

Illumination Space:

A Feature Space for Radiance Maps

by

Andrew Chalmers

A thesis submitted to the
Victoria University of Wellington
in fulfilment of the requirements for the degree of
Doctor of Philosophy
in Faculty of Computer Science and Engineering

Victoria University of Wellington
2018

Supervisory Committee

Primary Supervisor: Associate Professor Taehyun Rhee

Secondary Supervisor: Associate Professor John Lewis

School of Engineering and Computer Science,
Victoria University of Wellington

Abstract

From red sunsets to blue skies, the natural world contains breathtaking scenery with complex lighting which many computer graphics applications strive to emulate. Achieving such realism is a computationally challenging task and requires proficiency with rendering software. To aid in this process, radiance maps (RM) are a convenient storage structure for representing the real-world. In this form, it can be used to realistically illuminate synthetic objects or for backdrop replacement in chroma key compositing. An artist can also freely change a RM to another that better matches their desired lighting or background conditions. This motivates the need for a large collection of RMs such that an artist has a range of environmental conditions to choose from. Due to the practicality of RMs, databases of RMs have continually grown since its inception. However, a comprehensive collection of RMs is not useful without a method for searching through the collection.

This thesis defines a semantic feature space that allows an artist to interactively browse through databases of RMs, with applications for both lighting and backdrop replacement in mind. The set of features are automatically extracted from the RMs in an offline pre-processing step, and are queried in real-time for browsing. Illumination features are defined to concisely describe lighting properties of a RM, allowing an artist to find a RM to illuminate their target scene. Texture features are used to describe visual elements of a RM, allowing an artist to search the database for reflective or backdrop properties for their target scene. A combination of the two sets of features allows an artist to search for RMs with desirable illumination effects which match the background environment.

Acknowledgments

Thanks to my great friends and colleagues who have helped me over the years: Christopher Dean, Jin Jong Choi, Lohit Petikam, Tommy Iorns, Ben Allen, Kurt Wan-Duo Ma, Stephen Thompson, Kieran Carnegie, Joshua Scott, Ping Liu, Ian Loh, Richard Roberts, Diana Siwiak, Neil Dodgson, Fang-Lue Zhang, Rob Lindeman, Todd Zickler, Agata Migalska, Craig Spence, Chandan Pawaskar, Evgeny Patrikeev, Jen Gillespie, Charlie Tait, Peter Hillman, and the Victoria Computer Graphics group.

I also appreciate the supporting people from Victoria University including Peter Andreae, Stuart Marshall, Monique Damitio, Marcus Frean, Roman Klapaukh, Mark Davies, Roger Cliffe, Radek Hes, Siyun Thompson, and Zarinah Amin.

Particular thanks goes to my supervisors Associate Professor Taehyun Rhee and Associate Professor John Lewis for their generous guidance and insight during my study.

Finally, I thank my partner Juliet Mohi for her encouragement and support in this academic endeavour.

Publications

- L. Petikam, **A. Chalmers**, and T. Rhee, "Visual Perception of Real World Depth Map Resolution for Mixed Reality Rendering", *Proceedings of IEEE Virtual Reality*, Reutlingen, Germany, March 2018.
- T. Rhee, L. Petikam, B. Allen, and **A. Chalmers**, "MR360: Mixed Reality Rendering for 360-degree Panoramic Videos", in *IEEE Transactions on Visualization and Computer Graphics*, vol. 23, no. 4, pp. 1379-1388, April 2017, presented at IEEE Virtual Reality 2017 and SIGGRAPH 2017.
- W. Ma, T. Iorns, **A. Chalmers**, and T. Rhee, "Synthesizing Radiance Maps from Legacy Outdoor Photographs for Real-time IBL on HMDs", *Proceedings of the 30th International Conference on Image and Vision Computing New Zealand*, Auckland, New Zealand, November 2015.
- **A. Chalmers**, J. Choi, and T. Rhee, "Perceptually Optimised Illumination for Seamless Composites", *Proceedings of Pacific Graphics*, Seoul, South Korea, October 2014.
- **A. Chalmers**, J.P. Lewis, P. Hillman, C. Tait, and T. Rhee, "Sky Browser: Search for HDR Sky Maps", *Proceedings of Pacific Graphics*, Seoul, South Korea, October 2014.

