the processes including calibration, compression, rectification, and HDR creation. In this section we explain few of our best practices in terms of the process orders we tried throughout our experiments.

   The first obvious step is to calibrate the camera. We captured 8 left exposures and then 8 right exposures for each scene capturing. In order to make sure we do not loose any information, we save the raw images in Tiff format and demosaic them. next step was to rectify each left and right pair matched by their exposure times. This way we have 8 fully rectified raw image pairs. Next and last step was to use each set of images with different exposures to create the HDR image. We create a left HDR image using all rectified exposures for the left and separately repeat the process for the right pair.

   It should be mentioned that we also tried using JPEG images instead of raw images which resulted in less accurate data collection. In another attempt we tried to first create the HDR images for the left and right images and then do the rectification which also did not achieve the best results. We recommend using raw and fully rectified images to create HDR stereo pairs.

A complete LDR sample from our multi-exposed stereo data set. First row: left view, second row: right view, third row, left to right: LDR and HDR disparity maps. Exposure times from left to right: 1/1000s, 1/500s, 1/250s, 1/125, 1/60s and 1/30s. Baseline: 150 mm.

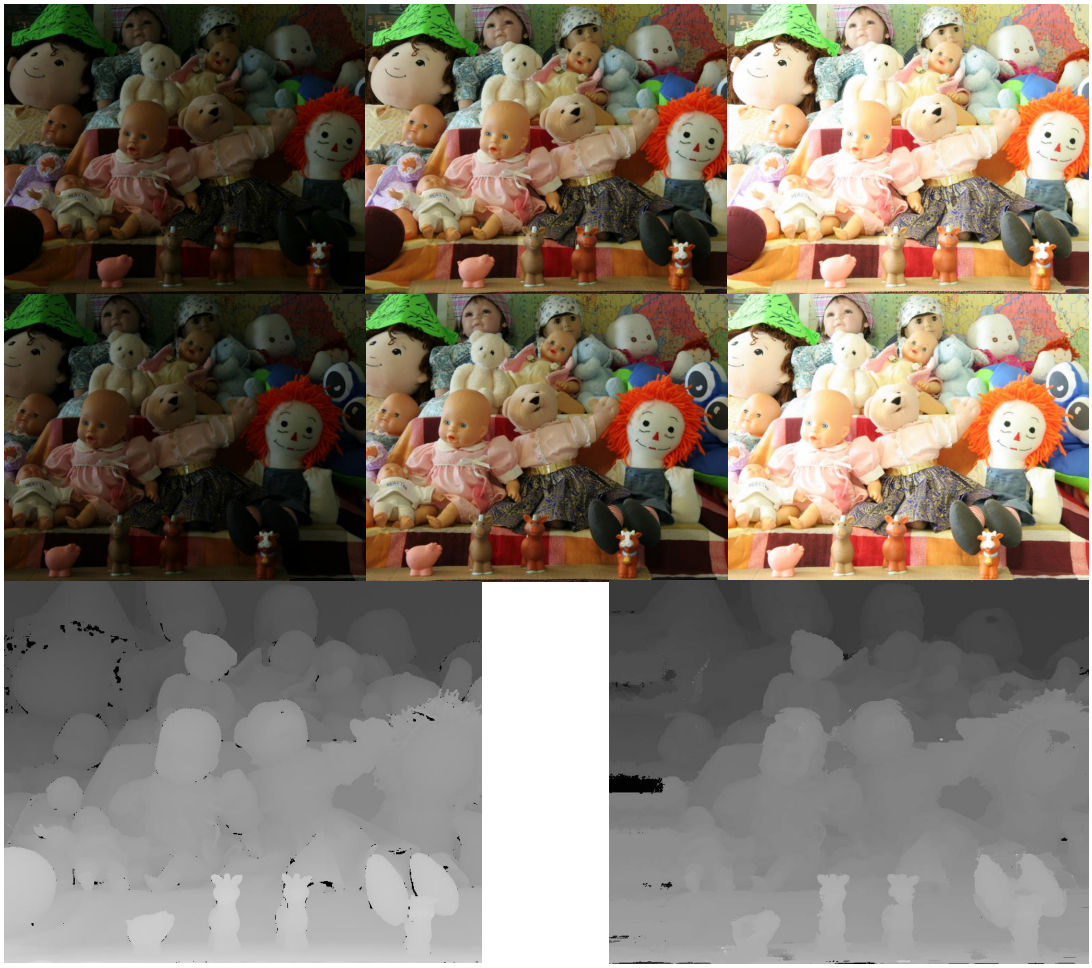


Figure 3.6: A complete LDR sample from the Middlebury multi-exposed stereo data set [HS07]. First row: left view, second row: right view, third row, left to right: LDR and HDR disparity maps. Exposure times from left to right: 1/4000s, 1/1000s and 1/250s. Baseline: 160 mm, between views 1 and 5 of Middlebury dataset. LDR disparity map is computed using the middle exposure, 1/1000s.

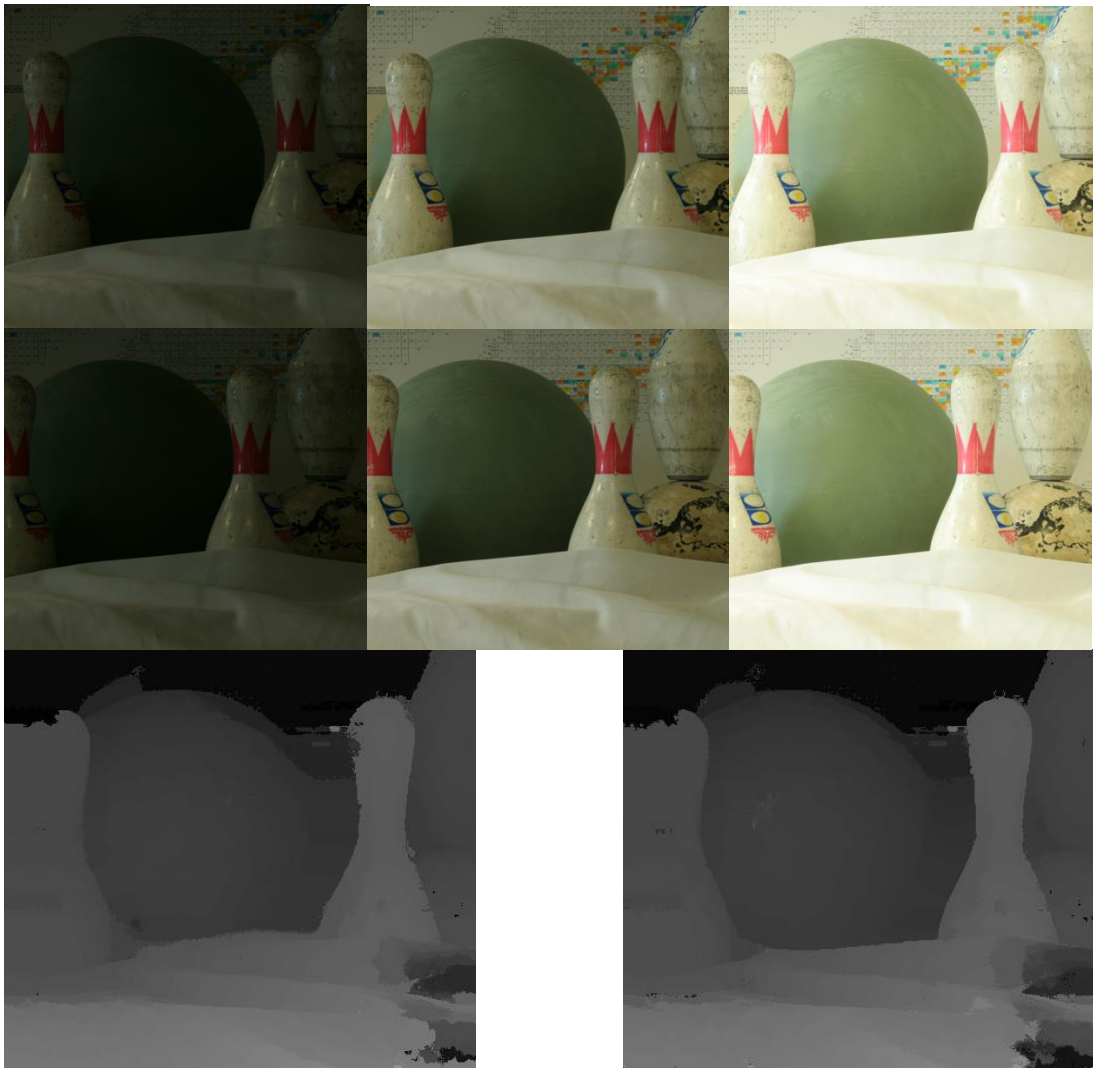


Figure 3.7: A complete LDR sample from the Middlebury multi-exposed stereo data set [HS07]. First row: left view, second row: right view, third row, left to right: LDR and HDR disparity maps. Exposure times from left to right: 1/2000s, 1/500s and 1/125s. Baseline: 160 mm, between views 1 and 5 of Middlebury dataset. LDR disparity map is computed using the middle exposure, 1/500s.

Figure 3.8: A complete LDR sample from our multi-exposed stereo data set captured with Bumblebee stereo camera. First and second row: left view, third and forth row: right view, fifth row, left to right: LDR and HDR disparity maps. Exposure times from left to right (shown in two rows): 1/1024s, 1/512s, 1/256s, 1/128s, 1/64s, 1/32s, 1/16s and 1/8s. Baseline: 227 mm. LDR disparity map is computed using the middle exposure, 1/64s.

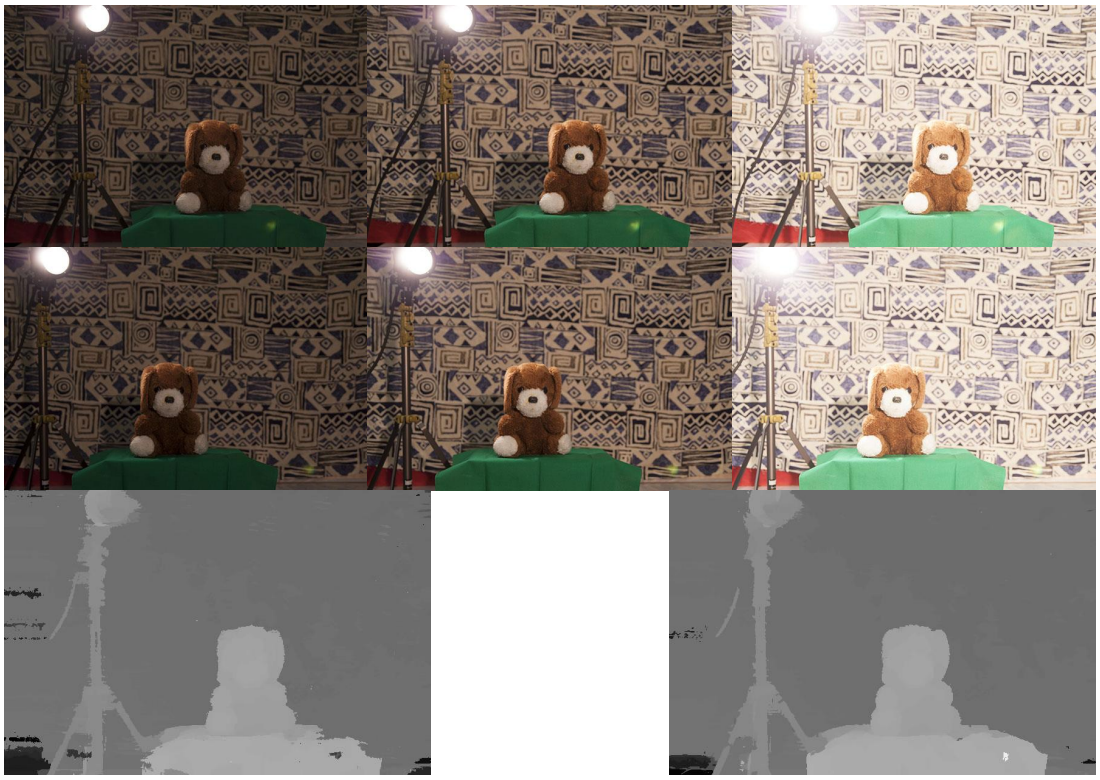


Figure 3.9: A complete LDR sample from our multi-exposed stereo data set. First row: left view, second row: right view, third row, left to right: LDR and HDR disparity maps. Exposure times from left to right: 1/200s, 1/100s and 1/50s. Baseline: 150 mm. LDR disparity map is computed using the middle exposure, 1/100s.

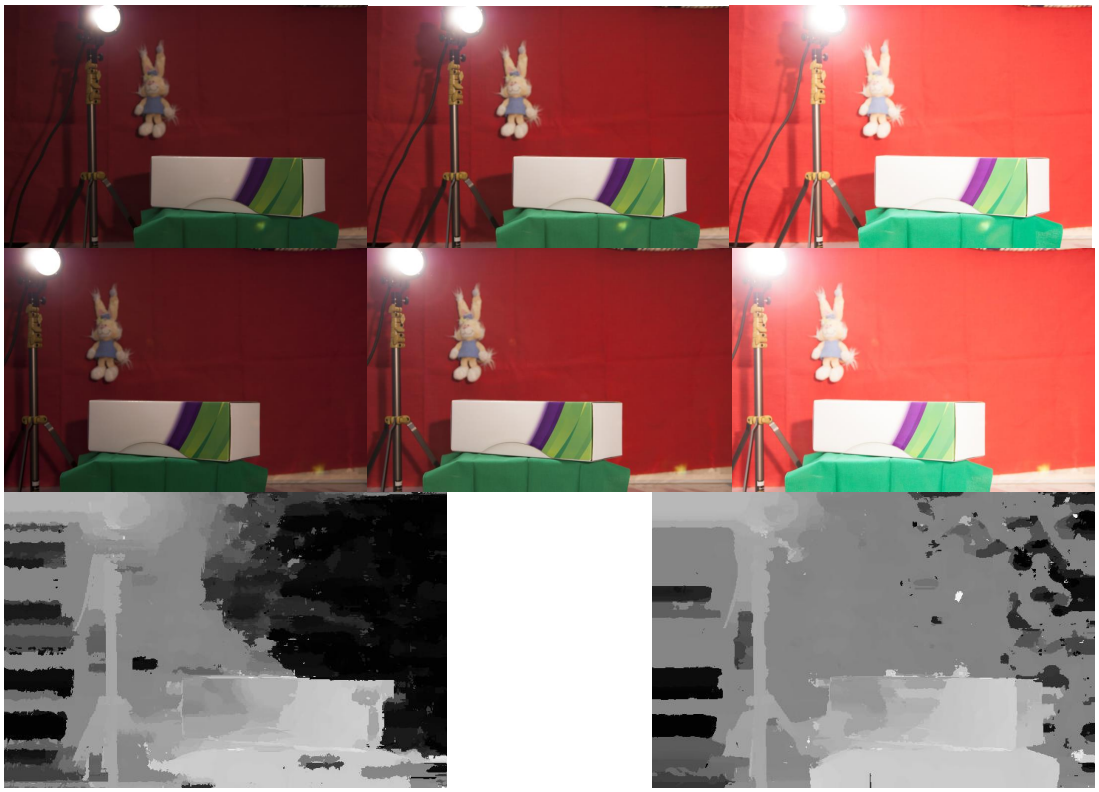


Figure 3.10: A complete LDR sample from our multi-exposed stereo data set. First row: left view, second row: right view, third row, left to right: LDR and HDR disparity maps. Exposure times from left to right: 1/160s, 1/60s and 1/20s. Baseline: 150 mm. LDR disparity map is computed using the middle exposure, 1/60s.

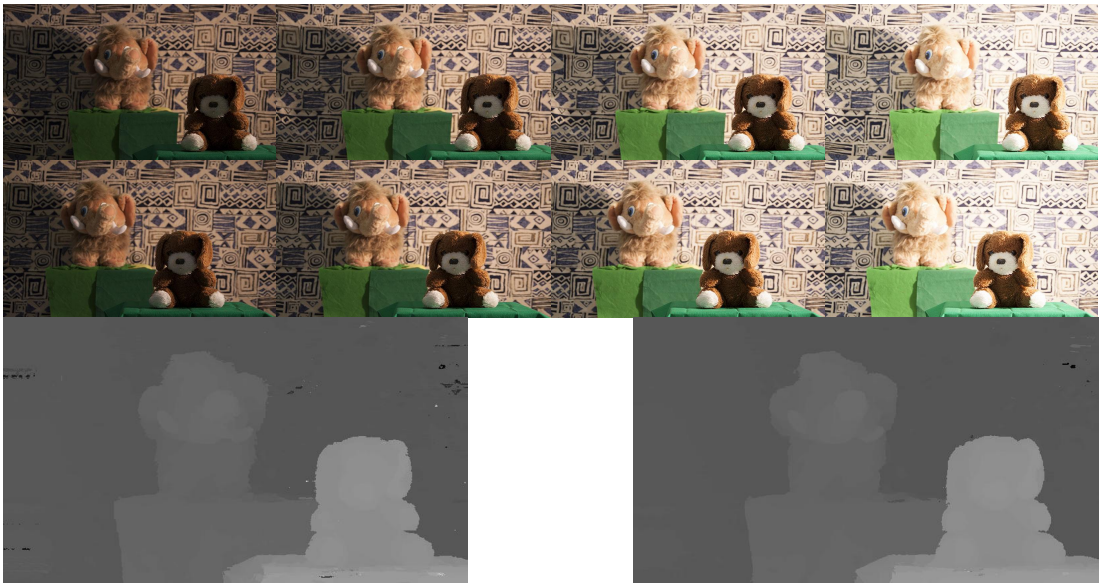


Figure 3.11: A complete LDR sample from our multi-exposed stereo data set. First row: left view, second row: right view, third row, left to right: LDR and HDR disparity maps. Exposure times from left to right: 1/200s, 1/100s, 1/50s and 1/25s. Baseline: 75 mm. LDR disparity map is computed using the middle exposure, 1/50s.

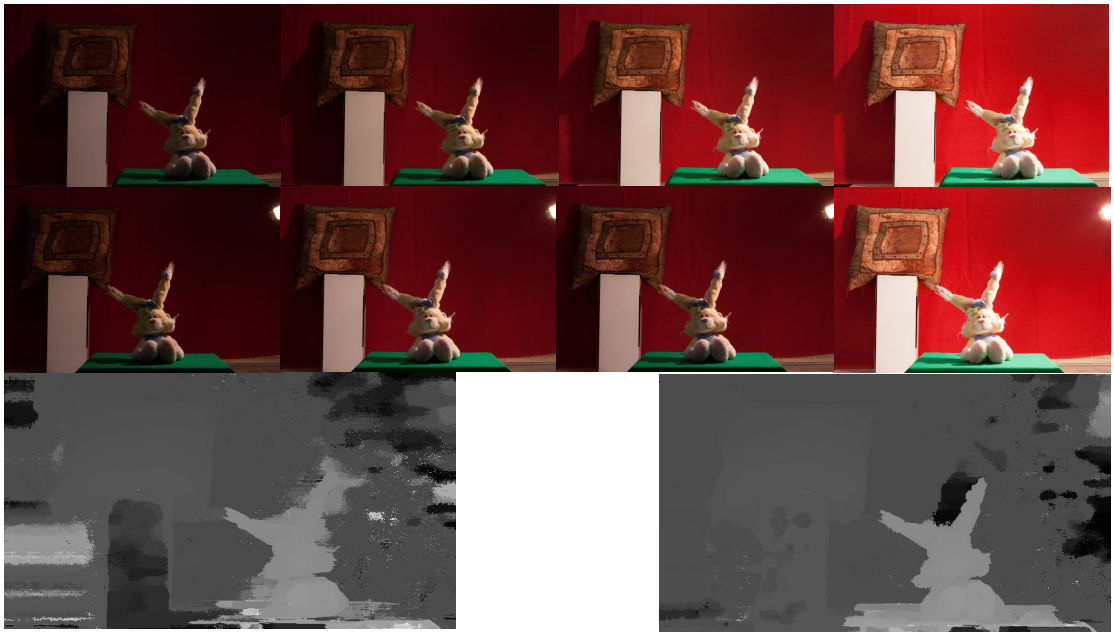


Figure 3.12: A complete LDR sample from our multi-exposed stereo data set. First row: left view, second row: right view, third row, left to right: LDR and HDR disparity maps. Exposure times from left to right: 1/125s, 1/60s, 1/30s and 1/15s. Baseline: 75 mm. LDR disparity map is computed using the middle exposure, 1/30s.