Contents

Supervisory Committee	i
Abstract	iii
Acknowledgments	v
Publications	vii
List of Figures	xiii
List of Tables	xxi
1 Introduction	1
1.1 Research Questions	3
1.2 Objectives and Contributions	3
1.3 Thesis Structure	5
2 Background	7
2.1 Lighting	7
2.1.1 Image Based Lighting	8
2.1.2 Radiance Map Creation	10

2.2	Feature selection	11
2.2.1	Texture based features	12
2.2.2	Illumination based features	12
3	The Human Visual System and Illumination from Radiance Maps	17
3.1	Introduction	17
3.2	Related Work	20
3.3	HVS in Local Illumination Changes	22
3.3.1	Experiment Setup	22
3.3.2	Stimuli	24
3.3.3	Procedure	25
3.3.4	HVS Observation and Analysis	26
3.4	HVS for Inverse Tone-Mapped Radiance Map	29
3.5	Perceptually Optimised Radiance Maps	33
3.6	Composition with a Photograph	33
3.7	Conclusion	39
4	Sky Browser: A Reflectance and Texture Feature Space for Radiance Maps	41
4.1	Introduction	41
4.2	Related Work	46
4.3	Overview	47
4.4	Features	49
4.4.1	GLCM Correlation	49
4.4.2	Entropy of the Laplacian (EL)	50
4.4.3	Spherical Harmonics Ratio of Red and Blue	50
4.4.4	Average Intensity	55

4.4.5	Feature Correlation	55
4.5	Results	56
4.6	Conclusion	59
5	Illumination Space: An Illumination Feature Space for Radiance Maps	61
5.1	Introduction	61
5.2	Related Work	65
5.3	Overview of Light Features	66
5.4	Dominant Light Feature Extraction	68
5.4.1	Dominant Light Model (DLM)	68
5.4.2	Dominant Light Detection	71
5.4.3	DLM Fitting	76
5.5	Ambient Light Feature Extraction	82
5.5.1	Ambient Light Model (ALM)	82
5.5.2	ALM Fitting	82
5.6	Illumination Space	83
5.6.1	Distance measure	83
5.6.2	Browsing	85
5.6.3	Searching	87
5.7	Results	87
5.7.1	Dominant Light Model	88
5.7.2	Complexity User Study	89
5.7.3	Dominant Light Detection	92
5.7.4	Browse using features	95
5.7.5	Search using a radiance map	97

5.7.6	Search using a shadow profile	101
5.8	Conclusion	104
6	Conclusion	107
6.1	Summary	107
6.2	Discussion	109
6.3	Limitations	111
6.4	Future Work	111
	Ethics	113
	Bibliography	121

List of Figures

2.1	Light sources commonly found in CG software. From left: ambient, point, directional, spot, and area light sources.	8
2.2	Image based lighting. The RM (a) is used to illuminate the scene (b), which consists of a glossy robot hand, reflective chrome ball, and a diffuse ball.	9
2.3	RMs (top row) and their corresponding rendered result (bottom row). From left: measured (photographed), analytical (Hosek-Wilkie model), and artist created RM.	11
2.4	Visualisation of the feature space using the ℓ_2 -norm and dimensionality reduction on RMs down to three dimensions (visualised with the x-axis, y-axis, and colour). Points a, b, c, d and e are sampled in the graph, and their corresponding RMs (left) and rendered images (right) are shown ([a-e, top-down]). There is no clear similarity between each RM, nor is there any sense of spatial continuity between RMs.	15
3.1	Examples of RM modifications in the user study, where the left column is an IBL rendering with the original HDR RM - this acts as a simulation of a photo. The right column has a synthetic object composited into the photo under improper illumination. The synthetic object of the first row is illuminated with an 8-bit RM, the second row's RM has its intensity reduced, and the third row modifies the direction of the RM.	23

3.2	Observation of the HVS based on (a) illumination direction and (b) intensity. The noticeability scale is the mean score of the participants. The graph is focused in the range of sensitivity, beginning at the moment that the illumination details were detected by the HVS.	28
3.3	The user study results of IBL using RM of an 8-bit LDR, extended LDR to HDR, and the extended LDR to HDR of optimized resolution.	30
3.4	The dynamic range of each scene. Column 1, 2 and 3 is the overcast, indoor and sunny scene respectively. Row 1 is the RM, and rows 2, 3 and 4 is the dynamic range of the LDR, LDR to HDR, and HDR respectively. The LDR luminance values have been normalised between 0 and 1.	32
3.5	Results of the survey with respect to RM resolution. The noticeability scale is the mean score of the participants, and each category represents a decrease in resolution.	34
3.6	Overcast RM lighting synthetic objects (green block, blue ball) composited into a photo with varying resolution and bit range. A seamless composition is achieved even with a low resolution LDR to HDR RM.	35
3.7	Indoor RM lighting synthetic objects (green block, blue ball) composited into a photo with varying resolution and bit range. A seamless composition is achieved even with a low resolution LDR to HDR RM.	36
3.8	Sunny RM lighting a synthetic objects (green block, blue ball) composited into a photo with varying resolution and bit range. The artefacts in the shadow is due to the fact that directly capturing the sun is difficult, even with many exposure levels. However, the lighting remains consistent in quality between each image.	37

3.9 Composition with a photograph: the synthetic objects (blue ball, green block, bunny, and teapot) are illuminated by a RM. From left: photograph, rendering with 5024x2512 HDR RM, and two examples of rendering with 320x160 LDR RM having expanded dynamic range using inverse tone mapping. The 2nd and 3rd images are visually equivalent, despite the large difference in RM resolution and dynamic range.	38
4.1 Three examples of backdrop composition. The synthetic object (Orc) is composited into a live action background. While maintaining the same illumination of the Orc, the distant backdrop (sky) is changed to adjust the mood of the scene. The sky colour has been modified for the composition.	44
4.2 Textural feature space examples for clouds.	51
4.3 The spherical harmonics expansion. Above the expansion is the input image, and below is the approximation of the input. To the right of each band is the corresponding approximation of the input image. . .	51
4.4 Textural feature space examples for tone.	52
4.5 Ambient tone colour distribution showing that the HDR database is dense around red, blue and grey (low saturation) values. Highly saturated areas tend to represent bright skies, which often take on red (sunset) or blue (clear sky) values. Low saturation areas tend to represent cloudy skies. Indoor scenes typically show low saturation values with varying hue.	54
4.6 Feature space distribution, where each point represents an image. Left: the textural feature space. Right-top: the red and blue spherical harmonic. Right-bottom: the colour of the sky in one dimension by taking the ratio of the red and blue value.	56

4.7	The search results (navigating the feature space). The transition sequence is as follows: increase intensity, slight move toward blue, decrease smoothness and increase edges, increase smoothness, and finally a large decrease of edges.	57
5.1	A pipeline overview for organising RMs into a queryable database based on the illumination properties of RMs. The user input and search function can vary based on the application, as outlined in Section 3.3. Examples includes an artist directed search function using the features or example based searching using an example RM or shadow profile.	62
5.2	(a) is the model fitting the integral of the light source and (b) is the model fitting the raw shadow data. We demonstrate a close correspondence between the two fits. An error of 0.824° , 1.321° and 0.01 for the light size, elevation and amplitude respectively. We also illustrate how our model accurately fits the entire curve compared with Mohan et al.'s shadow edge model.	70
5.3	On the left is a top-down orthographic camera of a cylinder on a plane. This setup is used to obtain a clear view of the shadow information cast in all directions on the horizontal plane, and on the right is an example rendering using an IBL as a light source.	74
5.4	From the left: rendered image, difference image showing shadow detail of the cylinder on the ground plane, and thresholding to remove shadows not perceivable to the human eye using Weber's Law. These show a zoomed in version of the data that was used.	75
5.5	Our light detection algorithm computes superpixels on the RM (left) and finds the set of superpixels which contribute as a dominant light source (right).	76

5.6 The fitting pre-process. (a) The input data is oversampled using cubic interpolation. (b) The derivative of the cubic curve (or a 1D representation of the original light source) is computed and a Gaussian distribution is approximated around the peak. (c) Using the Gaussian, the oversampled curve is resampled, producing a higher density around the penumbra. (d) The DLM approximates the original curve using the oversampled and high density points.	79
5.7 (a) Fitting the light model to the (b) dominant light source. The red horizontal lines in (b) visualise the parameters of the model in (a), showing a good fit around the dominant light source (in this case, the sun).	80
5.8 Browsing through the illumination space. Our light model describes important illumination properties of RMs. Light features are inferred from the model, allowing for RM databases to be sorted with semantically meaningful parameters. RMs within the same cluster share similar illumination properties. There is a mutual correspondence between the light source (light elevation, light area, and multiple lights) and the rendered image (shadow length, shadow softness, and shadows complexity).	84
5.9 Visualisation of the feature space with elevation and tone. Points a, b, c, d and e are sampled at a low elevation, varying from blue to red. The RMs and their corresponding rendered images are shown ([a-e, top-down]).	86
5.10 Fitness error comparing our DLM (green) with [Mohan07]’s SEM (red). The errors are normalised and compared horizontally (smaller is better). There are 2849 comparisons made (incrementing on the y axis), and the results are sorted in descending order (left to right) for readability.	88