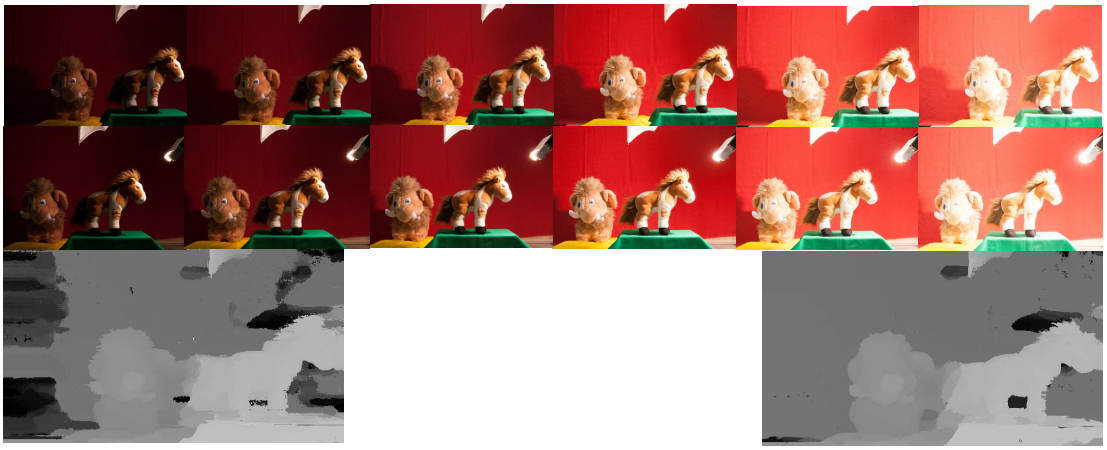


Figure 3.13: A complete LDR sample from our multi-exposed stereo data set. First row: left view, second row: right view, third row, left to right: LDR and HDR disparity maps. Exposure times from left to right: 1/1000s, 1/500s, 1/250s, 1/125, 1/60s and 1/30s. Baseline: 150 mm. LDR disparity map is computed using the middle exposure, 1/125s.

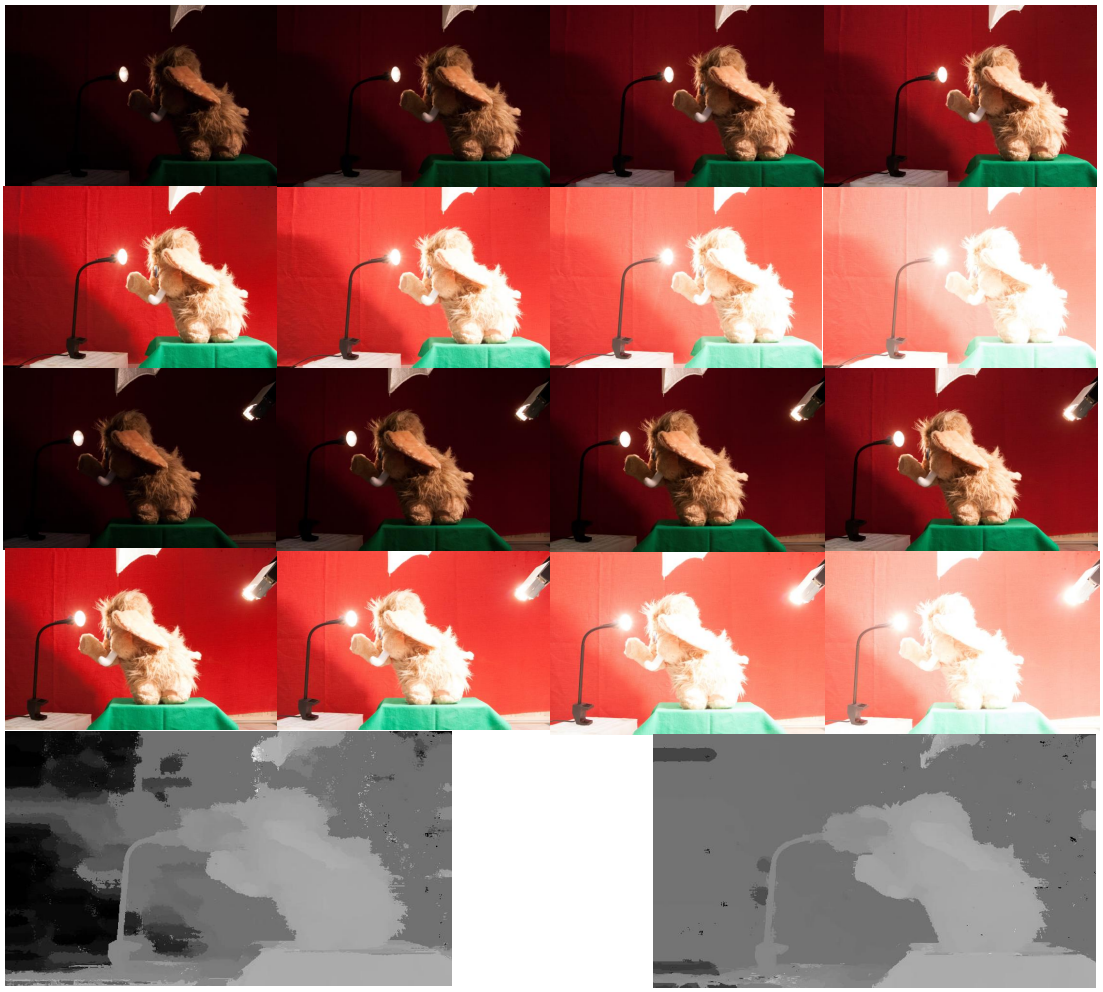


Figure 3.14: A complete LDR sample from our multi-exposed stereo data set. First and second row: left view, third and forth row: right view, fifth row, left to right: LDR and HDR disparity maps. Exposure times from left to right (shown in two rows): 1/2000s, 1/1000s, 1/500s, 1/250s, 1/125s, 1/60s, 1/30s and 1/15s. Baseline: 150 mm. LDR disparity map is computed using the middle exposure, 1/125s.

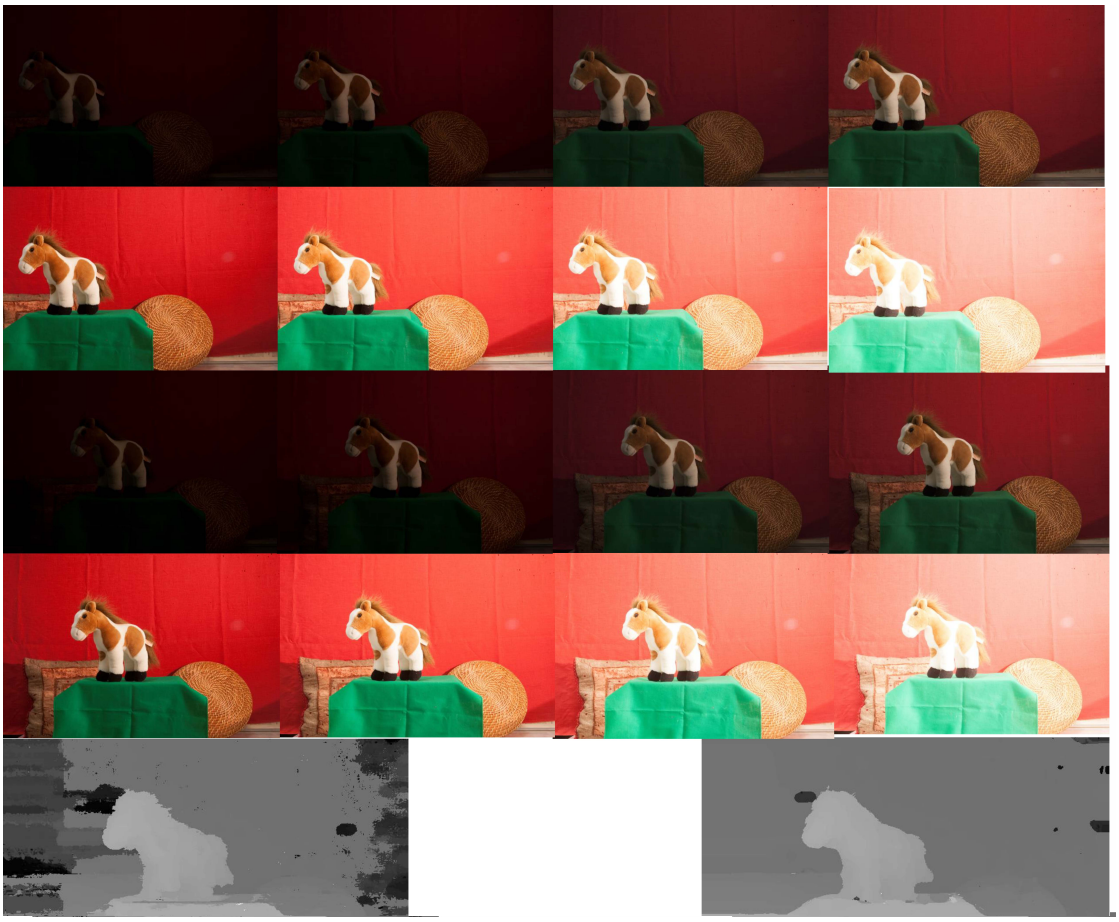


Figure 3.15: A complete LDR sample from our multi-exposed stereo data set. First and second row: left view, third and forth row: right view, fifth row, left to right: LDR and HDR disparity maps. Exposure times from left to right (shown in two rows): 1/400s, 1/200s, 1/100s, 1/50s, 1/25s, 1/12s, 1/6s and 1/3s. Baseline: 150 mm. LDR disparity map is computed using the middle exposure, 1/50s.

CHAPTER 4

HDR Stereo Matching

During the last few decades, stereo matching has become more and more important for many applications based on 3D reconstruction such as depth image based rendering, 3DTV, and multi-view autostereoscopic displays. The main goal of stereo matching is to reconstruct 3D real world information from a pair of images obtained from two different views. There are many stereo matching algorithms which perform well when there are no under-exposed or over-exposed areas in the scene. However, it is difficult to obtain a high quality disparity map using conventional photography, when a scene contains high dynamic range content. In this chapter, we address the question in which way the HDR and stereo processing steps should be combined in order to maximize the quality of the disparity maps that can be obtained from a given set of multi-exposed stereo input images.

We introduce two different ways of computing disparity maps for HDR scenes in this chapter. The first approach is to customize available stereo matching techniques to work on stereo HDR input data. Therefore, we use the HDR stereo data set introduced in Chapter 3. An alternative solution to calculate disparity maps in HDR scenes with a large luminance spectrum is to use a customized structured light approach for multi exposed images. These two approaches are described in detail in the rest of this chapter.

4.1 Stereo Matching in HDR Space

Our main contribution in this section is proposing a stereo matching approach that is tailored to work for HDR scenes. We replace color information with relative radiance information of the scene. Gradient values are calculated from luminance values in log domain rather than from luma as in conventional methods. We embed our HDR specific approach into a state of the art stereo matching technique based on cost volume filtering [HRB+13].

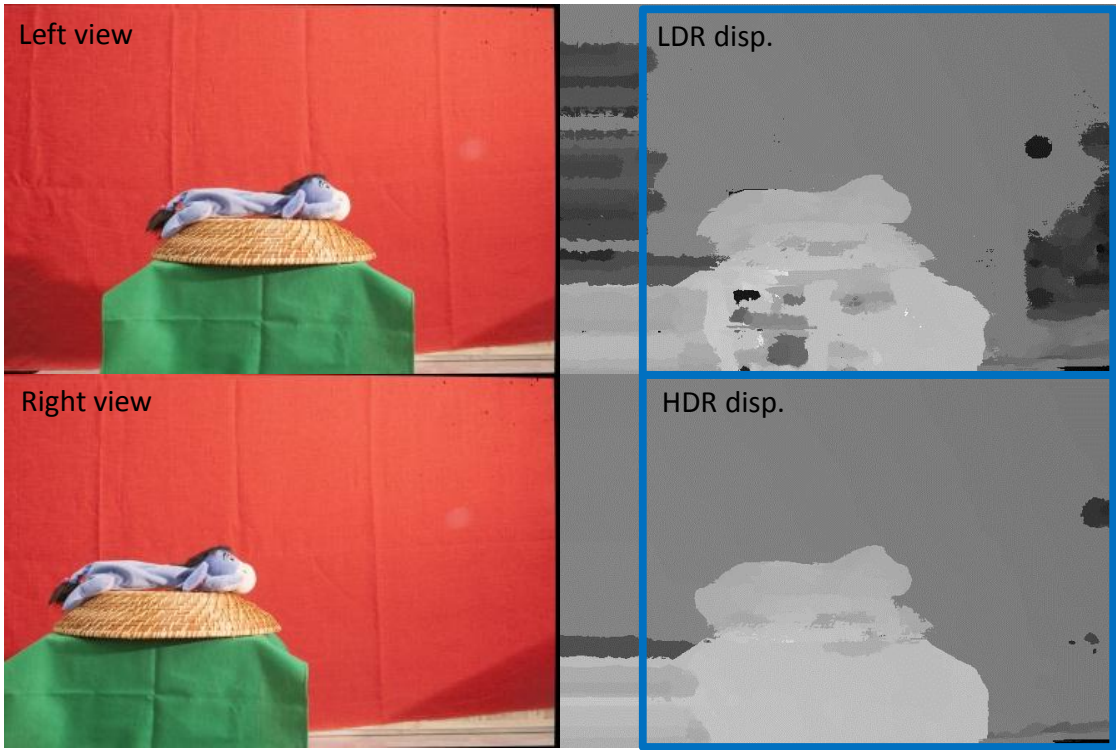


Figure 4.1: LDR-HDR stereo matching comparison. First column: left and right views of a stereo pair. Second column: LDR versus HDR disparity maps. Blue boxes highlight the regions visible in both views. Baseline: 150 mm.

A stereo matching pipeline is proposed based on the cost volume filtering approach for HDR images. The proposed approach can apply to other state or of the art stereo matching techniques. We chose the cost volume filtering approach based on it's proven performance in recent years. Our method aims to exploit the radiance information available in HDR scenes, in contrast to conventional stereo techniques that rely on intensity. Pixel intensity values are rarely the true measurement of the scene radiance, because of an unknown, nonlinear mapping of the radiance to intensity values while photographing. For our experiments, we use the data set consisting of multi-exposed images as well as HDR images for the left and right views introduced in Chapter 3. In our evaluation performed on the LDR-HDR stereo data set, we show that our proposed HDR stereo matching approach improves disparity maps in saturated regions and low-textured areas.

An LDR, middle-exposed sample from our data set is shown in Fig. 4.1. The first column shows the stereo pair input to the matching algorithm. Second column is a comparison between the disparity maps calculated using conventional stereo matching (cost volume filtering [HRB$^+$13]) and our customized HDR stereo matching. As shown in the disparity results, a substantial amount of miss-matched labels in the LDR disparity map are

corrected working in radiance space. Disparities are calculated for the left view. The cost volume filtering customization and experimental results are explained in details in the rest of this section.

4.1.1 Cost-volume filtering for HDR stereo images

Adapting HDR content to a stereo matching pipeline requires deep understanding of human perception as well as applying special perceptual techniques based on physiological research. The easiest way to adapt image processing algorithms to HDR images is to convert luminance values to the logarithmic domain [HRB+13], however this does not preclude the possibility of alternative approaches. This can be explained according to the Weber-Fechner law on human sensitivity to luminance [Roe98]. Unlike LDR luma values, HDR luminance values are not perceptually uniform. In our experiment we found that using *log domain luminance values*, results in more discriminative gradient information and disparity maps. Throughout this thesis, we work in log luminance level for our HDR and tone mapped stereo matching.

We modify the fast cost-volume filtering method introduced in [HRB+13] to work in radiance space. Hosni follows a discrete label-based approach by calculating costs for choosing a label, which corresponds to a disparity. These costs are determined using pixel correlations in the left and right image by calculating absolute differences in the intensity space. The approach consists of several steps:

1. constructing the cost-volume
2. filtering the cost volume using guided filter [HST10]
3. label selection (winner takes all, optimization problem)
4. occlusion detection and filling
5. post-processing

We transfer this approach to scene relative radiance space. Gradient information is more discriminative than color similarity (which lacks accuracy) for matching [CL08, BC10, DMVC11], since it is more robust to illumination changes. As gradient information is not available in low-textured areas, therefore we exploit HDR radiance values as a replacement for color information especially in low-textured regions.

One could use the conventional stereo matching methods with the conventional LDR stereo input images of the HDR scene. This would result in disparity maps such as the one presented in Fig. 4.1. We used the cost-volume filtering [HRB+13] approach. The matching cost calculation using the color intensity values (I_i) and luma gradient information ($\bigtriangledown_x I_i$) is formulated in Eq.4.1 [HRB+13].

$$C_{i,d} = \alpha \cdot min[\|I_i - I'_{i-d}\|, \tau_1] + (1-\alpha) \cdot min[\|\bigtriangledown_x I_i - \bigtriangledown_x I'_{i-d}\|, \tau_2]. \tag{4.1}$$

The cost-volume entry, $C_{i,d}$, determines how well a pixel i in the left image matches the same pixel in the right image shifted by vector (disparity) d in the x direction. Here, $\bigtriangledown_x$ is the gradient operator in the x direction. For weighting the color and gradient information α is used, and τ values are truncation values. We refer to this approach as *LDR stereo matching* in the rest of the paper.

Modified Cost-volume Computation: we calculate the cost values as shown in Eq. (4.2). The cost-volume, $C_{i,d}$, determines how well a pixel i in the left HDR image matches the same pixel in the right HDR image shifted by vector (disparity) d in the x direction. For this purpose, we calculate the truncated absolute differences between pixel radiance values R (in the left HDR image) and R' (in the shifted HDR image) as well as gradient differences calculated from the luminance image Lv (in the left HDR image) and Lv' (in the shifted HDR image). Since human vision perception of light and illumination changes could be imitated by a logarithmic function [Roe98], we simply apply the log operator on the luminance channel of stereo HDR images as in Eq. 4.2. We found it more effective to keep the original color radiance values and apply the log function only on luminance values as formulated below.

$$C_{i,d} = \alpha \cdot min[\|R_i - R'_{i-d}\|, \tau_1] + (1-\alpha) \cdot min[\|\bigtriangledown_x \log(Lv_i) - \bigtriangledown_x \log(Lv'_{i-d})\|, \tau_2]. \tag{4.2}$$

Here, $\bigtriangledown_x$ is the gradient operator in the x direction. For weighting radiance and gradient information α is used, and τ values are truncation values.

The easiest way to adapt image processing algorithms to HDR images is to convert luminance values to the logarithmic domain [HRB+13], however this does not preclude the possibility of alternative approaches. The Weber-Fechner law on human sensitivity to luminance [Roe98] is explained in Chapter **??**.

Smoothing the Cost-volume: Various smoothing approaches could be used for this purpose. In [HRB+13] guided filter [HST10] is found to be faster and more effective. We use the same filtering approach just on our scene radiance values. C' is the filtered cost volume.

$$C'_{i,d} = \sum_j W_{i,j}(R) C_{i,d}. \tag{4.3}$$

The filter weights $W_{i,j}$ (in pixel position i and j) are chosen according to the weights of guided filter as used in [HRB+13]. Having the radiance guided image R, weights are defined as:

$$W_{i,j} = \frac{1}{|w|^2} \sum_{k:(i,j)\in w_k} (1 + (R_i - \mu_k)^T (\sum k + \in U)^{-1} (R_j - \mu_k)), \quad (4.4)$$

where μ_k is the mean vector and $\sum_k$ is the co-variance vector calculated in a squared window w_k with dimensions of $(2r+1) \times (2r+1)$, centered at pixel k in radiance image R. Further details can be found in He et al. [HST10].

Label Selection: for each pixel the minimum cost among the calculated costs for different disparities ($d \in D = 1, ..., D$) is chosen after the cost values have been filtered. The filtered cost values of a specified pixel, for all disparity labels, could be seen in graphs shown in Fig. 4.3. The minimum label for pixel i is f_i.

$$f_i = \underset{d\in D}{\operatorname{argmin}} C'_{i,d} \quad (4.5)$$

Occlusion Detection and Filling: to detect unreliable labels we use a left-right cross-checking approach as in [HRB+13].

Post-processing: unreliable labels detected in the last step are

1. replaced first with the lowest disparity value of the spatially closest non-occluded pixel on the same scanline and
2. then are smoothed using the weighted median filter $W^{bf}_{i,j}$ with bilateral filter weights in Eq. (4.6). This filtering is only applied to the pixels with labels deemed unreliable.

$$W^{bf}_{i,j} = \frac{1}{K_i} exp(-\frac{|i-j|^2}{\sigma_s^2}) exp(-\frac{|R_i - R_j|^2}{\sigma_c^2}), \quad (4.6)$$

where σ_s represents the spatial similarity and σ_c represents radiance similarity. As normalizing factor K_i is used.

4.1.2 Implementation, Results and Discussions

Our experiments were performed in MATLAB 2012b. Multi-exposed images were taken with a Canon EOS-1D Mark III camera with a baseline of 75 and 150 mm and a focal length of 45 mm.

Parameters: even though the radiance value range of our images is much higher than conventional 256 values, for our computation we normalized all values to the range of [0-1].