5.11 The shadow complexity pairwise comparison (top) and Likert scale (bottom) test results. The results show that complexity is mostly influenced by the number of shadows rather than other factors such as shadow sharpness, symmetry, length, or a mixture of soft/sharp shadows.	90
5.12 The shadow complexity Likert scale test results, from simple (bottom) to complex (top). The results are sorted in ascending order (left to right) for readability.	91
5.13 Comparison of dominant light detection methods (top row) and their affects on the rendered image (bottom row). (a) is the ground truth, followed by various light detection algorithms. (b) is naive thresholding with a threshold value of $T=1.0$, (c) is Karsch's method, and (d) is our method. Notice that the rendered images look similar to the ground truth, yet our method has removed unnecessary light sources in the yellow dotted region to extract the dominant lights. (e) is thresholding by increasing the threshold to remove the unnecessary light sources, yet this introduces artefacts in the rendered scene due to removing other important light sources.	93
5.14 User study results comparing [Karsch14]'s light detection (red bars), thresholding (green bars) and our method (blue bars). The reduced amount of dominant light pixels averaged across for each RM is 79.3%, 80.1% and 95.6% for each light detection method respectively. ML are the RMs with multiple lights, SL has one soft light, and HL has one hard light.	94
5.15 Feature F_{ls} , the transition from hard to soft shadows. In the RM, this corresponds to the size of the dominant light source, transitioning from small to large. RMs with similar dominant light elevation F_{le} are chosen.	96

5.16 Feature F_{le} , the transition from long to short shadows. In the RM, this corresponds to the elevation of the dominant light source, transitioning from low to high. RMs with similar shadow softness F_{ls} are chosen.	96
5.17 Feature F_{la} , the transition from the lightest to the darkest shadow. Specifically, it is the darkest point in the shadow relative to the unoccluded diffuse intensity. In the RM, this corresponds to intensity of the dominant light source relative to the ambient light, transitioning from lowest to highest intensity. RMs with similar shadow softness F_{ls} are chosen.	96
5.18 Feature F_{lc} , the transition from low to high complexity shadows. The center row shows how the complexity measure corresponds to the number of detected dominant lights (red dots have been added to the center of each detected dominant light source). The rendered images have been normalised to show the detail.	96
5.19 The returned results from the ambient tone search. The left-most RM is the input, and the following 4 images are the closest images returned in terms of the ambient tone.	99
5.20 The top row is the input RM with its corresponding rendering of a cylinder on a plane. In the same column, the closest 3 RMs to the input RM are shown for each distance measure. Column 1 shows an example with sun light with cloud scattering, column 2 is an example of a red sunset, and column 3 is an overcast blue sky. For our distance measurement, column 2 and 3 are user directed toward tone and shadow detail respectively, whereas column 1 is fully automatic, aiming to find RMs with both similar shadows and tone. The rendered images have been normalised to show the detail.	100

- 5.21 Searching through the illumination space with a sunny scene. Given an input photograph and the shadow profile, our light model computes the light features from the shadow information. Using these features, a RM is found in the illumination space. The RM provides similar illumination to the input photograph, allowing for a realistic composition with synthetic objects. 102
- 5.22 Searching through the illumination space with an overcast scene. The RM provides similar illumination to the input photograph, allowing for a realistic composition with synthetic objects. 103

List of Tables

3.1	Groups of stimuli in the perceptual study	25
4.1	Feature Correlation	55
4.2	User Evaluation Results for <i>Sky Browser</i>	58
5.1	Each feature's semantic relationship between the light source and the rendered image. Above the dashed line are the dominant light features, below are the ambient light features.	68
5.2	Quantitative Evaluation: SSIM between input and returned images .	97

Chapter 1

Introduction

In computer graphics, the ultimate receiver of image information is the human eye. As such, there are two important elements to consider; the image and the observer. An image is comprised of objects, material properties, and light sources which reflect light off the surface for the observer to perceive. The observer is sensitive to various illumination effects and may discriminate against visual artefacts which are deemed unrealistic or unimpressive. Therefore, in order to present an image with desirable visual effects, it is important to control (by an artist or capturing devices) the lighting in the image.

To *create* illumination effects, there are many commercial and open source tools [1, 2, 3]. These tools provide artists with an interface to create effects of high quality in a short amount of time. As computer graphics in the film and games industry continues to grow rapidly in budget [4, 5] and visual fidelity, there is an increasing need for artists to provide high quality visual effects under strict time constraints [6]. Thus the research and development invested in tools for illumination creation is an important and challenging area of interest.

The creation of illumination can be simplified down to three main components: light sources, geometry and material properties. There are many ways for defining each component, but the core focus of this thesis is defining the light source. Abstract models of light sources, such as a directional light, achieve simplistic lighting effects with a few intuitive parameters. However, the end result does not contain high

quality lighting unless multiple light sources are added to the scene.

One solution to achieve high fidelity lighting is to store a large set of light sources in a spherical texture. The spherical texture surrounds the scene, and each pixel of the texture is treated as a separate light source. This texture is referred to as a radiance map (RM), and the rendering technique that uses a RM for lighting is referred to as image-based lighting (IBL).

Research has predominantly focused on developing analytical [7, 8, 9, 10], measured [11, 12, 13, 14, 15], and artist driven [16, 17, 18, 19, 20] methods to create RMs. These tools provide artists with many options for creating desired illumination effects, from stylistic to photo-realistic results. However, each method has its own drawbacks. Analytical models are constrained by the model itself, such that artists cannot create RMs outside of what is defined by the model. For example, day-time clear sky models do not parametrise night-time skies or skies with clouds. Artist driven methods are time consuming and are often physically inaccurate. This is particularly important when the artist is aiming to produce artistic yet physically accurate lighting. Measured methods accurately describe real-world lighting but are time consuming to produce as it requires users to manually capture the data.

While each method has its own drawbacks, it shows that there is a desire by artists to have a large variation of high fidelity light sources to choose from. This motivates the need for a large collection of RMs (comprised of analytic, measured and artistic RMs) to match various moods and lighting conditions. However, a comprehensive collection of RMs is not useful without a method for searching through the collection.

This thesis defines a novel set of features that concisely describes the important properties of RMs. The features then construct a novel '*illumination space*', which is a low dimensional feature space that organises the RMs with intuitive parameters. The features have clear semantic meaning, allowing users to browse for desired RMs intuitively, such as bright sunny skies, red sunsets or indoor scenes with the intended illumination.

1.1 Research Questions

This thesis addresses the following questions:

- i) *What are the most important characteristics of a radiance map?*

A RM is useful for high quality lighting for synthetic objects. This quality is achieved due to the fact that a RM is comprised of many light sources. However, describing many light sources concisely is a challenging problem. Finding a concise set of characteristics of RMs will reduce the complexity of defining the features.

- ii) *What are the features that describe the important characteristics of a radiance map, and how are they computed?*

The set of features extracted from a RM will need to be mathematically defined. The defined features should also concisely capture the important characteristics of a RM. Once the features are defined, they will make up the basis of the illumination space. An important element of this thesis is the application of the features to a database of RMs. However, a database may contain many entries such that it becomes impractical to manually attribute the features to each RM. In order to apply the features to a RM database, an algorithm should be able to iterate over the database and automatically extract the features.

- iii) *What applications or research areas can leverage the illumination space?*

An organised RM database will have implications on various applications and research areas. Applications such as browsing, searching, and database visualisation will be considered. Furthermore, research areas such as inverse rendering and distribution interpolation will be discussed.

1.2 Objectives and Contributions

The primary goal of this thesis is the organisation of RM databases. That is, defining how data point entries in the RM database relate to one another using

low-dimensional features. Once the illumination space is established, this thesis will explore uses of the illumination space in novel ways. This includes searching the database using parameters, exemplars, and inverse rendering. There are three primary research questions, and each research question will be addressed with an objective and solution which make up this thesis' contributions.

- i) *Perceptual User Study:* A novel perceptual user study is conducted to observe the important illumination and textural properties of RMs. Previous approaches to perceptual user studies in this research area are not appropriate for RMs. We introduce a new user study for this purpose. The study is setup to observe how parameters relevant to RMs affect the illuminated synthetic objects. From the results of the user study, various observations of RMs are made. These observations act as a guide for selecting a set of concise features for RMs. Additionally, the user study produces results which lead to two optimisations in RM rendering.
- ii) *Light and Textural Features:* A set of analytical values that describe the important properties of RM is guided by the observations from the user study. The novel features concisely describe RMs in a semantically meaningful way. The features are obtained automatically in order to accommodate large databases of RMs. To achieve this, a novel automatic light detection algorithm for HDR RMs is described. Furthermore, an automatic light model fitting algorithm is also described. The light features, as opposed to the texture features, are more important for describing the illumination properties of RMs. However, for a complete pipeline that also accounts for high specular reflections and backdrops, texture based features are covered as well.
- iii) *Distance measures and applications:* Distance measures are formulated to organise a RM database based on both the RM's illumination and textural properties. The distance measures define how RMs in the database relate to one another. Using the measures, applications can leverage the RM database. Applications include browsing the database through visualisation, searching the database using exemplars, and inverse rendering.