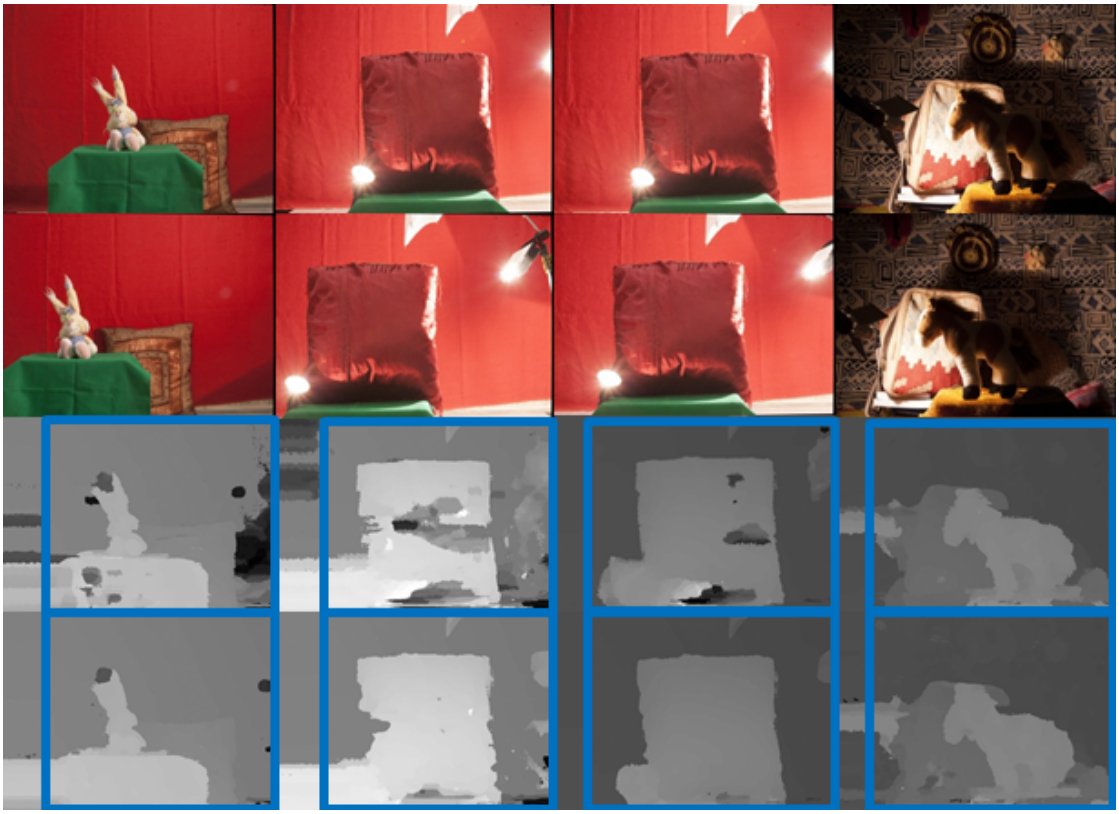


Figure 4.2: LDR-HDR stereo-matching comparison. Rows from top to bottom: left view, right view, LDR disparities, HDR disparities. Baselines, left to right: 150 mm, 150 mm, 75 mm, 75 mm.

It should be noted that despite this normalization the HDR properties are maintained by the finer quantization steps within the unit interval. We also used LDR and HDR images of the same size for our comparison. Therefore, we could apply the same fixed parameters used in [HRB+13] for stereo matching on both the LDR and HDR versions of the code to generate our comparison results.

LDR-HDR Comparison: all LDR disparity results shown in this paper were computed using the fast cost-volume filtering on one of the two middle exposures (1/60s or 1/125s), depending on which one achieved better results.

Experimental Results Discussion: Fig. 4.1 and Fig. 4.2 show the computed disparity maps using LDR and HDR images. It is visible that the disparity information in low textured regions in the foreground and background as well as filled in areas is improved in the HDR case. One can see that for the highly-textured, non-saturated example on the right column of Fig. 4.2, the LDR method performs almost as well as the HDR approach. Contrarily, for the less-textured, partly saturated images in columns 1 to 3 the improvement achieved by the HDR matching is clearly visible. Comparing the two middle

columns in Fig. 4.2, it is noticeable that the same left view processed with different right views (75 mm and 150 mm baselines) results in quite different disparity maps. From closer examination of our results on these and other image pairs, we found some indication that the improvement gained by matching in HDR space gets even more pronounced for larger baselines, as the illumination and color changes between the two images increase. Our experiments show that stereo images with bigger baselines, especially in HDR scenes, cause more challenges to the depth computation process. Higher baselines cause bigger illumination and color change between the pairs which causes problems in finding the correct matches. Working in radiance space, we reduce the mismatches as can be seen in the provided results. The last two columns on the right side of Fig. 4.2 present stereo images with a baseline 75 mm (smaller occluded region, smaller disparity range).

Due to lack of ground truth for stereo HDR images, we seek to get a deeper insight into the observed quality differences between the LDR and HDR matching results by comparing computed cost values in the intensity space versus radiance space. Details of the cost values computed in the intensity (LDR) and radiance space (HDR) for a specific sample are provided in Fig. 4.3. Graphs (g_1), (g_2), (g_3) and (g_4) show the computed matching costs for specified pixels in e and f. In all graphs, cost values (first rows) are calculated from color/radiance information (middle row) and gradient information (bottom row) using our LDR and HDR cost computation algorithms as described before. Pixels x_1 and x'_1 in the LDR and HDR disparity maps (e and f), respectively, achieve a plausible disparity in both cases. Contrarily, x_2 and x'_2 represent the same pixel in the LDR and HDR disparity map for which a wrong disparity in LDR but a correct disparity in HDR was achieved. The top row in graph (g_3) shows how working in conventional intensity space results in very similar matching costs inside low texture areas that are very likely to cause a mismatch. The top row in graph (g_4) presents the corresponding pixel matching costs in radiance space with only one minimum among matching costs. The same effect was observed on other image pairs too. Using scene radiance information instead of intensity values, we achieved one global minimum which is known to be a more reliable situation in optimization algorithms.

4.2 Stereo HDR Disparity Map Computation Using Structured Light

In this section, we present our work towards the generation of disparity maps for HDR scenes using a structured light approach. Structured light matching approaches are amongst the most accurate stereo matchings. They are mostly used for disparity ground truth generation since they are performed offline and are computationally heavy. The development and evaluation of novel stereo matching algorithms that are tailored to the characteristics of HDR images would greatly benefit from the availability of such reference data. We describe our laboratory setup and processing steps for acquiring multi-exposed stereo images along with a corresponding reference disparity map computed by a structured light approach. We discuss the special requirements and challenges which

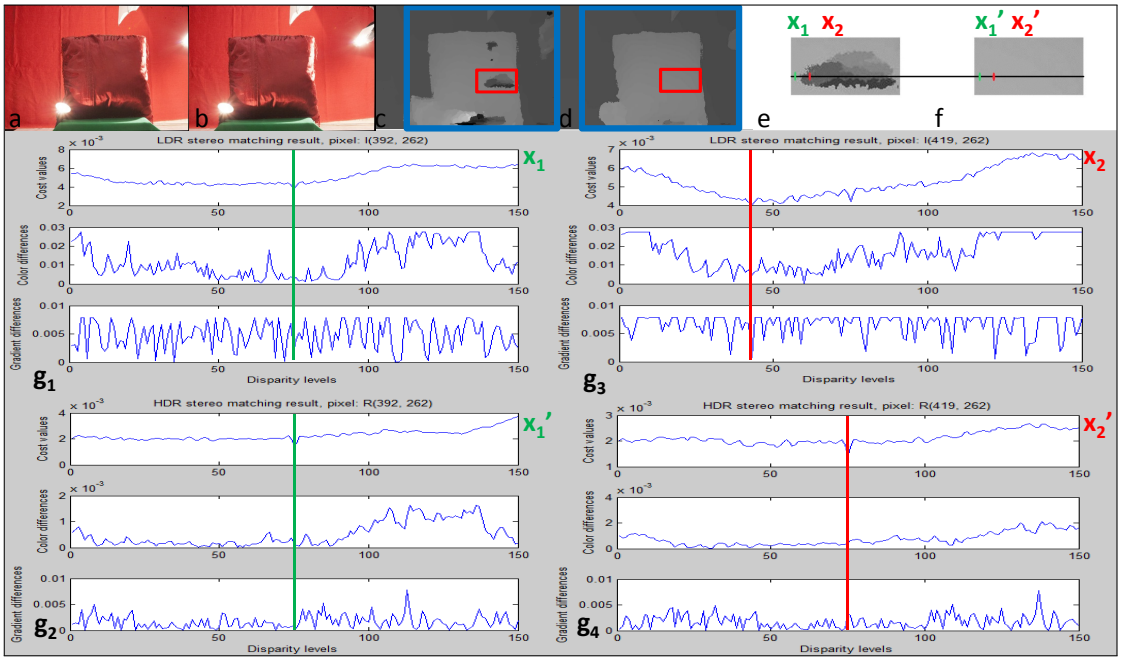


Figure 4.3: LDR-HDR cost calculation comparison. (a) and (b): left and right views; (c) and (d): LDR and HDR computed disparity maps; (e) and (f): zoomed areas from (c) and (d); (g_1), (g_2), (g_3) and (g_4): from top to bottom: cost values, color/radiance-differences and gradient-differences between pixels x_1, x_1', x_2 and x_2' and the pixel shifted by different disparity levels (1-150 pixels) in x direction of the same line. (g_2) and (g_4) show how by working in HDR space we achieved a more pronounced minimum among costs.

HDR test scenes impose on the computation of the reference disparities and show some preliminary results.

We use a structured light technique for our reference disparity computation. In [SFPL10] a quantitative and qualitative survey on different structured light approaches is carried out. We follow the spatio-temporal Gray-code approach because

1. we prioritize accuracy over acquisition time
2. we only capture still scenes
3. the capturing method used in the Middlebury data sets is well proven [SS03b]

In this section, we explain our experimental hardware and software setup (see Fig. 4.4, top left) for (1) the stereo HDR capturing process (see Fig. 4.4, top right) and (2) the disparity map computation process based on structured light (see Fig. 4.4, second row). In Section 4.2.3 the preliminary calculated reference disparity map for the HDR scene captured in Fig. 4.4 is shown and discussed.

Figure 4.4: Scene capturing process. Top left: experimental setup, top right: HDR image capturing, bottom left: example of a horizontal Gray-code image, bottom right: example of a vertical Gray-code image.

In the context of ongoing research towards the development of HDR stereo matching algorithms [AYG13], we seek to acquire a variety of HDR data sets along with corresponding reference disparity maps that can serve as ground truth for evaluating HDR stereo matching algorithms. When setting up the test scenes, we pay special attention to including scene characteristics that are representative of HDR scenes (for example, highly saturated regions) as well as features that are known to be challenging for stereo matching algorithms (for example, lack of texture or slanted surfaces).

4.2.1 Reference disparity computation using structured light

To generate the reference disparity maps, we use the structured light technique with Gray-codes as used for the Middleburry disparity map computation [SS03b] described in Fig. 4.5. A Panasonic projector PT-AH1000AE is placed beside the camera (see Fig. 4.4, top left). It is capable of projecting resolutions up to 1920x1080 pixels with a dynamic range of 500000:1. We also use a desktop PC to control the camera, rail and projector

during the capturing process. The projector is used to generate the binary Gray-code patterns that contain black and white (on/off) pixel values.

The first step according to Fig. 4.5 is to capture all patterned images for both views. To distinguish among n locations, $\log_2(n)$ patterns are required. Each pattern represents one bit plane, and all $\log_2(n)$ patterns together achieve one unique coded value for each pixel. As described in [SS03b], using a projector resolution of 1024 x 768, 10 horizontal and 10 vertical patterns are sufficient to uniquely code each pixel in the scene. This sums up to 40 patterned images, 20 for each of the left and right views. To threshold the pixels more reliably into on or off categories, patterns and their *inverse patterns* are being used, which doubles the image numbers into 80. For more accurate decoding, we capture the same patterns with two different exposures, which results in 160 patterned images to be processed.

As in the stereo HDR capturing process, we first need to rectify the corresponding left and right views for all patterns and exposures [Bra00]. Using the decoded Gray-code patterns, we assign a unique label to each pixel of the scene. Corresponding pixels in the left and right view are found using the unique codes at each pixel (in 1D search, since images are rectified). The result of this correspondence process is called *view disparity*. Using view disparities and code labels, projection matrices for the illumination (pattern) sources can be determined. Reprojecting the code labels into the two view geometry results in *illumination disparities*. By combining all disparities, a final accurate disparity map is achieved. More details about the process can be found in [SS03b].

4.2.2 Post processing

Once the disparity map has been computed as described in 4.2.1, appropriate post processing is required. It is suggested in [SS03b] to use some interpolation method to fill small holes in the ground truth disparity while no solution for recovering large holes is offered. Small holes are caused by noise or wrongly matched pixels which are usually easy to repair or fill in. Large holes appear in (1) areas that are shadowed under all illumination patterns and (2) specular surfaces.

We apply the post processing in two separate phases to, (1) fill small holes and (2) to cover big holes which are the result of shadows in the scene. It should be noted that in this context shadows refers to areas that do not receive any projection pattern, and not to shadowing effects related to other illumination sources, such as lamps, inside the HDR scene. The latter type of shadows, however, may pose problems in stereo matching due to poor texture and noise. All holes have been marked as invalid disparities during a left-right consistency check to be post processed later. We repair invalid pixels by using their corresponding disparity value in a morphologically closed version of the current disparity map [**?**]. Any other filtering method with the same purpose could be used as well.

4.2.3 Results and Discussion

First results of our processing are shown in Fig. 4.6. Areas that are shadowed under all illuminated patterns create large holes which cannot be easily removed or filled in. In the bottom row of Fig. 4.4 a shadowed (non illuminated) area in the top left corner of the scene is visible, which is caused by the lamp. For better visibility, shadow regions are illustrated in the fully illuminated (using projector) image shown in Fig. 4.6 (a). In these regions a lack of illumination codes caused unknown disparity values (holes) in the disparity result (see Fig. 4.6 (b)). At the moment, we manually mark these regions and fill in the holes with values in the closest Euclidean distance. The final disparity map generated in the mentioned way from the left point of view is shown in Fig. 4.6 (d). It should be noted that this is just a preliminary result of our ongoing research.

For comparing the calculated reference disparity of the HDR scene to state of the art stereo matching results, we show the disparity map achieved from the fast cost-volume filtering approach [HRB$^+$13]. Fig. 4.6 (e) shows the disparity calculated from the middle exposed stereo images and Fig. 4.6 (f) represents the disparity computed from the Drago tone mapped [DMAC03] stereo pairs. It is visible that the disparity information gained from the structured light technique is of better quality than the results of the stereo matching approach on LDR and tone mapped stereo images.

As part of our work towards the generation of stereo HDR data sets and ground truth, we explained the hardware setup and visualized the data set generation and disparity computation approach step by step on an example scene. In principle, the chosen structured light approach is capable of delivering high-quality disparity maps that can be used as reference for stereo matching. However, our HDR test scene has also demonstrated the problem of missing disparity information caused by lamps placed in the scene foreground, which block the light pattern emitted by the projector. Since such light sources constitute an important component of meaningful HDR indoor test scenes, suitable ways for alleviating this problem (for example, by using a second structured light projector) should be explored.

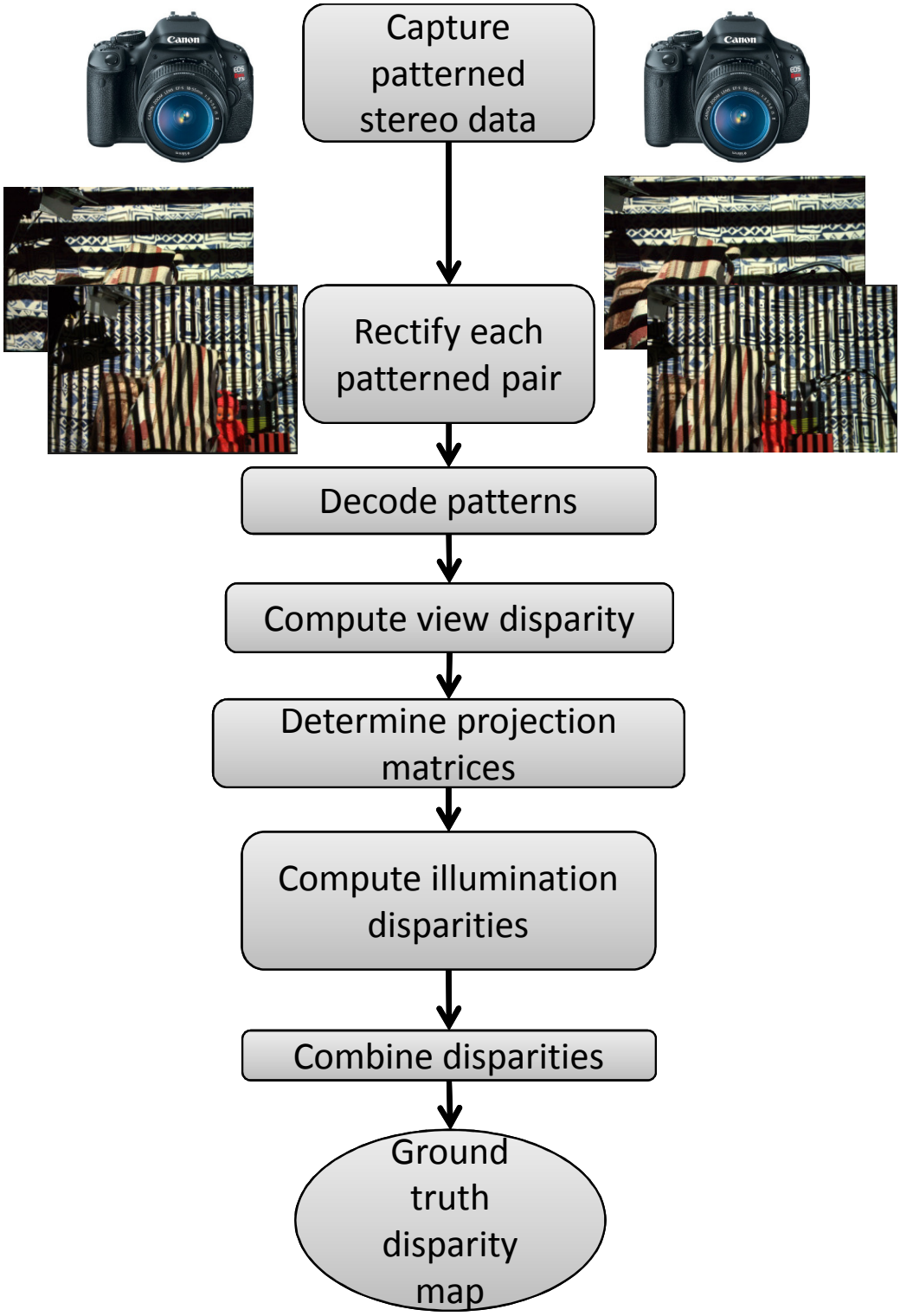


Figure 4.5: Ground truth computation process [SS03b].

Figure 4.6: Reference disparity map computation. (a): Full projected scene, with non-illuminated regions (shadows are marked black) specified; (b): computed reference map before manual post processing (shadow holes); (c): left LDR mid-exposed image; (d): calculated reference disparity after post processing; (e): LDR stereo matching result corresponding to (c); (f): tone mapped stereo matching result corresponding to top row of Fig. 3.4.

CHAPTER 5

Tone-mapped Stereo Matching

HDR images provide greater detail and larger brightness levels than conventional LDR ones. Working in HDR space can lead to better results by using more detailed brightness information. It is not hard to predict that working in radiance (HDR) space provides more informative disparities as discussed in Chapter 4. This is especially true in challenging lighting conditions and low-textured areas. However, two important challenges when switching calculations from LDR to HDR domain are *lack of data* and *backward compatibility*. We talked about how we compensate in this thesis for *lack of data* for HDR stereo matching in Chapter 3 by introducing our own data set. In this chapter we focus on the second challenge which is *backward compatibility* for available stereo matching techniques.

Why do we need a backward compatible HDR stereo matcher? Most of the available stereo matching software approaches are hard-coded to work with 8-bit input images, including the top ranked methods in Middlebury [Sch07]. Although a few methods including some hardware solutions implemented on microprocessors have the advantage of running on floating point data [Kon97]. Furthermore, most of the available stereo matching codes assume a fixed range of intensities for both left and right images while an HDR image pair most probably has different maximum luminance values for each image due to angular differences. See Table 5.1 for information regarding the maximum luminance values in the left and right images of our data set pairs used in the rest of this chapter. The middle exposed LDR images for the left and right views are shown in Fig.5.2.

To generate a backward compatible solution to conventional stereo matchers, we use Tone Mapping Operations (TMO) to compress the dynamic range into conventional range while preserving details of an HDR image. After comparing disparity maps achieved from different TMOs, we propose a fusion framework for achieving more informative disparity maps from tone-mapped, HDR image pairs. In order to estimate corresponding ground truth disparity maps, we make use of our HDR stereo images and an implementation of a

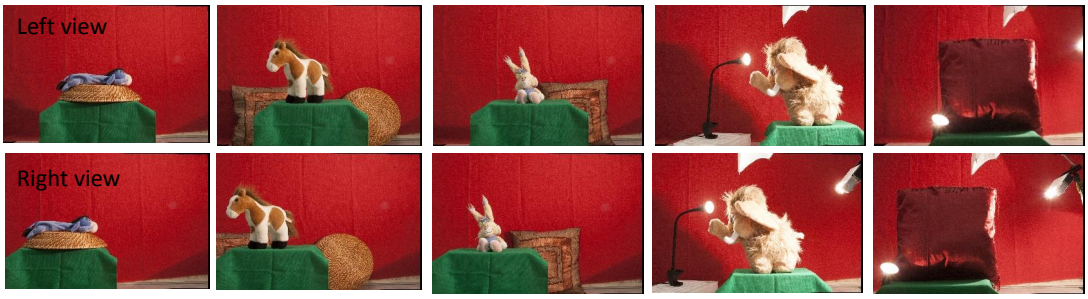


Figure 5.1: Left and right views of the dataset used in this chapter. The maximum luminance values of the scenes for each left and right view are specified in Table 5.1.

Scene	**Left, Max. Luminance (in cd/m^2)**	**Right, Max. Luminance (in cd/m^2)**
Donkey	5412	5536
Horse	5423	5543
Rabbit	5401	5570
Elephant	6802	7153
Pillow	8150	8221

Table 5.1: Maximum luminance value of the left and right images in our data set. The rows of this table are arranged in respect to the data set shown in Figure 5.2 from left to right: Donkey, Horse, Rabbit, Elephant and Pillow.

customized stereo matching approach introduced in Chapter 4. Based on these estimated disparities, we perform an objective evaluation of stereo matching approaches.

Since stereo matching is an old research and industrial field, it is important for new stereo data sets to be runnable on conventional approaches. Therefore, we use tone-mapped image pairs with the standard interface to legacy stereo matchers and propose an algorithm to optimize the disparity maps to achieve closest results to HDR stereo matching.

Even though we captured HDR stereo data, we only used them for our ground truth disparity calculation and used HDR *tone-mapped* data in our proposed framework in this chapter. The first reason for using tone-mapped images (not the original HDR image pairs) is described in the previous paragraph regarding backward compatibility. The second reason is, although our HDR stereo matching approach outperforms the tone-mapped and LDR approaches, it is not fully automatic. We used a manual in-painting post processing to fill-in some of the holes in the disparity maps as shown in Fig. 5.2. Manual post processing and in-painting of disparity maps for the purpose of ground truth generation (evaluation) is a common approach [SS03c, AKCG14], but usually not a scalable or real-time method. We used the original HDR disparity maps just as a reference for our evaluation method for the same mentioned reasons.

In this chapter, we present a graph-cut based disparity map fusion framework using different tone-mapped stereo matching results in order to take into account the best features for stereo matching from several different TMOs. Stereo matching in HDR scenes introduces new challenges to the matching state-of-the-art. Computing the disparity on the tone-mapped image pairs is an approach to solve these challenges, but not so many tone-mapping operators are suitable to be applied on more than one frame or image while keeping the consistency of the images or frames. This problem has recently been addressed in video tone-mapping [EWMU13a].

The remainder of the chapter is structured as follows. We compare the disparity results obtained from different tone-mapped image pairs with a focus on edge-aware filtering based TMOs. Then, our proposed framework to combine several computed disparity maps is introduced in. Finally, Section 4.2.3 debates our experimental results, evaluation and discussions. This Section discusses the HDR stereo image pairs as well as our reference disparity map generation approach used for the quantitative evaluation.

5.1 Highly-ranked TMOs

According to most of the comparisons on TMOs mentioned before, Reinhard TMO [RSSF02], Fattal TMO [FLW02] and Drago TMO [DMAC03] are among the most effective ones. Therefore, we choose them to be among our short listed TMOs which we are evaluating for stereo matching application. In the following subsections we compare and discuss these TMOs from the stereo matching point of view.

Reinhard TMO: is a global TMO which took the inspiration from traditional wet-film photography techniques.

Fattal TMO: represents a local TMO which is based on gradient domain operators.

Drago TMO: is an example of a global TMO which is extending the logarithmic response curves to perform on a wider dynamic range.

5.1.1 Edge-aware TMOs

Our experiments show that edge-aware TMOs achieve more discriminative disparity maps since they are more robust to lighting changes between the left and right images. Some of the TMOs based on the edge aware filtering approach such as Durand TMO [DD02] are also listed among the highly ranked operators in the literature. Based the same idea proposed by Durand, there has been other filtering approaches which can replace the Bilateral filter solution used by Durand. We implement and compare TMOs based on the most popular filtering methods.

Durand TMO: Based on the idea of an image being consisted of a high spatial frequency (LDR), and a low frequency (HDR).

Gastal, He, and Farbman TMOs: Based on the same idea as Durand TMO but using different edge-aware filtering approaches.

The tone-mapped disparity maps are calculated using the cost-volume filtering stereo matching [HRB+13] on every tone-mapped image pair. The TMOs which are used in this experiment are: Reinhard, Durand, Fattal, Gastal, Drago, He and Farbman TMOs. Our main contribution is the introduction of the combination approach to take all the different characteristics of TMOs into account. The TMOs which we used in our benchmark are interchangeable.

TMO Parameters: Most TMOs have tunable parameters. We used the default parameters suggested by the authors in most of the cases for TMOs. Here we list the parameter values which we modified to tune the TMOs for tone-mapped disparity calculation according to our experiments. For more details of the parameters definitions please refer to the references.

- Guided Filter (He TMO [HST10]): $\{r = 9, \epsilon = 0.0001\}$
- WLS Filter (Farbman TMO [FFLS08]): $\{\alpha = 1.2, \lambda = 1\}$
- Domain Transform Filter (Gastal TMO [GO11]): $\{\sigma_s = 60, \sigma_r = 0.33\}$

5.2 Backward Compatible Tone mapped Matching Algorithm: Hybrid Tone mapped Disparity Calculation

The discontinuities in a disparity map often co-occur with the color or brightness changes in the associated camera image [DT05] as illustrated in Fig. 4.3. Therefore, using tone-mapped image pairs which contain more accurate information of the brightness and color than LDR images, helps the disparity estimation. Our results, illustrated in Table 5.2, show that some of the TMOs provide better data for stereo matching but here we propose an approach to take into account all the positive features of the different tone-mapped images. We tie together the available information from different tone-mapped image pairs. Each TMO is based on some specific image features. The probability distribution based on each disparity map (obtained from different TMOs) provides a practical platform to effectively obtain the most probable disparity value for each pixel.

5.2.1 Tone-mapped Disparity Map Calculation

We use a state-of-the-art local stereo matching technique based on cost-volume filtering [HRB+13] for all of our disparity estimations. The matching cost calculation based on the color intensity values (Ic_i) and luma gradient information ($\bigtriangledown_x I_i$) is formulated in Eq.5.1 [HRB+13]. The seven different tone-mapped image pairs are used as different inputs to our stereo matcher.

$$C_{i,d} = \alpha \cdot min[\|Ic_i - Ic'_{i-d}\|, \tau_1] + (1-\alpha) \cdot min[\|\bigtriangledown_x I_i - \bigtriangledown_x I'_{i-d}\|, \tau_2]. \tag{5.1}$$

The cost-volume entry, $C_{i,d}$, determines how well a pixel i in the left image matches the same pixel in the right image shifted by vector (disparity) d in the x direction. Here, $\bigtriangledown_x$ is the gradient operator in the x direction. For weighting the color and gradient information α is used, and τ values are truncation values. We apply this approach on different tone-mapped stereo inputs in the rest of the paper. The minimum cost disparity value (among the d disparities) is then estimated for each pixel i to be saved in f_i as the final disparity as in Eq. 5.2.

$$f_i = \underset{d}{\operatorname{argmin}}\, C_{i,d}. \tag{5.2}$$

5.2.2 Combination method

Markov Random Fields (MRFs) are being widely used in computer vision during the last two decades because of their enormous power in modeling visual perception problems [WKP13]. We model our combining problem using a pairwise MRF by defining the measurement and smoothness potentials.

To our best knowledge, this is the first time tone-mapped disparities are being combined to provide more discriminative disparity information in HDR, low-textured scenes. We apply a graphical model to the problem of fusing several disparity maps using a Markov Random Field approach for integrating the disparities. We propose a modified version of the MRF approach used in [DT05] for integrating our seven different tone-mapped disparity maps. They used a Maximum A-Posteriori (MAP) which can be mapped to the usage of graph cuts to solve an energy minimization problem (See Eq. 5.8). Our MRF formulation of the problem works with seven layers of information, one layer per TMO disparity.

The target disparity value y is estimated using the prior seven disparity information $Z = \{z_1, z_2, ..., z_7\}$ and reference guided image (here, left image) x from the likelihood function $p(y|Z, x)$. Since the stereo images are available in high resolution, this insight is used to enhance the accuracy of the disparity estimation. We used the left image as a guided image in our approach. The MRF is defined in the form of:

$$p(y|Z, x) = \frac{1}{C} \exp(-\frac{1}{2}(\psi + \phi)). \tag{5.3}$$

where, C is a normalization factor. The disparity measurement potential ψ, and the disparity smoothness potential ϕ, are calculated as follows. The measurement potential is based on a quadratic distance between the estimated disparity value and the seven

measured ones. The set of indexes for which different disparity values are available is shown by L. A constant weight of K can be placed on the depth measurements. In our calculations we used $K = 1$.

$$\psi = \sum_{i \in L} K(y - z_i)^2, L = 1, 2, ..., 7. \tag{5.4}$$

The neighboring nodes to pixel i are considered in $N(i)$ and ϕ calculates the weighted quadratic distance between neighboring disparity information. A various number of neighbors can be used depending on the purpose of the application. We used values from eight neighboring pixels. The weighting values w_{ij} determine the correspondence between two adjacent pixels, using the constant c as a penalty for smoothing the edges in the image.

$$\phi = \sum_{i} \sum_{j \in N(i)} w_{ij}(y_i - y_j)^2. \tag{5.5}$$

The weights w_{ij} are calculated from the guidance reference image x, which in our case is the left view image since we calculate the disparity for the left view.

$$w_{ij} = exp(-c \quad u_{ij}). \tag{5.6}$$

$$u_{ij} = \|x_i - x_j\|^2. \tag{5.7}$$

Now that the MRF model is defined, there are many ways for solving the optimization problem. This problem can be solved using MAP or Energy Minimization (see Eq. 5.8). In [DT05] the conjugate gradient is used for solving the MRF. Here, we minimize the energy with the help of the graph cuts using α-expansion moves [BVZ01] as formulated in Eq. 5.10, since one of the classical usages of energy minimization is for assigning labels (here, disparity values) to pixels. Minimizing the energy/cost has gained a lot of popularity especially in solving low-level vision problems such as stereo matching.

$$E(y|Z, x) = -\log p(y|Z, x). \tag{5.8}$$

Equations 5.8, 5.9 and 5.10 show the relation between the energy and likelihood and their optimization approaches, where E is the energy function [WKP13].

$$\hat{y} = \underset{y}{\text{argmax}}\{exp \sum_{p} \in (p(y|z))\}. \tag{5.9}$$

$$\hat{y} = \underset{y}{\text{argmin}}\{-\sum_{p} \ln (p(y|z))\}. \tag{5.10}$$

5.3 Evaluation and Comparison of TMOs For Stereo Matching Purpose

In this section, we discuss our results based on running the same stereo matching algorithm on various tone mapped stereo images and compare the calculated disparity maps. Our goal is to specify the most suitable TMO for stereo matching in HDR scenes.

5.3.1 Reference/Ground Truth Disparity Map

Disparity maps obtained from HDR stereo image pairs are post processed and used as our reference disparity information. We computed the HDR disparity maps by replacing the color intensity information and luma gradient values in Eq. 5.1 with radiance (R_i) and radiance gradient values ($\bigtriangledown_x R_i$), respectively [AYG14], as shown in Eq. 5.11. These HDR disparity maps are shown as *post processed disparity maps* in Fig. 5.2.

$$\begin{aligned} C_{i,d} = \alpha \cdot min[\|R_i - R'_{i-d}\|, \tau_1] + \\ (1-\alpha) \cdot min[\|\bigtriangledown_x R_i - \bigtriangledown_x R'_{i-d}\|, \tau_2]. \end{aligned} \tag{5.11}$$

In Fig. 5.2, the first two rows demonstrate the left and right views of the images which contain the low-textured background as well as tricky lighting conditions for stereo matching. The third row shows the *gradient images* corresponding to the left view, showing the lack of gradient information. In the fourth row the *matched points* from two views using the customized HDR cost-volume stereo matcher [AYG14] are shown. These disparity maps are not post processed. *The post processed disparity maps* are presented in the fifth row. In the last row *the in painted*, improved reference disparity maps are depicted using the gradient information to fill in the holes in the disparities presented in the fifth row.

We seek to get a deeper insight into the observed quality differences between the LDR and HDR matching results by comparing computed cost values in the intensity space versus HDR radiance space. Details of the cost values computed in the intensity (LDR) and radiance space (HDR) for a specific sample are provided in Fig. 4.3. Graphs (g_1), (g_2), (g_3) and (g_4) show the computed matching costs for specified pixels in e and f. In all graphs, cost values (first rows) are calculated from color/radiance information (middle row) and gradient information (bottom row) using the LDR and HDR cost computation algorithms as described before. Pixels x_1 and x'_1 in the LDR and HDR disparity maps (e and f), respectively, achieve a plausible disparity in both cases. Contrarily, x_2 and x'_2 represent the same pixel in the LDR and HDR disparity map for which a wrong disparity in LDR but a correct disparity in HDR was achieved. The top row in graph (g_3) shows how working in conventional intensity space results in very similar matching costs inside low texture areas that are very likely to cause a mismatch. The top row in graph (g_4) presents the corresponding pixel matching costs in radiance space with only one minimum among matching costs. The same effect was observed on other image pairs too. Using

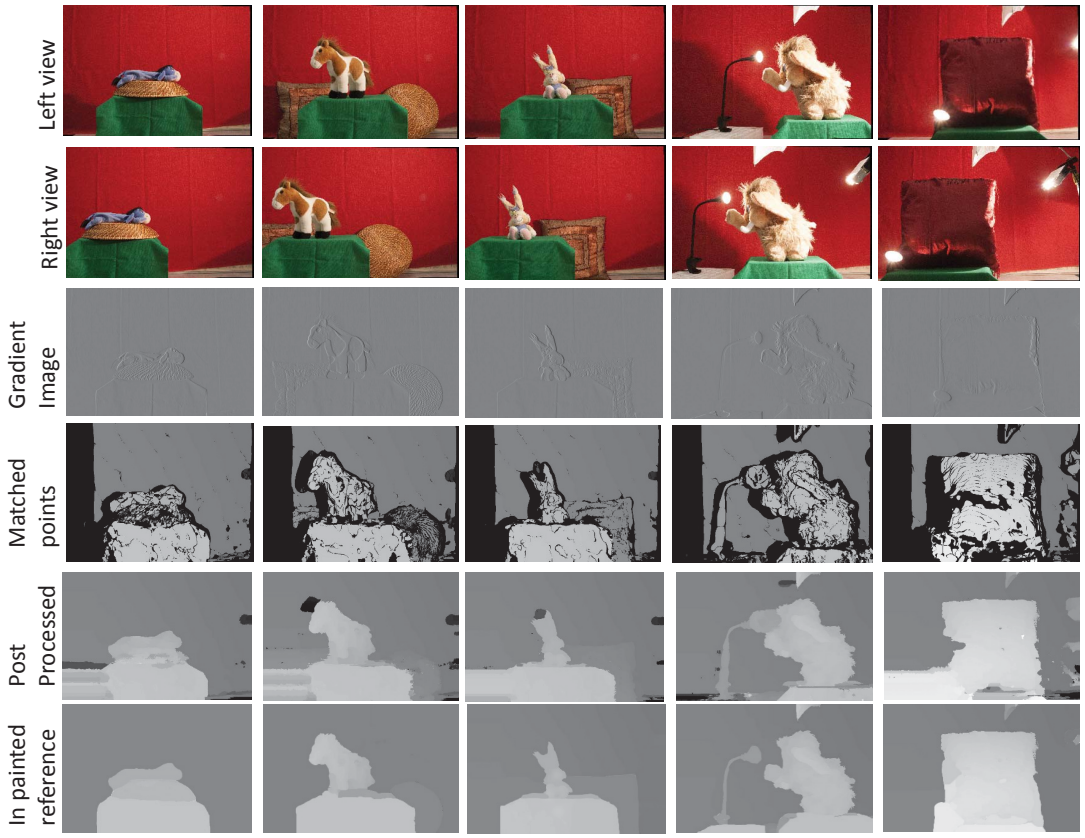


Figure 5.2: HDR stereo data set and their corresponding reference disparity maps. First two rows: the left and right views of the images, Third row: the *gradient images* corresponding to the left view, Forth row: the *matched points* from two views using the customized HDR cost-volume stereo matcher, Fifth row: *the post processed disparity maps* and, the last row: *the in painted* reference disparity maps.

scene radiance information instead of intensity values, we achieved one global minimum which is known to be a more reliable situation in optimization algorithms. Therefore, we use the HDR disparity maps (forth row in Fig. 5.2) with small modifications (last row in Fig. 5.2) as reference disparity information for our quantitative evaluation.

5.3.2 Qualitative Evaluation

Fig. 5.3 compares the disparity maps achieved from different input stereo data. The first row shows the HDR reference disparity maps calculated as described in Section 5.3.1. The following rows show the matching results on different tone-mapped stereo images as well as the LDR stereo input. The last row represents our combination approach disparities. Five HDR scenes with low-textured background are chosen for our experiments. The lighting and shadow situation in all of the scenes creates challenging situation for stereo

matching using conventional LDR images.

As it is noticeable from the results, some TMOs perform worse than the traditional LDR approach. Reinhard TMO, Durand TMO and Fattal TMO cause more artifacts and mismatched pixels than the traditional LDR matching. Our results illustrate that Gastal, Drago, He and Farbman TMOs are robust to some changes in the lighting condition and therefore achieve more informative disparity maps from stereo pairs. The combined approach outperforms all of the other calculated disparities.

5.3.3 Quantitative Evaluation

The Root Mean Square Error (RMSE) value indicating the difference between each calculated disparity map and the corresponding HDR reference disparity is shown below the results in Fig. 5.3. Table 5.2 contains the average RMSE values on all of the five data set shown in Fig. 5.2 and Fig. 5.3. Based on the information presented in the Table 5.2, we derive some conclusions:

- Not all of the TMOs perform well for stereo matching. Three out of seven tone-mapped stereo images used in our experiment resulted in more mismatched disparities compared to conventional LDR image pairs. Choosing an effective TMO for the stereo matching and 3D reconstruction purposes is a completely different task than evaluating TMOs for visual and display purposes.
- Local TMOs based on edge-aware filters perform better for stereo matching than some of the other TMOs which are based on a global curve estimation. Among the four edge-aware filtering approaches implemented and studied in this paper, the Farbman filter suits disparity estimation better than the others.
- As expected, the combined disparities outperform all of the other tone-mapped stereo matchers. This is based on the fact that the disparities with mismatched pixels in some regions, might carry valuable information in some other specific areas of the image. Our experiments show that combining the four best performing tone-mapped disparities (Gastal, Drago, He, Farbman) does not achieve as informative result as when we used all the seven disparities.

Table 5.3 shows the effectiveness of the MRF model. The simplest way to combine several results is to use average or median. We averaged our seven disparity maps and achieved an error which is very close to the LDR result. We also show the average RMSE on choosing four random tone-mapped disparities among the seven in our MRF solution. Even though combining four disparities results in less RMSE but our experiments point out the fact that using more information to combine results in more informative results. One can use more than seven tone-mapped disparities, however, both of our qualitative and quantitative evaluations illustrate that seven proposed results are discriminative. There is a trade off between the complexity and reducing the RMSE. We suggest to

Approach	Average RMSE
Reinhard	27.9986
Durand	26.1757
Fattal	24.4647
LDR	16.6513
Gastal	14.1403
Drago	11.9454
He	10.6072
Farbman	9.0224
Our Combined Approach	5.1754

Table 5.2: Our combined tone-mapped based approach in comparison to other approaches. Our results were compared to conventional LDR stereo matching introduced in [HRB$^+$13] and 7 other well-known tone-mapped stereo matching approaches. The error is calculated as the average RMSE on the 5 introduced stereo image pairs.

Combining Approach	Average RMSE
Average	16.6087
Median	15.3736
Our graph cut based combination method combining 4 random tone-mapped disparities	10.5418
Our graph cut based combination method combining 7 tone-mapped disparities	5.1754

Table 5.3: Comparison of our combined graph-cut based approach to several other combination methods.

keep the proposed results for combination below ten to be able to achieve near real-time disparity estimation. However, the question of how much improvement one could gain using the proposed fusion approach, remains open.

5.3.4 Implementation

The implementation was done using MATLAB R2014a code for each of the tone-mapping methods as well as the combinational approach and performed on a Lenovo X220 laptop with Intel Corei5 CPU (2250-second generation) and 8G of RAM. In this section we will expand on the tone-mapping operators implementation but mostly on our proposed combinational approach and graph cut solution. Our combinational method takes almost 5 seconds using our MATLAB code on the specified hardware. Further optimizations are possible to make the code faster and close to real time.

TMO Implementation

We used the available MATLAB codes from original publications for most of the TMOs used in this work. Some of the parameters of TMOs were modified as shown in Chapter 5.1.1. All of the codes were performing near real time on our stereo image pairs. A few milliseconds is the average time to run the TMOs on the hardware specified before.

Gastal, He, and Farbman TMOs were implemented by ourselves following the edge-aware filtering approach introduced in Durand TMO. The same concept of separating the image into the base and detail layers was taken into account.

Graph Cut, Alpha Expansion Implementation

For using a graph cut approach to combine few available results, there are few important concepts which need to be determined and defined properly based on the nature of the problem. Data Term, Cost Functions, Number of labels, Smoothness Term and Move Making Algorithm are key for achieving satisfactory results. What we used in this work can be replaced but based on our experiments the following definitions lead to the best outcome. For more details about the general definitions and some use case examples please refer to [BVZ01]. We used the available Graph Cut MATLAB code in the mentioned reference as a base for the standard functions. After defining the problem, data term and smoothness term correctly, one can use the open source graph cut solution to solve the optimization. The main function of our MATLAB code as well as smoothness definition can be found in the Appendix A.

Data Term:

Defining the data term is depending on the problem and is key for the optimization process. Stereo matching has been known as a labeling problem in the literature. Starting the labeling with defining a base cost volume disparity is also a common approach as discussed before in [HRB+13]). To start with, we define a normal distribution function on each pixel of each tone-mapped disparity candidate which results in a 3D cost volume disparity map. Our unary cost or our graph cut definition is the stereo matching cost. We define the data term as the accumulated summation of cost volume disparities. The number of labels that we are choosing the most optimized one from is in the range of [0-max disparity of the data set].

Smoothness Term:

The smoothness term will help finding the areas which are connected or have the same label value. We define the four neighbouring pixels as neighbours to each label. In other words, each pixel has a link to it's four neighbours. Potts model is a very effective smoothness term of penalizing neighbouring values for graph cut problems. Using Potts model, discontinuities between any pair of labels are penalized equally. This is amongst the simplest discontinuity preserving models and it is especially useful when the labels are unordered or the number of labels is small. There are variations of Potts model

using some static cues or weighting system. We use the gradient image for weighing the standard Potts model. Our smoothing function code is also available in Appendix A.

Move Making Algorithm:

Alpha expansion and Alpha-Beta swap are the two effective moves introduced in [BVZ01]. Both moves are trying to define a standard move for the minimization problem in order to find a local minimum. Alpha-expansion suits problems with one label; like stereo matching when we are seeking the best disparity label; and Alpha-Beta swap suits problems with a pair of labels. Our labels are disparity numbers and we are seeking the disparity label with minimum error and closest to the real distance between the left and right images. For this problem Alpha-expansion is a good optimization solution.

Below is a pseudo code on Alpha-expansion from [BVZ01]:

1. Start with an arbitrary labelling f
2. Set success = 0
3. For each label $\alpha \in$ L
   Find $\hat{}$f = arg min E(f$\acute{}$) among f$\acute{}$within one $\alpha - expansion$ of f
   If E($\hat{}$f) $<$ E(f), set f = $\hat{}$f and success = 1
4. If success = 1 goto 2
5. Return f

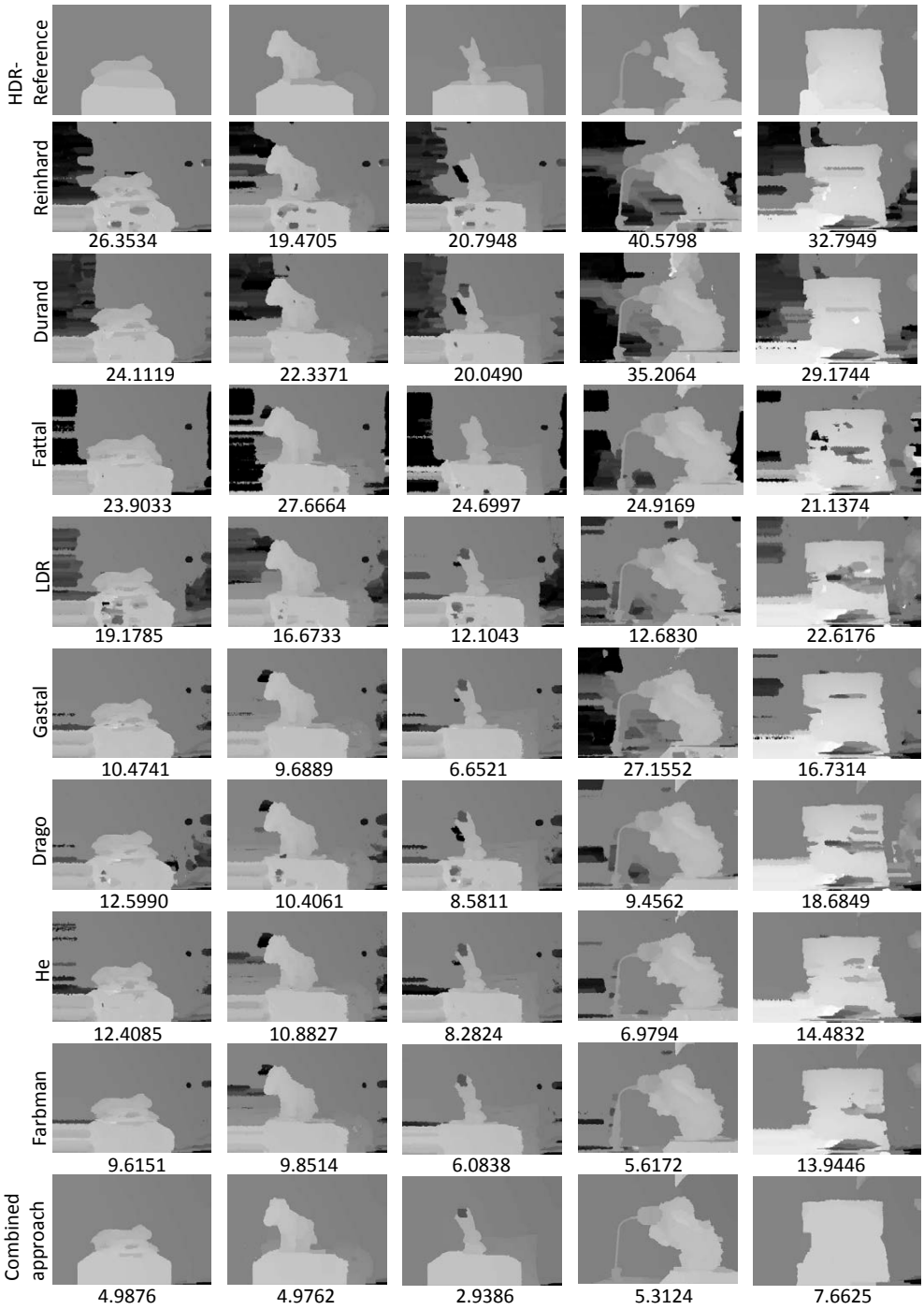


Figure 5.3: Disparity map comparison. The RMSE values are shown below each disparity. First row: HDR reference disparity maps illustrated in Fig. 5.2.Reinhard, Fattl and Durand TMOs achieve disparities with bigger error than conventional LDR stereo matching. Gastal, Drago, He and Farbman TMOs perform better than LDR stereo matching while our combined graph-cut approach outperforms all of the LDR and tone-mapped disparity maps.

CHAPTER 6

Tone Reproduction and Color Retargeting for Mesopic Vision

Not only the visible color gamut depends on the luminance level, but also our visual system works differently in different luminance conditions. Visible luminance levels in different situations are shown in Figure 2.3. The human eye is capable of perceiving a huge range of brightness but only a small part of the range is covered with current cameras and display panels. HDR displays [SHS+04] provide a greater range closer to the range perceived in real world. However, the fact that our visual color perception also changes based on the luminance level has not been taken into account [CPSZ08a]. As shown in Figure2.4, in lower luminance levels, known as mesopic and schotopic vision, visual perception differs. The fact that rod photoreceptors contribute to our vision in mesopic, as well as the change in our contrast sensitivity result in the need to compensate the content and displays to match our vision in mesopic. HDRI knowledge, tone reproduction methods and color retargeting approaches are combined to achieve this goal in this Chapter.

The human visual system works in three different modes called photopic, mesopic and scotopic vision. Photopic vision refers to our vision at day light situations (high light levels) at which only cones are responsible for our vision. As the light level falls on to a luminance of 10 $cd/m2$ [KP04a], the visual system smoothly goes from photopic vision to mesopic vision, in which both cones and rods contribute to visual perception. In the so called scotopic situations, the light level is lower than the absolute threshold of cone photoreceptors and the human vision is only mediated by rods. Photopic vision has attracted the most focus of the reserach since most of the camera and display applications in real world are in photopic and daylight environment. Mesopic and scotopic vision, are correspondinlgy attracting more attention in the industry [EVH+05], [Fai13b], [CHVW10].

In this Chapter, we first discuss some of the most highly-ranked approaches for both color

retargeting and tone reproduction with focus on mesopic vision. We also introduce a method to compensate for color perception in mesopic rendering. Our method is based on the inverse of a well-known Color Appearance Model(CAM) which simulates our mesopic visual system. The inverse of the CAM is required for color retargeting since for rendering content on dimmed displays in the dark compensating for color details is the key. CAMs **simulate** the behaviour of our visual system and inverse of CAMs **compensate** for what our visual system looses in the dark. Finally we conduct a subjective evaluation on the effectiveness of each of the color retargeting methods and furthermore the combination of both color and contrast.

Our result is very consistent with some of the prior art in this field which insists on the importance of taking advantage of both color and contrast corrections to provide the end user with an optimized visual experience [WM14b], [Rei11].

6.1 Mesopic Vision Characteristics

Our visual system reacts to changes in the spectral, spatial, and temporal properties of the illuminant across ~10 log units of dynamic range [KB86]. This is partially accomplishable because of the switching operations between two main photoreceptors in the retina, rods and cones. Rods and cones' operations overlap in some light ranges. In this chapter, we focus on mesopic vision. There is a considerable amount of literature on photopic vision because of it's huge range of applications for the camera and display industry. Mesopic is attracting more attention in recent years from computer scientists. Matching the displays to our visual system taking into account the neurological changes in the dark is an important factor which we focus on it in this work. A good understanding of the eye physiology is needed for this matter. More detailed explanation on visual perception is described in Chapter 2, Section 2.4.

In the rest of this chapter we briefly describe some of the most important perceptual effects in low light conditions and especially mesopic vision. Then, we propose compensation algorithms to help adjust display contents to human visual system. Subjective evaluations are performed and reported to compare and evaluate our methods.

Purkinje Shift: In 1825, physiologist Johannes Evangelista von Purkinje ran experiments on passing white light through a prism checking for the resulting spectrum during the early and late stages of dark adaptation. As a result of his studies, he reported that the brightest region shifted toward the shorter wavelength end of the spectrum after about 10 min in the dark. This is because of the fact that the retinal cones are maximally sensitive to a wavelength of 555 nm whilst the rods are maximally sensitive to 505 nm. In simple words, as cone vision switches over to rod vision during dark-adaptation, our visual system gets more sensitive to shorter wavelengths and blues look relatively lighter than reds [EWMU02].

6.2 Tone and Color Retargeting: Two Sides of the Same Coin

In recent years, mesopic vision and perception in the dark is getting a greater traction from both academia and industry [App, EVH+05, Fai13b]. One of the most important reasons for this traction is the advances in Virtual Reality (VR) [Occ] as well as the growth in using smart devices before bed time and the effect of late stimulation on vision and circadian rhythms [con]. The automotive industry and night time driving is another important industrial problem which is driving many research in the mesopic field [EVH+05]. Our visual color and contrast perception changes in the dark. In the mesopic range, rods and cones both contribute to color vision [PC10]. As a result, some color retargeting or reproduction is needed to match the current standards -which are mostly based on cones response [MFH+02]- to take into account the mesopic vision characteristics [RAS+15].

Contrast retargeting or tone reproduction to compensate for the contrast in the dark environment has been also investigated [WM14b], [KML03] in the last decade. The HDRI pipeline knowledge has been a reference for tone reproduction or contrast retargeting for content rendering on specific displays since the concept is close to tone mapping. The mesopic content rendering can be viewed as an application of the design showed in 2.2, in Chapter 2.

A more detailed design for the mesopic rendering solution including contrast retargeting and color retargeting was proposed by Wanat et al., shown in Figure 6.1. As illustrated in this Figure, the input to this algorithm is either of a scene-referred image or a display-referred image. In the global retargeting step, Y stands for the luminance channel, a global tone curve is applied that minimizes the contrast distortions when the image is shown on a target display; in this case a dimmed display in the dark. In this work Laplacian pyramid is used on frequency-based separated luminance levels, after a base-detailed layer separation of the image luminance is conducted. The color retargeting is taking into account the rod contribution in mesopic vission and outputs new linear RGB values. The display inverse model is used eventually to transform pixel values into sRGB color space.

In this Chapter, we use the general approach presented by Wanat et al. [WM14b] with few modifications and improvements based on our subjective evaluation results. The modifications are: first, we focus on compensation for mesopic and not night vision simulation even though the same model can be used for night vision simulation as well. Second, we show that using Shin CAM [SMYS04b] is more effective for mesopic compensation rather than Cao CAM [CPSZ08a] used by Wanat et al. and we propose an inverse of the Shin CAM for display rendering for mesopic compensation.

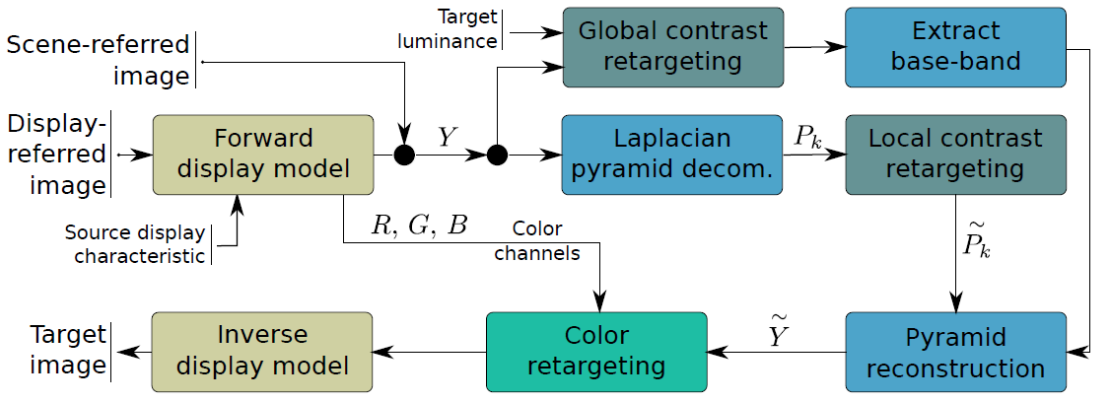


Figure 6.1: Design of the combined tone and color retargeting method for mesopic rendering [WM14b]

6.3 Color Retargeting for Mesopic Rendering Compensation

As mentioned at the beginning of this Chapter, perceptual rendering necessitates involving both a color appearance model and its inverse. The first step is to determine the luminance of the original scene-referred image. The forward color appearance model - in our case Shin's model- predicts the color perceptual attributes for a standard human observer. We then feed these attributes into the backward model (i.e. inverse Shin's model) along with the luminance of the target display and obtain the excitation values of the primaries as an output. The overview of this perceptual model is shown in Fig.6.2.

In simpler words, using the forward model, we get the values for simulating the night vision. Using this simulation, we measure the amount of required compensation in each color channel to be applied on the target dimmed display to keep the perceived color quality of the user in the dark.

In the rest of this Section, we first introduce the Shin CAM (forward model) and then the inverse of the Shin CAM. Furthermore, a detailed discussion of our subjective evaluation and methods used in the evaluation are explained.

6.3.1 Shins Color Appearance Model for Mesopic Vision

Shin et al. proposed a modified version of the Boynton two-stage model with fitting parameters to account for the rod intrusion in mesopic vision [SMYS04a]. The parameters of the model are obtained as a function of illuminance based on the asymmetric color matching experimental data. In their experiment, the observer is presented with a Munsell color chip under the mesopic condition and is asked to match the appearance of that patch with the simulated image reproduced by this model in the CRT display under photopic condition. The model is as follows:

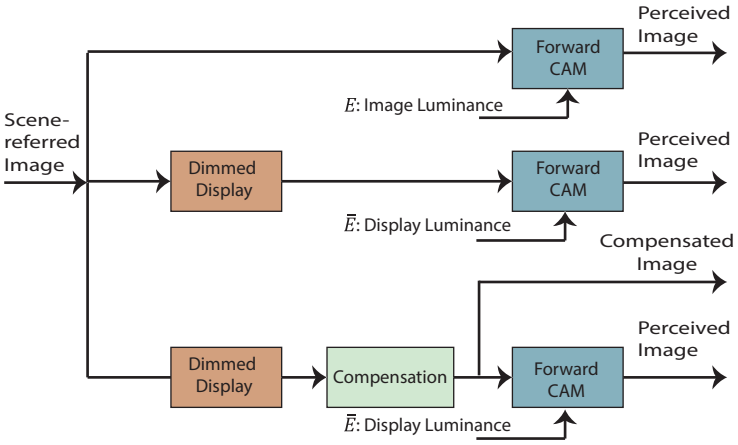


Figure 6.2: Proposed design for using inverse CAM for mesopic rendering color compensation.

1. The XYZ image (i.e. the RGB image which is transformed to the XYZ color space) is input to the model and is converted to the LMS space in the first step.

$$\begin{aligned} XYZ &= M_{rgb2xyz} \times RGB \\ LMS &= [L_p \ M_p \ S_p]^t = M_{xyz2LMS} \times XYZ \end{aligned} \tag{6.1}$$

2. The LMS signals are substituted into the opponent channel equations of the Boynton two stage model:

$$\begin{aligned} A(E) &= \alpha(E) K_w((L_p + M_p)/(L_{pw} + M_{pw})) \\ &\quad + \beta(E) K'_w (Y'/Y'_w)^{\gamma} \\ r/g(E) &= l(E)(L_p - 2M_p) + a(E)Y' \\ b/y(E) &= m(E)(L_p + M_p - S_p) + b(E)Y' \end{aligned} \tag{6.2}$$

where $A(E), r/g(E),$ and $b/y(E)$ are achromatic, red/green and blue/yellow opponent responses respectively; indices p and w indicate "photopic" and "white point" respectively; Y' represents the scotopic luminance; $\alpha(E), \beta(E), l(E), a(E), m(E),$ and $b(E)$ are the fitting parameters indicating the relative contribution of the rod's response to the opponent channels; and K_w and K'_w are the maximum responses of the luminance channel at photopic and scotopic conditions.

3. Then, opponent responses, $A(E), r/g(E)$, and $b/y(E)$, are back transformed to the XYZ space and then to the RGB space.

$$XYZ = M_{opp2xyz} \times [A(E) \ \ r/g(E) \ \ b/y(E)]^t \tag{6.3}$$

The parameters of the Shin model are selected according to Table 6.1. Functions ($\alpha(E)$, $\beta(E)$, $l(E)$, $a(E)$, $m(E)$, $b(E)$) are evaluated based on interpolation over the given points

Table 6.1: Parameters of the Shin model

Parameter	value
K_w	1
K'_w	78.4
γ	0.77

Table 6.2: Transformation matrices used in the Shin Model

Parameter	value
$M_{rgb2xyz}$	$\begin{bmatrix} 0.4124 & 0.3576 & 0.1805 \\ 0.2126 & 0.7152 & 0.0722 \\ 0.0193 & 0.1192 & 0.9505 \end{bmatrix}$
$M_{xyz2LMS}$	$\begin{bmatrix} 0.155 & 0.543 & -0.03286 \\ -0.155 & 0.457 & 0.033 \\ 0 & 0 & 1 \end{bmatrix}$
$M_{opp2xyz}$	$\begin{bmatrix} 1.008 & 2.149 & -0.212 \\ 1 & 0 & 0 \\ 1 & 0 & -1 \end{bmatrix}$

in Table 1. of [SMYS04a]. Moreover, transformation matrices used in the model are listed in Table 6.2.

6.3.2 Inverse of The Shin Color Appearance Model for Compensation

Figure 6.3 shows our general approach for the inverse model based on the target display luminance level. Our goal is to calculate an estimation of the LMS values for the amount of compensation required in LMS color space. Estimated LMS values can then be easily transformed to XYZ and eventually RGB color spaces to be rendered on the display.

To develop the inverse of the Shin Color Appearance Model, we first calculate the opponent responses of the forward model $(A(E), r/g(E), b/y(E))$ that are fed to the inverse model as shown in Figure 6.3. We assume that the compensated image based on the display luminance, E, produces the same opponent responses as the opponent responses of the forward model to make a perfect match to the perceived image at the intended luminance, E.

Then the fitting functions as introduced in Section 6.3.1, are also evaluated for the average display luminance of E. These fitting functions are indicating the relative contributions of the rod's response to the calculated opponent channels $(A(E), r/g(E), b/y(E))$. Fitting functions include $\alpha(\overline{E})$, $\beta(\overline{E})$, $a(\overline{E})$, $l(\overline{E})$, $b(\overline{E})$ and $m(\overline{E})$.

Third, the computed functions and opponent responses are substituted in the forward model using Equation 6.2 and the $\overline{LMS}$ values of the compensated image can be obtained as follows:

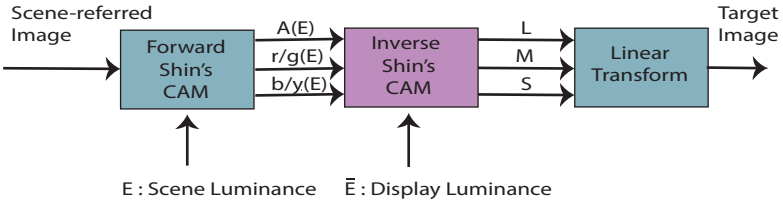


Figure 6.3: Overview of our inverse perceptual rendering framework based on Shin's model

$$\begin{aligned}
\overline{L_p} + \overline{M_p} &= ((L_{pw} + M_{pw})/(\alpha(\overline{E})K_w))\times \\
&\quad (A(E) - \beta(\overline{E})K'_w(\overline{Y'}/\overline{Y'_w})^{\gamma}) \\
\overline{L_p} - 2\overline{M_p} &= \frac{(r/g(E) - a(\overline{E}) \times \overline{Y'})}{l(\overline{E})} \\
\overline{L_p} + \overline{M_p} - \overline{S_p} &= \frac{(b/y(E) - b(\overline{E}) \times \overline{Y'})}{m(\overline{E})}
\end{aligned} \tag{6.4}$$

Fourth, the left-hand side variables of Equation 6.4 are transformed to $\overline{L_p}$, $\overline{M_p}$ and $\overline{S_p}$ using a simple linear transformation as shown below:

$$\begin{bmatrix} \overline{L_p} \\ \overline{M_p} \\ \overline{S_p} \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 \\ 1 & -2 & 0 \\ 1 & 1 & -1 \end{bmatrix}^{-1} \times \begin{bmatrix} \overline{L_p} + \overline{M_p} \\ \overline{L_p} - 2\overline{M_p} \\ \overline{L_p} + \overline{M_p} - \overline{S_p} \end{bmatrix} \tag{6.5}$$

And finally, a standard linear transformation is applied to convert the $\overline{LMS}$ values to XYZ and subsequently to RGB values.

6.4 Subjective Evaluation on Combination of Color and Tone Retargeting Approaches for Mesopic Vision

In this section, five different tone and color retargeting approaches for visual compensation in mesopic vision are analyzed, compared and evaluated on different content by naive observers in a dark room. We reviewed the tone reproduction approaches for mesopic vision as well as the color retargeting methods and show in the subjective evaluation that a suitable combination of tone and color retargeting methods is the most preferred result by observers on a dimmed display. The five chosen compensation methods as well as the unprocessed, non compensated image on a dimmed display in a dark room

were compared by 22 observers. Fifteen pairs of images processed with different tone and color compensation methods were shown to the observers. Observers were asked to choose their preferred image in terms of detail presentation in each scenario. Our pairwise comparison results show that the combination of a tone reproduction method proposed by Wanat et al. [WM14b] and a Shin-based color retargeting method proposed in this work in the previous Section achieves the best results on an OLED tablet.

The main goals of our subjective evaluation are i) to verify our hypothesis that tone and color reproduction methods need to be combined to provide the most effective compensation solution fer mesopic vision and ii) to find the most preferred combination of the tone and color retargeting approaches according to the subjects.

6.4.1 Subjective Evaluation Design

We used an OLED display tablet with a bigger color gamut than conventional LCD displays. Twenty two subjects which are not image processing experts participated in our subjective study. Fourteen male subjects and eight female subjects were located in a dark environment with only the experiment tablet in front of them for 10-15 minutes. Each observer was presented by fifteen pairs of images within an Android application and was asked to choose the image which is preferred by them. They are always presented with the same image but processed differently as shown in Figure 6.4. The observers were instructed to select one of the two displayed images as their preferred image based on the overall feeling of the color and tones.

We used eight images from different scenes such as indoor, outdoor and portraits with different color themes. Each observer compared each two method combinations (combinations of two out of six results - 5 methods and an unprocessed image) for all eight images. Observers were in the age group of 25 to 62 years old, with normal color vision and from different backgrounds. They were asked to sit in a dark room of around 2 cd/m^2 and use a Samsung Galaxy Tab AMOLED-based tablet to pick their preferred image in each case. The tablet display is a 10.5 inches with a resolution of $2560X1600$ and luminance level of 2 cd/m^2 to match the environment. During the experiment, observers were able to control their viewing angle and distance from the display.

The images for the pairwise comparison were computed beforehand using our OpenGL shader code of our proposed Shin-based color retargeting and an OpenGL shader implementation of the contrast retargeting approach introduced by [WM14a]. For other methods we used the open source available code in Matlab by the authors of the corresponding papers. The full processed images after applying the contrast and color retargeting method were organized to be randomly presented in pairs to users in our Android application.

6.4.2 Compared Methods in Our Subjective Evaluation

In all of the compared approaches except the unprocessed image and iCAM06, the tone retargeting approach introduced by Wanat et al. is used. We chose their method to

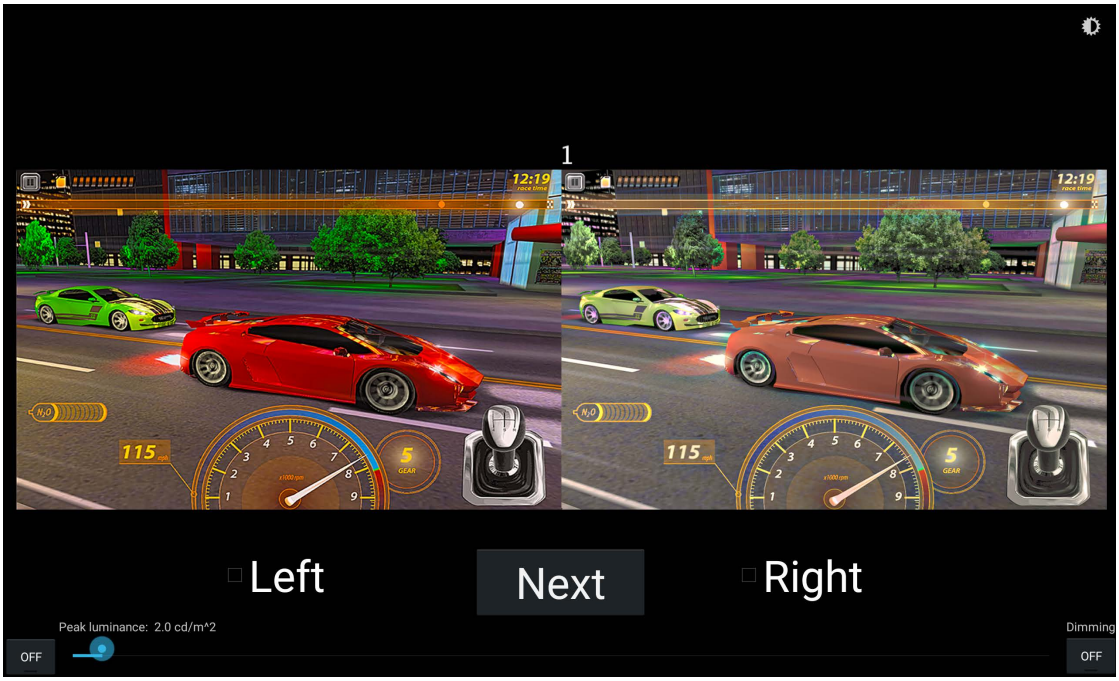


Figure 6.4: An example of our Android application used for pairwise evaluation.

retarget luminance contrast since it has been proven to work very effectively and it is one of the most recent approaches, which proved to work better compared to the well-known state of the art approaches in [WM14a]. We did not use this tone retargeting for iCAM06 since iCAM06 is a complete model of luminance and color retargeting together and we would like to compare it as is to our proposed framework. A simple overview of our framework which is a simplified version of the one introduced in [WM14a] is shown in Figure 6.5.

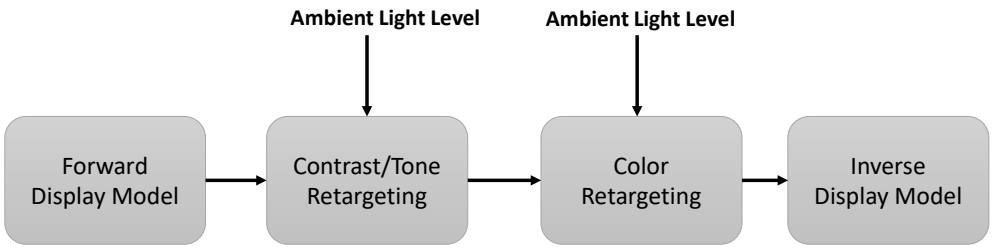


Figure 6.5: Simple overview of our combinational framework for contrast and color retargeting for compensation in mesopic vision.

Method 1: iCAM06 [KJF07a]

This approach is one of the most well-known image appearance methods in the literature. The iCAM06 tone mapping technique accounts for mesopic condition by including the rod response in its tone compression operator [KJF07a] which is summarized as follows.

1. The chromatic adapted image is input to the tone compression unit and in the first step, this input image is converted to the Hunt-Pointer-Estevez space. Then the cone responses (R'_a, G'_a, B'_a) are obtained using the cone response functions introduced by Hunt [Hun95] from the (R', G', B') inputs which is coming from the previous step.

$$\begin{aligned} R'_a &= \frac{400(F_L R'/Y_w)^p}{27.13 + (F_L R'/Y_w)^p} + 0.1 \\ G'_a &= \frac{400(F_L G'/Y_w)^p}{27.13 + (F_L G'/Y_w)^p} + 0.1 \\ B'_a &= \frac{400(F_L G'/Y_w)^p}{27.13 + (F_L G'/Y_w)^p} + 0.1 \\ F_L &= 0.2k^4(5L_A) + 0.1(1 - K^4)^2(5L_A)^{1/3} \\ k &= 1/(5L_A + 1) \end{aligned} \tag{6.6}$$

where Y_w refers to the luminance of the local adapted white image, p is a user adjustable parameter and L_A is the adaptation luminance.

2. The adapted rod response (A_S) is calculated using the Hunt model [Hun95].

$$\begin{aligned} A_s &= 3.05B_s[\frac{400(F_{LS}S/S_w)^p}{27.13 + (F_{LS}S/S_w)^p}] + 0.3 \\ F_{LS} &= 3800j^2(5L_{AS}/2.26) \\ &+ 0.2(1 - j^2)^4(5L_{AS}/2.26)^{1/6} \\ L_{AS} &= 2.26L_A \\ j &= 0.00001/[(5L_{AS}/2.26) + 0.00001] \\ B_S &= \frac{0.5}{1 + 0.3[(5L_{AS}/2.26)(S/S_w)]^{0.3}} \\ &+ \frac{0.5}{1 + 5[5L_{AS}/2.26]} \end{aligned} \tag{6.7}$$

where S and S_w are the luminance of the chromatic adapted image and the that of the reference white respectively and L_{AS} is the scotopic luminance value.

3. The tone compression output (RGB_{TC}) is computed as a linear combination of the cone responses (RGB'_a) and the rod response (A_S).

$$RGB_{TC} = RGB'_a + A_s \tag{6.8}$$

It is assumed that rod cells contribute to all cone responses with the same weights which is questionable based on the recent findings [KO11].

The processed results of eight sample images with the iCAM06 approach are shown in the second column of Fig 6.6. In our experiment we set the input parameters of this model as: maximum luminance, $max_L = 2$ (cd/m^2); overall contrast, $p = 0.7$; surround adjustment, $gamma_{value}$=1. A Matlab code which is available from the authors of this method is used for processing our eight images.

Method 2: The Wanat contrast retargeting approach combined with saturation color retargeting method [WM14b]

As discussed earlier, the luminance retargeting method proposed by Wanat and Mantiuk [WM14b] consisted of tone-curve optimization, spatial contrast processing, and hue/saturation shifts modelling.

The color retargeting model of this method is as follows:

$$\begin{aligned} \hat{R}_i &= \tilde{Y} \times (\frac{\tilde{R}_i}{\tilde{Y}})^{\frac{s(Y)}{s(\tilde{Y})}} \\ s(Y) &= \frac{Y}{(Y+\kappa)} \end{aligned} \tag{6.9}$$

where Y and $\tilde{Y}$ are the image luminance and the luminance of the tone-mapped image respectively, κ is an adjusting factor and $\tilde{R}_i$ refers to the i_{th} channel of the tone-mapped image.

While their tone mapping algorithm showed improved performance compared to many other methods [WM14b], the color retargeting method did not show significant contribution for image reproduction at low light levels [RAS+15]. The processed results of eight sample images with Wanat combined contrast and color approach are shown in the third column of Fig. 6.6.

It is worth noting that iCAM06 and Wanat and Mantiuk's luminance retargeting method take into account both tone and color reproduction, while the other two models (Cao and Shin) focus on color appearance only. Thus, the state-of-the-art tone mapping technique [WM14b] was combined with Cao and Shin's color appearance models to reproduce images at Mesopic light levels.

Method 3: The Wanat contrast retargeting approach [WM14b]

In our subjective evaluation, we also considered using only the contrast/tone reproduction method proposed by Wanat and Mantiuk, since that is introduced as their main contribution in the publication. Considering only a tone reproduction method in our study, we evaluate the relative importance of the tone reproduction algorithm compare to the combined approaches. According to the subjective evaluation done by [WM14b], their approach using a simplified color retargeting works considerably better than the state of the art approaches.

The color retargeting approach of this method is summarized in the following:
1. The image RGB values are transformed to the LMSR photoreceptor space which gives the cone and rod responses.

$$[E_L\ E_M\ E_S\ E_R]^t = M_E.[R\ G\ B]^t \tag{6.10}$$

where M_E is the corresponding transformation matrix.
2. Since rods and cones share their pathway to the visual cortex, the photoreceptor responses are combined according to the following equation:

$$\begin{aligned}\begin{bmatrix} L \\ M \\ S \end{bmatrix} &= \begin{bmatrix} 1 & 0 & 0 & \kappa_1(Y) \\ 0 & 1 & 0 & \kappa_1(Y) \\ 0 & 0 & 1 & \kappa_2(Y) \end{bmatrix} \begin{bmatrix} E_L \\ E_M \\ E_S \\ E_R \end{bmatrix} \\ &= M_c(Y) \begin{bmatrix} E_L \\ E_M \\ E_S \\ E_R \end{bmatrix}.\end{aligned} \tag{6.11}$$

The coefficients $\kappa_1(Y)$ and $\kappa_2(Y)$ are set according to the luminance level.
3. The result of retargeting for a new target luminance value $\tilde{Y}$ can be obtained by:

$$\begin{bmatrix} \tilde{R} \\ \tilde{G} \\ \tilde{B} \end{bmatrix} = \frac{\tilde{Y}}{Y}(M_c(\tilde{Y})M_E)^{-1}M_c(Y)M_E \begin{bmatrix} R \\ G \\ B \end{bmatrix} \tag{6.12}$$

The processed results of the eight sample images with Wanat contrast approach are shown in the fourth column of Fig. 6.6.

Method 4: The Wanat contrast retargeting approach combined with Cao color retargeting model [WM14b]

In this algorithm, the Cao algebraic model and it's inverse are employed in the retargeting method. This algorithm is implemented by ourselves and used for processing images as explained in [WM14b]. In this publication Wanat and Mantiuk proposed a complete Cao-based color retargeting method which was reported not to work as effective as expected and the results were not presented in their subjective evaluation results. As explained in Section 6.4.2, another color retargeting approach called saturation was considered in the evaluation. The processed results of eight sample images with the Wanat combined contrast and Cao-based color approach are shown in the fifth column of Fig. 6.6.

Method 5: The Wanat contrast retargeting approach [WM14b] Combined with our proposed Shin-based color retargeting model

This is a combined approach of one of the most recent tone retargeting approaches for mesopic vision and our proposed Shin-based color retargeting. To create an image appearance retargeting framework which works in photopic and Mesopic range for both simulation and compensation purposes, the contrast retargeting approach proposed by Wanat and Mantiuk [WM14b] is combined with the color retargeting approach proposed by us in Section 6.3. Both methods are shown to be very effective using separate subjective evaluations published in [WM14b] and [RAS+15]. In our subjective evaluation we replace the Wanat and Mantiuk color retargeting approach with the one proposed by us, keeping the same proposed architecture shown in Fig.**??**. The processed images using this combinational approach are shown in the sixth column of Fig. 6.6.

Unprocessed Images

We also show the unprocessed, original images on a dimmed display in a dark environment ($2\ cd/m^2$) as one of the compared methods. The unprocessed images are shown in the first column of Fig. 6.6. In the next section we explain the process of our pairwise comparison evaluation in more detail.

6.5 Results and Discussions

Figure 6.6 shows the output of the different approaches used in our subjective evaluation. Columns from left to right respectively show the result of applying no processing, iCAM06 (method 1) , the Wanat tone and color retargeting approach (method 2), the Wanat tone reproduction method with no color retargeting (method 3) , the Wanat tone reproduction combined with Cao color retargeting method (method 4) and the Wanat tone reproduction combined with our Shin-based color retargeting method (method 5).

We use the pairwise comparison approach with Just-Noticeable-Difference (JND) evaluation to subjectively evaluate the different combinations of our color and contrast retargeting methods. JND is known as the difference threshold, a value is changed in order for a difference to be noticeable. This approach has been used recently for subjective evaluation in the literature [WM14b, EWMU13b]. We used the Bayesian method of Silverstein and Farrell [SF01], which maximizes the probability of the pairwise comparison result account for the experiment under the Thurstone Case V assumptions. During optimization procedure, a quality value for each image is calculated maximizing the probability, modelled by the binomial distribution. Since we have 5 conditions for comparison, this Bayesian approach is suitable, which is robust to unanimous answers and common when a large number of conditions are compared, not only 2.

Figure 6.6 shows the result of our subjective evaluation calculating the JND values as defined in [EWMU13b]. The average JND value of the 6 different compared methods for overall images presented in the Figure 6.6 from left to right are 3.5, 2.08, 3.1, 4.4,

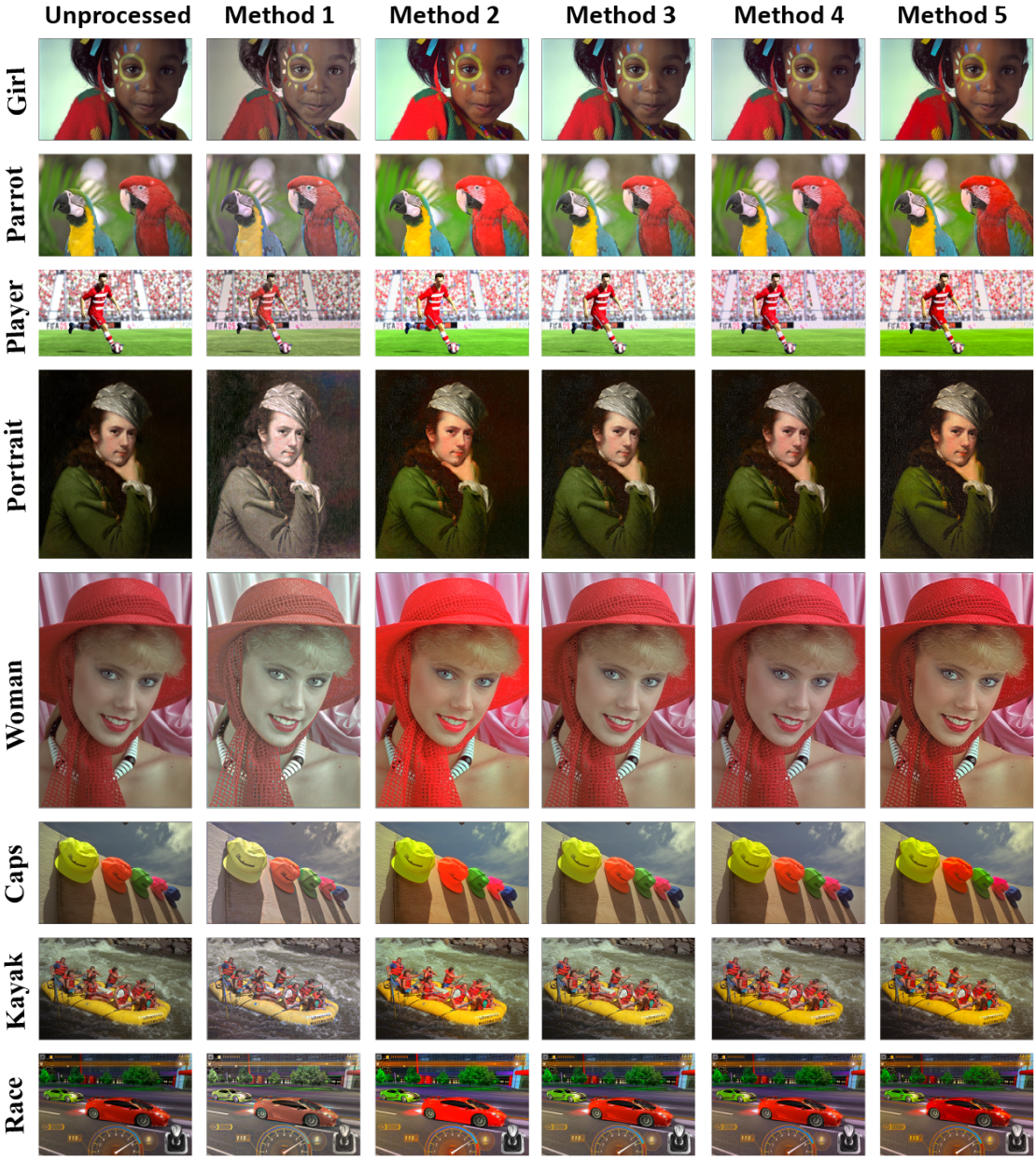


Figure 6.6: The original images and the results of different approaches applied to each image are shown. Images are processed for the luminance level of $2cd/m^2$.

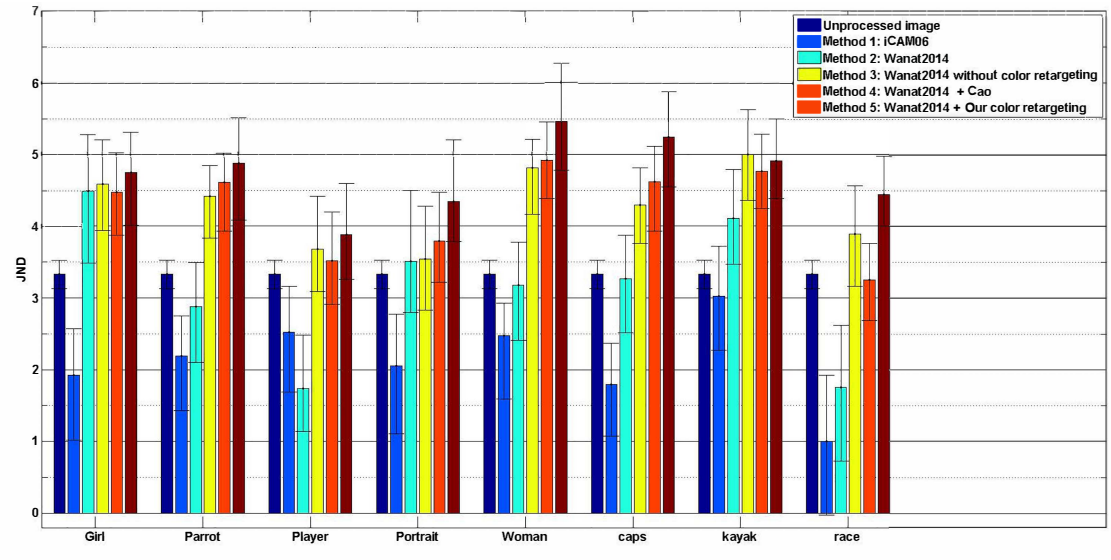


Figure 6.7

4.2 and 4.7. The absolute JND values are not meaningful and only relative difference can be used for discriminating different choices. JND is a relative value that can be also meaningful when compared with others. A method with higher JND is preferred over methods with smaller JND values, where 1 JND corresponds to 75 percent discrimination threshold. According to the average JND values, the combination of the Wanat contrast retargeting and our proposed Shin-based color retargeting provides the most preferred mesopic rendering. The Figure 6.7 represents each JND value for each of the eight images.

Not only the average JND values for method 5 is achieving the best results but also for 7 out of 8 of test images, JND of method 5 is higher that all the other approaches. The only image that the method 5 JND value is not winning the comparison is for "Kayak" image. For this image JND values for methods 3, 4 and 5 are very close to each other.

6.6 Implementation

The implementation was done using Matlab R2014a code for each of the methods as well as the combinational approach and performed on a Lenovo X220 laptop with Intel Corei5 CPU (2250-second generation) and 8GB of RAM.

We have two implementations of the Shin inverse color retargeting method proposed in Section 6.3.2, our first implementation was in Matlab R2014a and takes few seconds to process a full HD image. The color retargeting needs to be applied on every pixel of the image so we expected it to take a lot of processing power. To get a faster implementation, we did a GPU OpenGL shader implementation of the color reargeting to leverage parallel processing. The GPU implementation of our code is adding 2 milliseconds latency to

process each frame and is real-time for up to 60 frames per second full HD videos or games. This implementation is owned and used by IRYStec Software Inc. and used in a few different products being commercialized by them.

To conduct the subjective evaluation, we used the iCAM06 Matlab code available from the authors as well as Wanat [WM14b] Matlab code (owned by IRYStec Software Inc.) with no color retargeting and with their proposed saturation color retargeting. We also implemented the proposed Cao-based color retargeting method from Wanat in Matlab for method 4. None of the Matlab codes worked in real time since all of them require pixel modification. We pre processed all the images for our subjective evaluation using a simple Java based Android application to show random pairs of the images to the observers and collect their desired data.

CHAPTER 7

Conclusion and Future Work

This thesis has presented novel methods for using HDRI techniques to improve results of two important computer vision problems. HDRI techniques offer massive potential for other computer vision and graphics applications to leverage the higher range of information. In this work, we focused on two computer vision problems and introduced new techniques and algorithms to take HDRI into account to solve them. We developed combinational algorithms for tone mapped HDR stereo matching. Our results show three times decreased error in stereo matching disparity map results. Furthermore, we developed a compensation method for rendering in mesopic vision using physiological research in color and contrast perception in the dark. Our Subjective evaluation results using JND show a great preference in subjects for using our approach in a mesopic environment. The following section summarizes the particular contributions and findings.

7.1 Summary

In this work, we leverage state of the art High Dynamic Range Imaging pipeline and techniques for two important computer vision problems: stereo matching and mesopic rendering. HDR images provide a greater level of information for image processing and computer vision problems. Although the majority of the state of the art HDRI focus so far has been on capturing and displaying HDR images.

Stereo matching and 3D reconstruction approaches do not work as well in HDR scenes and low textured regions. The greater luminance information available in HDR content is taken into account to provide more informative disparity maps in challenging matching regions of stereo images such as highly exposed regions and low textured areas. The first challenge we faced in leveraging HDRI techniques for stereo matching was lack of data. Therefore, in Chapter 3 we introduced our setup for capturing an HDR stereo data set. We shared some of our best practices capturing stereo images from HDR scenes as well as successful and failed captured stereo image examples.

Later in Chapter 4 we customized one of the state of the art stereo matching cost volume approaches to work with HDR left and right images. The results were used as ground truth HDR stereo matching data. Our work in the two mentioned chapters provide the ground for the research on comparing different tone-mapped methods for achieving better disparity maps in stereo matching. In Chapter 5 we compare the results of well-known tone mapping operators used for stereo matching and propose a new combination method for combining the disparity maps calculate using different tone mapped HDR stereo pairs. To evaluate our proposed method we used subjective and objective approaches. Our results show three times reduced disparity map error in the stereo matching calculation.

As the second application for tone mapped HDR images, we describe the challenges for rendering content on dimmed displays in dark environments. Mesopic vision introduces a challenge in rendering since our visual perception changes in dark ambient lighting. In chapter 6 we compare and evaluate different contrast and color compensation methods for mesopic rendering and propose a method that was ranked the most effective and preferred by subjects in an extensive subjective evaluation.

7.2 Future Work

Despite the high quality provided by the presented stereo matching and mesopic rendering methods, there are still a few areas for investigation that can potentially provide more accurate results. This section discusses these problems and suggestions for future research. We discuss our suggestions in different sub sections based on the research area.

7.2.1 HDR Stereo Data Set

As mentioned before in Chapter 3, there is no HDR stereo data set available containing real world HDR scenes. HDR scenes introduce special challenges to stereo matching. Introducing a comprehensive HDR stereo data set as well as ground truth disparity maps requires more in depth research. A dataset such as Middlebury for HDR scenes is required to benchmark HDR stereo matching methods.

In this thesis, we set up a laboratory setting for capturing HDR stereo images with the luminance levels of max 10^3 cd/m^2. We successfully captured 12 data sets using 8 different exposures. Two ground truth generating techniques were developed in Chapter 3 to evaluate HDR stereo matching methods. What we explained and experimented in this work can be considered as a starting point for a massive research area of HDR stereo matching and tone mapped stereo matching. Below we list a few of our suggestions for future work in this field:

1. **Stereo Pairs With Higher Dynamic Range:** Capturing outdoor real world scenes with a greater range of luminance levels under different exposures will provide a valuable ground for benchmarking and research in the HDR stereo matching field. Luminance levels in the order of 10^4 cd/m^2 or 10^5 cd/m^2 containing cloudy

weather, sunny days and night time captures including spot lights will provide real world data for stereo matching.

2. **Number of Data Sets and Scene Diversity:** With the goal of providing a real world data set and moving away from controlled conditions for stereo matching, a diversity of scenes are required. Bright indoor scenes like the scenes we captured in this work, as well as low textured objects, outdoor scenes, various baselines, and including different light sources in the scenes, each introduce specific challenges for stereo matching. A diverse data set will help researchers to be able to further test the limitations of their algorithms and compare approaches. More robust methods will be developed and will help to move one step forward towards calculating depth information for real world conditions.

3. **Ground Truth Disparity Maps:** Structured light approach or manual mismatch area filling of automatic disparity estimation were both used in this work. These approaches or any other standard method for creating ground truth disparity maps can be effective for HDR scenes.

7.2.2 Tone Mapped Combinational Approach

In this section, we discuss a few possibilities for combining HDR tone mapped disparity maps.

1. **Other Tone Mapping Operators:** Each TMO works best under certain circumstances and on a specific type of input images. Our proposed combinational approach can be used with a different set of TMOs. The more diverse collection of TMOs picked, the better the result of our combinational approach will be. To choose the more diverse collection of TMOs, comparison and benchmarking of the approaches are required. We suggest research on different TMOs and evaluation of the combinational method on different TMOs and also different number of TMOs. We tried testing 3, 4, 5, 6 and eventually 7 TMOs to be combined. The combination of a few more TMOs should help with gaining more informative disparity maps.

2. **Other Combinational Methods:** As mentioned in Chapter 5 there are so many ways to combine results. The simplest approaches such as average, weighted average and median were compared to our MRF graph cut based solution proposed in this work. Using machine learning techniques such as Neural Networks, Random Forest or Deep Learning approaches can help with achieving better results. One of the challenges with using any of the mentioned learning techniques is providing the learning data set. A big enough learning data set and ground truth disparity maps are required for any of the learning methods to converge into satisfactory results.

   We cited a few state of the art methods for disparity map calculation using machine learning techniques in the mentioned Chapter. We believe that once some diverse data set along with ground truth disparities becomes available, machine learning

will be an effective solution for the combination of different disparity information available from using different TMOs on HDR stereo pairs.

7.2.3 Mesopic Rendering

Mesopic rendering on different displays which are being used in dynamic lighting conditions is of interest nowadays. The HDRI pipeline provides a great analysis of human perception in mesopic and very bright lighting conditions as well as photopic vision. Leveraging this understanding to render images differently based on the lighting condition and display back light to match the perception in the best way is considered to be a new field of research.

1. **Display Panels and Available Color Gamuts:** Color gamut and contrast range of different display panels such as LCD, OLED or quantum dot panels can be measured using available tools. However, the **perceived** color and contrast of a display is not a fixed parameter and it varies in different lighting conditions and for different people. Coming up with a way to measure and apply Contrast Sensitivity Functions (CSF) and Color Appearance Models to display rendering systems is an interesting future work.

2. **Color and Contrast Compensation in Bright Photopic Lighting Condition:** In this work we focused on mesopic vision and compensating the color and contrast loss on dimmed displays in a dark environment. A complementary research can be developed on compensating for color and contrast loss in very bright sunny environments.

APPENDIX A

Combination Method Using Graph Cuts and Alpha-Expansion

The code below shows the main MATLAB code for defining the cost volume, labels, data term, smoothness function and running graph cuts Alpha-expansion for solving the minimization problem.

```
% example of using graphcut on "elephant" data set to optimize
    result of different TMO disparity maps
flag = 1; % for elephant
%  flag = 2; % for rabbit
%  flag = 3; % for rest (horse, pillow, donkey)


d1 = imread('.\images\elephant_2lamps\ldr.jpg');
d1= d1(1:433, 1:667);
d2 = imread('.\images\elephant_2lamps\Drago.jpg');
d3 = imread('.\images\elephant_2lamps\Durand.jpg');
d4 = imread('.\images\elephant_2lamps\Fattal.jpg');
d5 = imread('.\images\elephant_2lamps\He.jpg');
d6 = imread('.\images\elephant_2lamps\Gastal.jpg');
d7 = imread('.\images\elephant_2lamps\Farbman2.jpg');



%% make a cost volume with likelihood for each pixel disparity,
     a 3D result
max_disp = 200;%different in different data sets based on the
    baseline
```

```
20
21 d1 = initialize_costvolume_tara (d1 , max_disp);
22 d2 = initialize_costvolume_tara (d2 , max_disp);
23 d3 = initialize_costvolume_tara (d3 , max_disp);
24 d4 = initialize_costvolume_tara (d4 , max_disp);
25 d5 = initialize_costvolume_tara (d5 , max_disp);
26 d6 = initialize_costvolume_tara (d6 , max_disp);
27 d7 = initialize_costvolume_tara (d7 , max_disp);
28
29
30 % combine data terms
31   Dc = d1 + d2 + d3 + d4 + d5 + d6 + d7;
32
33 % fill in the occluded area (visible only in the left view)
34 Dc(:, 1:165, :) = 0.2;
35
36
37 [n m z] = size(d1);
38
39
40 NumLabels = max_disp;
41 NumSites = n*m;
42
43 h = GCO_Create(NumSites,NumLabels);              % Create new
     object with NumSites=pixels, NumLabels=max_disp
44 Dc = reshape(Dc, m*n, z);
45 Dc = Dc';
46 Dc = scaledata (Dc, 0, 255);
47 GCO_SetDataCost(h,Dc);
48
49
50 %% Define smoothness term and neighbours, Potts model with a
     linear weight
51
52 img1 = imread('.\images\elephant_2lamps\left.jpg');
53
54 img2 = imread('.\images\elephant_2lamps\ldr.jpg');
55 img2 = img2 (1:433, 1:667);
56 N1 =  set_smoothness_tara(img1, img2);
57 img2 = imread('.\images\elephant_2lamps\Drago.jpg');
58 N2 =  set_smoothness_tara(img1, img2);
59 img2 = imread('.\images\elephant_2lamps\Durand.jpg');
60 N3 =  set_smoothness_tara(img1, img2);
```

```
61 img2 = imread('.\images\elephant_2lamps\Fattal.jpg');
62 N4 =  set_smoothness_tara(img1, img2);
63 img2 = imread('.\images\elephant_2lamps\He.jpg');
64 N5 =  set_smoothness_tara(img1, img2);
65 img2 = imread('.\images\elephant_2lamps\Gastal.jpg');
66 N6 =  set_smoothness_tara(img1, img2);
67 img2 = imread('.\images\elephant_2lamps\Farbman2.jpg');
68 N7 =  set_smoothness_tara(img1, img2);
69
70 N = N1 + N2 + N3 + N4 + N5 + N6 + N7;
71
72 GCO_SetNeighbors(h,N);
73
74 %% Alpha expansion
75 GCO_Expansion(h); % Compute optimal labeling via alpha-
       expansion
76 label = GCO_GetLabeling(h);
77
78
79 result = reshape(label, 433 ,667);
80 result = mat2gray(result);
81 result = scaling (result, flag);
82
83 figure, imshow(result);
84
85
86
87
88 [E D S] = GCO_ComputeEnergy(h  % Energy = Data Energy + Smooth
       Energy
89
90 GCO_Delete(h);
```

The code below shows the "set smoothness" MATLAB code using different variations of Potts model to solve Alpha-expansion. In Chapter 5 of this work, Potts model with a linear weighting using the gradient information from the original image is used.

```
1
2  function [NE] = set_smoothness_tara(image, disparity_map)
3      GCparam.neighborOrder = 1;
4      GCparam.useEdge = true;
5      GCparam.smoothFunc = 3;
6      GCparam.sCue = 8;
7      GCparam.linearTrunc = 40;
8      GCparam.linearModelThr = 40;
9
10     [M N d] = size(image);
11     if GCparam.useEdge == true
12         canny_edge = max(compute_color_edge(image), [], 3);
13     end
14     if d > 1
15         image = double(rgb2gray(image));
16     else
17         image = double(image);
18     end
19     tic;
20     g_image = get_gradients(image, GCparam.neighborOrder);
21     g_disp  = get_gradients(double(disparity_map), GCparam.
           neighborOrder);
22
23     switch GCparam.smoothFunc
24         case 1 % Potts model (standard).
25             cost1 = ones(1, M*N);
26             cost2 = ones(1, M*N);
27             if GCparam.neighborOrder == 2
28                 cost3 = ones(1, M*N);
29                 cost4 = ones(1, M*N);
30             end
31         case 2 % Potts model X Static Cue.
32             cost1 = (g_image(:, :, 1) <= GCparam.sCue) + 1;
33             cost2 = (g_image(:, :, 2) <= GCparam.sCue) + 1;
34             if GCparam.neighborOrder == 2
35                 cost3 = (g_image(:, :, 3) <= GCparam.sCue) + 1;
36                 cost4 = (g_image(:, :, 4) <= GCparam.sCue) + 1;
37             end
38         case 3 % Potts model X linear weight.
39             g_image(g_image < GCparam.sCue) = 0;
```

```
40              cost1 = max(GCparam.linearModelThr - g_image(:, :,
                    1), 1);
41              cost2 = max(GCparam.linearModelThr - g_image(:, :,
                    2), 1);
42
43              if GCparam.neighborOrder == 2
44                  cost3 = max(GCparam.linearModelThr - g_image(:,
                        :, 3), 1);
45                  cost4 = max(GCparam.linearModelThr - g_image(:,
                        :, 4), 1);
46              end
47          case 4 % Linear Truncated Model (standard).
48              cost1 = min(g_disp(:, :, 1), GCparam.linearTrunc);
49              cost2 = min(g_disp(:, :, 2), GCparam.linearTrunc);
50              if GCparam.neighborOrder == 2
51                  cost3 = min(g_disp(:, :, 3), GCparam.
                        linearTrunc);
52                  cost4 = min(g_disp(:, :, 4), GCparam.
                        linearTrunc);
53              end
54          case 5 % Linear Truncation X Static Cue.
55              cost1 = min(g_disp(:, :, 1), GCparam.linearTrunc).*
                    ...
56                                  (g_image(:, :, 1) <= GCparam.
                                        sCue) + 1;
57              cost2 = min(g_disp(:, :, 2), GCparam.linearTrunc).*
                    ...
58                                  (g_image(:, :, 2) <= GCparam.
                                        sCue) + 1;
59              if GCparam.neighborOrder == 2
60                  cost3 = min(g_disp(:, :, 3), GCparam.
                        linearTrunc).* ...
61                                  (g_image(:, :, 3) <= GCparam.
                                        sCue) + 1;
62                  cost4 = min(g_disp(:, :, 4), GCparam.
                        linearTrunc).* ...
63                                  (g_image(:, :, 4) <= GCparam.
                                        sCue) + 1;
64              end
65          case 6 % Linear Truncation X linear weight.
66              g_image(g_image < GCparam.sCue) = 0;
67
68              cost1 = min(g_disp(:, :, 1), GCparam.linearTrunc).*
```

```
                    ...
69                          max(GCparam.linearModelThr - g_image(:,
                                :, 1), 1);
70              cost2 = min(g_disp(:, :, 2), GCparam.linearTrunc).*
                    ...
71                          max(GCparam.linearModelThr - g_image(:,
                                :, 2), 1);
72              if GCparam.neighborOrder == 2
73                  cost3 = min(g_disp(:, :, 3), GCparam.
                        linearTrunc).* ...
74                          max(GCparam.linearModelThr - g_image(:,
                                :, 3), 1);
75                  cost4 = min(g_disp(:, :, 4), GCparam.
                        linearTrunc).* ...
76                          max(GCparam.linearModelThr - g_image(:,
                                :, 4), 1);
77              end
78      end
79
80      % Do indexing stuffs.
81      index1 = M+1:M*N;
82      index2 = 2:M*N;
83      if GCparam.neighborOrder == 2
84          index3 = M+2:M*N;
85          index4 = M:M*N;
86          index = [index1 index2 index3 index4];
87      else
88          index = [index1 index2];
89      end
90
91      % Re-weight edges using canny edge information.
92      if GCparam.useEdge == true
93          cost3 = ((1 - double(canny_edge))*(1-eps) + eps);
94          cost1 = cost1.*cost3;
95          cost2 = cost2.*cost3;
96      end
97
98      % Create a sparse neighbor graph.
99      if GCparam.neighborOrder == 1
100         NE = sparse([1:length(index1) 1:length(index2)], index,
                [cost1(index1) cost2(index2)], M*N, M*N);
101     else
102         NE = sparse([1:length(index1) 1:length(index2) 1:length
```

```
            (index3) 1:length(index4)], index, ...
103             [cost1(index1) cost2(index2) cost3(index3) cost4(
                    index4)], M*N, M*N);
104     end
105     toc;
106 end
107
108 function g_image = get_gradients(image, order)
109     [M N] = size(image);
110     g_image = zeros(M, N, order*2);
111     g_image(:, :, 1) = abs(conv2(image, [-1 1], 'same'));
112     g_image(:, :, 2) = abs(conv2(image, [-1 1]', 'same'));
113
114     if order == 2
115         g_image(:, :, 3) = abs(conv2(image, [-1 0;0 1], 'same')
                );
116         g_image(:, :, 4) = abs(conv2(image, [0 -1;1 0], 'same')
                );
117     end
118
119 end
120
121 function edge_image = compute_color_edge(image)
122     [M, N, ~] = size(image);
123     edge_image = zeros(M, N, 3);
124     for k = 1:size(image, 3)
125         edge_image(:, :, k) = edge(image(:, :, k));
126     end
127 end
```

List of Figures

List of Tables

Bibliography

[AKCG14] T. Akhavan, C. Kappeler, J. Cho, and M. Gelautz. Stereo HDR disparity map computation using structured light. In *HDRi2014 - Second International Conference and SME Workshop on HDR imaging*, 2014.

[AM10] Tara Akhavan and Mohsen Ebrahimi Moghaddam. A new combining learning method for color constancy. In *2nd International Conference on Image Processing Theory Tools and Applications*, pages 421–425, 2010.

[AM11] T. Akhavan and M. Moghaddam. A color constancy method using fuzzy measures and integrals. *Optical Review*, 18(3):273–283, 2011.

[Ans02] Stuart Anstis. The purkinje rod-cone shift as a function of luminance and retinal eccentricity. *Vision Research*, 42(12):2485–2491, 2002.

[App] Perception in low light environment. `http://www.theverge.com/2016/3/29/11326194/apple-night-shift-blue-light-sleep`. Accessed: 2016-05-02.

[ASC+14] Tunç Ozan Aydin, Nikolce Stefanoski, Simone Croci, Markus Gross, and Aljoscha Smolic. Temporally coherent local tone mapping of hdr video. *ACM Trans. Graph.*, 33(6):196:1–196:13, 2014.

[AYG13] T. Akhavan, H. Yoo, and M. Gelautz. A framework for HDR stereo matching using multi-exposed images. In *HDRi2013 - First International Conference and SME Workshop on HDR imaging*, 2013.

[AYG14] T. Akhavan, H. Yoo, and M. Gelautz. Evaluation of LDR, Tone Mapped and HDR Stereo Matching Using Cost-volume Filtering Approach. In *22th European Signal Processing Conference (EUSIPCO 2014)*, pages 1–6, 2014.

[BBM87] Robert C. Bolles, H. Harlyn Baker, and David H. Marimont. Epipolar-plane image analysis: An approach to determining structure from motion. *International Journal of Computer Vision*, 1(1):7–55, 1987.

[BC10] M. Bleyer and S. Chambon. Does color really help in dense stereo matching? In *3DPVT*, pages 1–8, 2010.

[BLV+12] J. Bonnard, C. Loscos, G. Valette, J. Nourrit, and L. Lucas. High-dynamic range video acquisition with a multiview camera. In *SPIE*, volume 8436, pages 84360A–1–84360A–11, 2012.

[Bra00] G. Bradski. The OpenCV Library. *Dr. Dobb's Journal of Software Tools*, 2000.

[BRG+14] M. Bätz, T. Richter, J. Garbas, A. Papst, J. Seiler, and A. Kaup. High dynamic range video reconstruction from a stereo camera setup. *Signal Processing: Image Communication*, 29:191–202, 2014.

[BRR11] M. Bleyer, C. Rhemann, and C. Rother. PatchMatch stereo - stereo matching with slanted support windows. In *BMVC*, pages 1–11, 2011.

[BSKZ14] Karlheinz Blankenbach, Andrej Sycev, Sascha Kurbatfinski, and Martin Zobl. Optimizing and evaluating new automotive hmi image enhancement algorithms under bright light conditions using display reflectance characteristics. *Journal of the Society for Information Display*, 22(5):267–279, 2014.

[BVZ01] Yuri Boykov, Olga Veksler, and Ramin Zabih. Fast approximate energy minimization via graph cuts. *PAMI*, 23(11):1222–1239, 2001.

[CHVW10] Cynthia L Clark, Joseph L Hardy, Vicki J Volbrecht, and John S Werner. Scotopic spatiotemporal sensitivity differences between young and old adults. *Ophthalmic and Physiological Optics*, 30(4):339–350, 2010.

[CL08] J. Choi and K. M. Lee. Accurate stereo matching using pixel response function. In *Workshop on Image Processing and Image Understanding (IPIU)*, 2008.

[con] Using mobile phones before bed. `http://www.dailymail.co.uk/sciencetech/article-2194806/Using-mobile-phones-tablets-bed-affecting-sleep-warn-scientists.html`. Accessed: 2016-02-29.

[CPSZ08a] Dingcai Cao, Joel Pokorny, Vivianne C Smith, and Andrew J Zele. Rod contributions to color perception: linear with rod contrast. *Vision research*, 48(26):2586–2592, 2008.

[CPSZ08b] Dingcai Cao, Joel Pokorny, Vivianne C Smith, and Andrew J Zele. Rod contributions to color perception: linear with rod contrast. *Vision research*, 48(26):2586–2592, 2008.

[CS09] James J Clark and Sandra Skaff. A spectral theory of color perception. *JOSA A*, 26(12):2488–2502, 2009.

[ČWNA06] M. Čadík, M. Wimmer, L. Neumann, and A. Artusi. Image attributes and quality for evaluation of tone mapping operators. In *14th Pacific Conference on Computer Graphics and Applications*, pages 35–44, 2006.

[DD00] Frédo Durand and Julie Dorsey. *Interactive tone mapping.* Springer, 2000.

[DD02] F. Durand and J. Dorsey. Fast bilateral filtering for the display of High-dynamic-range images. *ACM TOG*, 21:257–266, 2002.

[Dee05] Michael F. Deering. A photon accurate model of the human eye. *ACM Trans. Graph.*, 24(3):649–658, July 2005.

[Dev02] Kate Devlin. A review of tone reproduction techniques. *Computer Science, University of Bristol, Tech. Rep. CSTR-02-005*, 2002.

[DM97] P. Debevec and J. Malik. Recovering high dynamic range radiance maps from photographs. In *ACM SIGGRAPH*, pages 369–378, 1997.

[DMAC03] F. Drago, K. Myszkowski, T. Annen, and N. Chiba. Adaptive logarithmic mapping for displaying high contrast scenes. *Computer Graphics Forum*, 22:419–426, 2003.

[DMVC11] L. De-Maeztu, A. Villanueva, and R. Cabeza. Stereo matching using gradient similarity and locally adaptive support-weight. *Pattern Recognition Letters*, 32(13):1643–1651, 2011.

[DT05] J. Diebel and S. Thrun. An application of markov random fields to range sensing. In *Proceedings of Conference on Neural Information Processing Systems (NIPS)*, pages 291–298, 2005.

[EVH+05] Marjukka Eloholma, M Viikari, L Halonen, H Walkey, T Goodman, JWAM Alferdinck, A Freiding, P Bodrogi, and G Várady. Mesopic models—from brightness matching to visual performance in night-time driving: a review. *Lighting Research and Technology*, 37(2):155–173, 2005.

[EWMU02] Gabriel Eilertsen, Robert Wanat, Rafal K. Mantiuk, and Jonas Unger. The Purkinje rod-cone shift as a function of luminance and retinal eccentricity. *Vision Research*, 42(22):2485–2491, 2002.

[EWMU13a] Gabriel Eilertsen, Robert Wanat, Rafal K. Mantiuk, and Jonas Unger. Evaluation of tone mapping operators for hdr-video. *Comput. Graph. Forum*, 32(7):275–284, 2013.

[EWMU13b] Gabriel Eilertsen, Robert Wanat, Rafal K. Mantiuk, and Jonas Unger. Evaluation of Tone Mapping Operators for HDR-Video. *Computer Graphics Forum*, 2013.

[eye] How to combat eye fatigue right now. http://www.eizoglobal.com/library/Guide_to_preventing_eye_fatigue.pdf. Accessed: 2016-02-29.

[Fai13a] Mark D Fairchild. *Color appearance models*. John Wiley & Sons, 2013.

[Fai13b] Mark D Fairchild. *Color appearance models*. John Wiley & Sons, 2013.

[FFLS08] Zeev Farbman, Raanan Fattal, Dani Lischinski, and Richard Szeliski. Edge-preserving decompositions for multi-scale tone and detail manipulation. *ACM TOG*, 27(3):67:1–67:10, 2008.

[FLW02] Raanan Fattal, Dani Lischinski, and Michael Werman. Gradient domain high dynamic range compression. SIGGRAPH, pages 249–256, 2002.

[FPSG96a] James A Ferwerda, Sumanta N Pattanaik, Peter Shirley, and Donald P Greenberg. A model of visual adaptation for realistic image synthesis. In *Proceedings of the 23rd annual conference on Computer graphics and interactive techniques*, pages 249–258. ACM, 1996.

[FPSG96b] James A Ferwerda, Sumanta N Pattanaik, Peter Shirley, and Donald P Greenberg. A model of visual adaptation for realistic image synthesis. In *Proceedings of the 23rd annual conference on Computer graphics and interactive techniques*, pages 249–258. ACM, 1996.

[GO11] Eduardo S. L. Gastal and Manuel M. Oliveira. Domain transform for edge-aware image and video processing. *ACM TOG*, 30(4):69:1–69:12, 2011.

[HRB+13] A. Hosni, C. Rhemann, M. Bleyer, C. Rother, and M. Gelautz. Fast cost-volume filtering for visual correspondence and beyond. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 35(2):504–511, 2013.

[HS07] H. Hirschmüller and D. Scharstein. Evaluation of cost functions for stereo matching. In *CVPR*, pages 1 – 8, 2007.

[HST10] K. He, J. Sun, and X. Tang. Guided image filtering. In *ECCV*, pages 1–14, 2010.

[Hun91] R. W. G. Hunt. Revised colour-appearance model for related and unrelated colours. *Color Research & Application*, 16(3):146–165, 1991.

[Hun94a] RWG Hunt. An improved predictor of colourfulness in a model of colour vision. *Color Research & Application*, 19(1):23–26, 1994.

[Hun94b] RWG Hunt. An improved predictor of colourfulness in a model of colour vision. *Color Research & Application*, 19(1):23–26, 1994.

[Hun95] Robert William Gainer Hunt. *The reproduction of colour*. Fountain Press, England, 1995.

[IKH+11] Shahram Izadi, David Kim, Otmar Hilliges, David Molyneaux, Richard Newcombe, Pushmeet Kohli, Jamie Shotton, Steve Hodges, Dustin Freeman, Andrew Davison, and Andrew Fitzgibbon. Kinectfusion: Real-time 3d reconstruction and interaction using a moving depth camera. 2011.

[KB86] L. Kaufman J. Thomas K. Boff. Sensitivity to light. *Handbook of Perception and Human Performance*, 1, 1986.

[KJF07a] Jiangtao Kuang, Garrett M Johnson, and Mark D Fairchild. icam06: A refined image appearance model for hdr image rendering. *Journal of Visual Communication and Image Representation*, 18(5):406–414, 2007.

[KJF07b] Jiangtao Kuang, Garrett M Johnson, and Mark D Fairchild. icam06: A refined image appearance model for hdr image rendering. *Journal of Visual Communication and Image Representation*, 18(5):406–414, 2007.

[KML03] Youngshin Kwak, Lindsay W MacDonald, and M Ronnier Luo. Prediction of lightness in mesopic vision. In *Color and Imaging Conference*, volume 2003, pages 301–307. Society for Imaging Science and Technology, 2003.

[KMS05] Grzegorz Krawczyk, Karol Myszkowski, and Hans-Peter Seidel. Perceptual effects in real-time tone mapping. In *Proceedings of the 21st spring conference on Computer graphics*, pages 195–202. ACM, 2005.

[KO11] Adam G Kirk and James F O'Brien. Perceptually based tone mapping for low-light conditions. *ACM Trans. Graph.*, 30(4):42, 2011.

[Kon97] Kurt Konolige. Small vision system: Hardware and implementation. In *Proc. of the Intl. Symp. of Robotics Research (ISRR*, pages 111–116, 1997.

[KP04a] Saad Masood Khan and Sumanta N Pattanaik. Modeling blue shift in moonlit scenes by rod cone interaction. *Journal of VISION*, 4(8):316a, 2004.

[KP04b] Saad Masood Khan and Sumanta N Pattanaik. Modeling blue shift in moonlit scenes by rod cone interaction. *Journal of VISION*, 4(8), 2004.

[Kul76] J.J. Kulikowski. Effective contrast constancy and linearity of contrast sensation. *Vision Research*, 16(12):1419–1431, 1976.

[KWK09] Min H Kim, Tim Weyrich, and Jan Kautz. Modeling human color perception under extended luminance levels. In *ACM Transactions on Graphics (TOG)*, volume 28, page 27. ACM, 2009.

[LC09] Huei-Yung Lin and Wei-Zhe Chang. High dynamic range imaging for stereoscopic scene representation. In *ICIP*, pages 4305–4308. IEEE, 2009.

[LCTS05] P. Ledda, A. Chalmers, T. Troscianko, and H. Seetzen. Evaluation of tone mapping operators using a high dynamic range display. In *ACM SIGGRAPH*, pages 640–648, 2005.

[Li09] Stan Z. Li. *Markov Random Field Modeling in Image Analysis*. Springer, 3 edition, 2009.

[LJDE11] F. Lu, X. Ji, Q. Dai, and G. Er. Multi-view stereo reconstruction with high dynamic range texture. In *ACCV*, pages 412–425, 2011.

[McC07] John J McCann. Colors in dim illumination and candlelight. In *Color and Imaging Conference*, volume 2007, pages 313–318. Society for Imaging Science and Technology, 2007.

[MDNW12] Zicong Mai, Colin Doutre, Panos Nasiopoulos, and Rabab Kreidieh Ward. Rendering 3-d high dynamic range images: Subjective evaluation of tone-mapping methods and preferred 3-d image attributes. *J. Sel. Topics Signal Processing*, 6(5):597–610, 2012.

[MFH+02] Nathan Moroney, Mark D Fairchild, Robert WG Hunt, Changjun Li, M Ronnier Luo, and Todd Newman. The ciecam02 color appearance model. In *Color and Imaging Conference*, volume 2002, pages 23–27. Society for Imaging Science and Technology, 2002.

[MS15] Myszkowski K. Mantiuk, R. K. and H.-P Seidel. High dynamic range imaging. pages 1–42, 2015.

[MSTK09] Michihiro Mikamo, Marcos Slomp, Toru Tamaki, and Kazufumi Kaneda. A tone reproduction operator accounting for mesopic vision. In *ACM SIGGRAPH ASIA 2009 Posters*, page 41. ACM, 2009.

[MWDG13] Belen Masia, Gordon Wetzstein, Piotr Didyk, and Diego Gutierrez. A survey on computational displays: Pushing the boundaries of optics, computation, and perception. *Computers & Graphics*, 37(8):1012–1038, 2013.

[Occ] Perception in low light environment. `https://www.oculus.com/en-us/connect/`. Accessed: 2016-02-29.

[PC10] Joel Pokorny and Dingcai Cao. Rod and cone contributions to mesopic vision. *Proceedings of CIE 2010 Lighting Quality & Energy Efficiency*, pages 9–20, 2010.

[PDKK03] P. Paclik, R. P. W. Duin, G. M. P. Van Kempen, and R. Kohlus. Segmentation of multi-spectral images using the combined classifier approach. *Image and Vision Computing*, 21(6):473–482, 2003.

[PFFG98] Sumanta N Pattanaik, James A Ferwerda, Mark D Fairchild, and Donald P Greenberg. A multiscale model of adaptation and spatial vision for realistic image display. In *Proceedings of the 25th annual conference on Computer graphics and interactive techniques*, pages 287–298. ACM, 1998.

[RAS+15] Mehdi Rezagholizadeh, Tara Akhavan, Afsoon Soudi, Hannes Kaufmann, and James J Clark. A retargeting approach for mesopic vision: Simulation and compensation. *Journal of Imaging Science and Technology*, 2015.

[RC13] Mehdi Rezagholizadeh and James J Clark. Maximum entropy spectral modeling approach to mesopic tone mapping. In *Color and Imaging Conference*, volume 2013, pages 154–159. Society for Imaging Science and Technology, 2013.

[Rei11] E. Reinhard. Tone reproduction and color appearance modeling: two sides of the same coin? In *Final Program and Proceedings - IS & T/SID Color Imaging Conference*, pages 171–176, 2011.

[RLMA13] R. Ramirez, C. Loscos, Ignacio M., and . Artusi. Patch-based registration for auto-stereoscopic HDR content creation . In *HDRi2013 - First International Conference and SME Workshop on HDR imaging*, 2013.

[Roe98] J.E. Roeckelein. *Dictionary of Theories, Laws, and Concepts in Psychology*. Greenwood Press, 1998.

[ROLMA15] Raissel Ramirez Orozco, Céline Loscos, Ignacio Martin, and Alessandro Artusi. Multiscopic hdr image sequence generation. In *Winter School of Computer Graphics*, 2015.

[RR98] Robert W Rodieck and Robert W Rodieck. *The first steps in seeing*. Sinauer Associates Sunderland, 1998.

[RSSF02] Erik Reinhard, Michael Stark, Peter Shirley, and James Ferwerda. Photographic tone reproduction for digital images. SIGGRAPH, pages 267–276, 2002.

[Rüf11] D. Rüfenacht. Stereoscopic high dynamic range video. Master's thesis, EPFL, Lausanne, Switzerland, 2011.

[RWPD05] E. Reinhard, G. Ward, S. Pattanaik, and P. Debevec. *High Dynamic Range Imaging: Acquisition, Display, and Image-based Lighting*. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 2005.

[Sch07] D. Scharstein. Learning conditional random fields for stereo. In *CVPR*, pages 1 – 8, 2007.

[SDBRC12] E. Selmanovic, K. Debattista, T. Bashford-Rogers, and A. Chalmers. Backwards compatible JPEG stereoscopic high dynamic range imaging. In *TPCG*, pages 1–8, 2012.

[SDBRC13] E. Selmanović, K. Debattista, T. Bashford-Rogers, and A. Chalmers. Generating stereoscopic HDR smages using HDR-LDR image pairs. *ACM Trans. Appl. Percept.*, 10(1):3:1–3:18, 2013.

[SF01] D. Amnon Silverstein and Joyce E. Farrell. Efficient method for paired comparison. *Journal of Electronic Imaging*, 10(2):394–398, 2001.

[SFPL10] Joaquim Salvi, Sergio Fernandez, Tomislav Pribanic, and Xavier Llado. A state of the art in structured light patterns for surface profilometry. *Pattern Recognition*, 43(8):2666–2680, 2010.

[SHS+04] H. Seetzen, W. Heidrich, W. Stuerzlinger, G. Ward, L. Whitehead, M. Trentacoste, A. Ghosh, and A. Vorozcovs. High dynamic range display systems. In *ACM SIGGRAPH*, pages 760–768, 2004.

[SKCP11] Donghwa Shin, Younghyun Kim, Naehyuck Chang, and Massoud Pedram. Dynamic voltage scaling of oled displays. In *Design Automation Conference (DAC), 2011 48th ACM/EDAC/IEEE*, pages 53–58. IEEE, 2011.

[SMW10] N. Sun, H. Mansour, and R. K. Ward. HDR image construction from multi-exposed stereo LDR images. In *ICIP*, pages 2973–2976, 2010.

[SMYS04a] JaeChul Shin, Naoki Matsuki, Hirohisa Yaguchi, and Satoshi Shioiri. A color appearance model applicable in mesopic vision. *Optical review*, 11(4):272–278, 2004.

[SMYS04b] JaeChul Shin, Naoki Matsuki, Hirohisa Yaguchi, and Satoshi Shioiri. A color appearance model applicable in mesopic vision. *Optical review*, 11(4):272–278, 2004.

[SS98] Daniel Scharstein and Richard Szeliski. Stereo matching with nonlinear diffusion. *International Journal of Computer Vision*, 28(2):155–174, 1998.

[SS02a] D. Scharstein and R. Szeliski. A taxonomy and evaluation of dense two-frame stereo correspondence algorithms. *Int. J. Comput. Vision*, 47(1-3):7–42, 2002.

[SS02b] Daniel Scharstein and Richard Szeliski. A Taxonomy and Evaluation of Dense Two-Frame Stereo Correspondence Algorithms. 47:7–42, 2002.

[SS03a] D. Scharstein and R. Szeliski. High-accuracy stereo depth maps using structured light. In *Proceedings of the 2003 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, CVPR'03, pages 195–202, 2003.

[SS03b] Daniel Scharstein and Richard Szeliski. High-accuracy stereo depth maps using structured light. In *CVPR*, pages 195–202, 2003.

[SS03c] Daniel Scharstein and Richard Szeliski. High-accuracy stereo depth maps using structured light. In *CVPR*, pages 195–202, 2003.

[STDT09] Sebastian Schuon, Christian Theobalt, James Davis, and Sebastian Thrun. Lidarboost: Depth superresolution for tof 3d shape scanning. In *CVPR*, pages 343–350, 2009.

[Sze10] Richard Szeliski. *Computer Vision: Algorithms and Applications*. Springer-Verlag New York, Inc., New York, NY, USA, 1st edition, 2010.

[TKS06] A. Troccoli, S. B. Kang, and S. M. Seitz. Multi-view multi-exposure stereo. In *3DPVT*, pages 861–868, 2006.

[Wan95] B. A. Wandell. Foundations of visi*on*. Sinauer Associates, Inc., Sunderland, MA, US, 1995.

[War01] Greg Ward. High dynamic range imaging. In *Color and Imaging Conference*, volume 2001, pages 9–16. Society for Imaging Science and Technology, 2001.

[WHM+05] Helen C Walkey, Antony Hurden, Ian R Moorhead, Julie AF Taylor, John L Barbur, and J Alister Harlow. Effective contrast of colored stimuli in the mesopic range: a metric for perceived contrast based on achromatic luminance contrast. *JOSA A*, 22(1):17–28, 2005.

[WKP13] Chaohui Wang, Nikos Komodakis, and Nikos Paragios. Markov random field modeling, inference & learning in computer vision & image understanding: A survey. *Comput. Vis. Image Underst.*, 117(11):1610–1627, 2013.

[WM14a] Robert Wanat and Rafał K Mantiuk. A comparison of night vision simulation methods for video. In *Proceedings of the 11th European Conference on Visual Media Production*, page 16. ACM, 2014.

[WM14b] Robert Wanat and Rafał K Mantiuk. Simulating and compensating changes in appearance between day and night vision. *Proceedings of SIGGRAPH 2014*, 33:147, 2014.

[WM14c] Robert Wanat and Rafał K Mantiuk. Simulating and compensating changes in appearance between day and night vision. *Proceedings of SIGGRAPH 2014*, 33:147, 2014.

[WT14] C. Wang and C. Tu. A multi-exposure images fusion approach for very large dynamic range scenes. *International Journal of Signal Processing, Image Processing and Pattern Recognition*, 7(5):217–228, 2014.

[YBMS05] A. Yoshida, V. Blanz, K. Myszkowski, and H. Seidel. Perceptual evaluation of tone mapping operators with real-world scenes. In *Human Vision and Electronic Imaging X, SPIE*, 2005.

[YYBR14] Charles Yaacoub, Cendrella Yaghi, and Christine Bou-Rizk. Fusion of tone-mapped high dynamic range images based on objective range-independent quality maps. In *ICASSP*, pages 1195–1199, 2014.