1.3 Thesis Structure

This thesis investigates important properties of RMs for both illumination and textural purposes. As such, this thesis analyses the human visual system to derive the important elements of the RM to describe. From the analysis, the illumination and aesthetic features are established. This thesis covers topics in human vision, rendering, and texture-based feature detection. Each topic consists of extensive related work, therefore, each chapter addressing each topic will be self-contained with an introduction, related work and contributions.

To establish the problem domain, Chapter 2 provides a general overview and describes the limitations of the current work. The contributions of this thesis are in the following chapters:

- i)* Chapter 3 outlines a perceptual user study to obtain the important characteristics of RMs. In order to achieve this, a user study is conducted where the RMs are used to illuminate synthetic scenes. The participants are then asked questions regarding the rendered images in order to observe important characteristics that the participants identify. Novel contributions in this chapter include a novel user study, identifying key RM characteristics, as well as various RM optimisations obtained from the user study.
- ii)* Chapter 4 describes novel texture based features which are used to organise the RM database based on its textural properties. This is particularly focused on describing cloud properties such that an artist can search for RMs with desirable clouds or clear skies. The tone and overall intensity of the RM are also described such that artists can find red sunsets to blue skies.
- iii)* Chapter 5 describes novel light features which are used to organise the RM database based on its illumination properties. Additionally, this chapter outlines a novel light model as well as a novel light detection algorithm.

Finally, Chapter 3, 4 and 5 are summarised, discussed and concluded in Chapter 6.

Chapter 2

Background

This chapter is separated into two sections. Section 2.1 is an overview on *lighting* and describes general methods for capturing or modifying light sources. This includes measured, analytical and artist driven methods. Following this, research on RM *features* is then presented in Section 2.2. This section describes current methods for organising RM databases and discusses their limitations. These two sections describe the necessary background and limitations of the problem that this thesis addresses. The following three chapters contain further background and related work which is specific to the problem domain addressed by each respective chapter.

2.1 Lighting

In computer graphics (CG), lighting is important for conveying the mood and shape of geometry. There are numerous types of light sources that are used for differing effects, as shown in Figure 2.1. These light sources are typically used for *direct illumination* (as opposed to indirect or global illumination (GI)), where light transport does not consider how light propagates in the scene after bouncing off an object. Because direct illumination limits the light transport, it is a useful technique for real-time lighting. GI strategies on the other hand consider all incoming light from every direction, including light bouncing off surrounding objects, resulting in

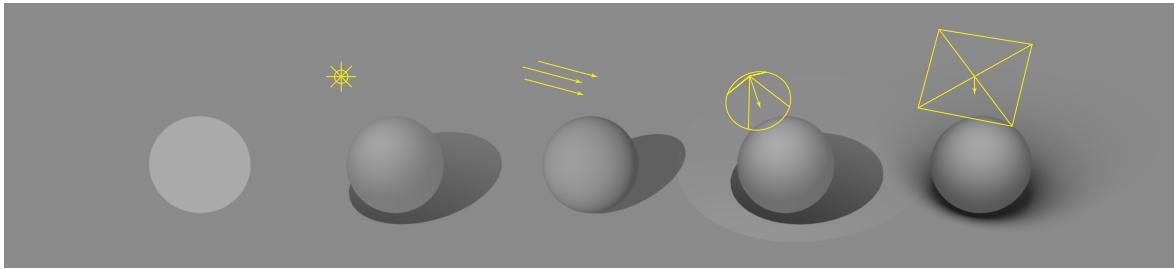


Figure 2.1: Light sources commonly found in CG software. From left: ambient, point, directional, spot, and area light sources.

realistic lighting. The rendering equation [21] is used to achieve this effect

$$L_o(x, w) = L_e(x, w) + \int_{\Omega} fr(x, w, w_i)L_i(x, w_i)(w_i \cdot \mathbf{n})dw_i \quad (2.1)$$

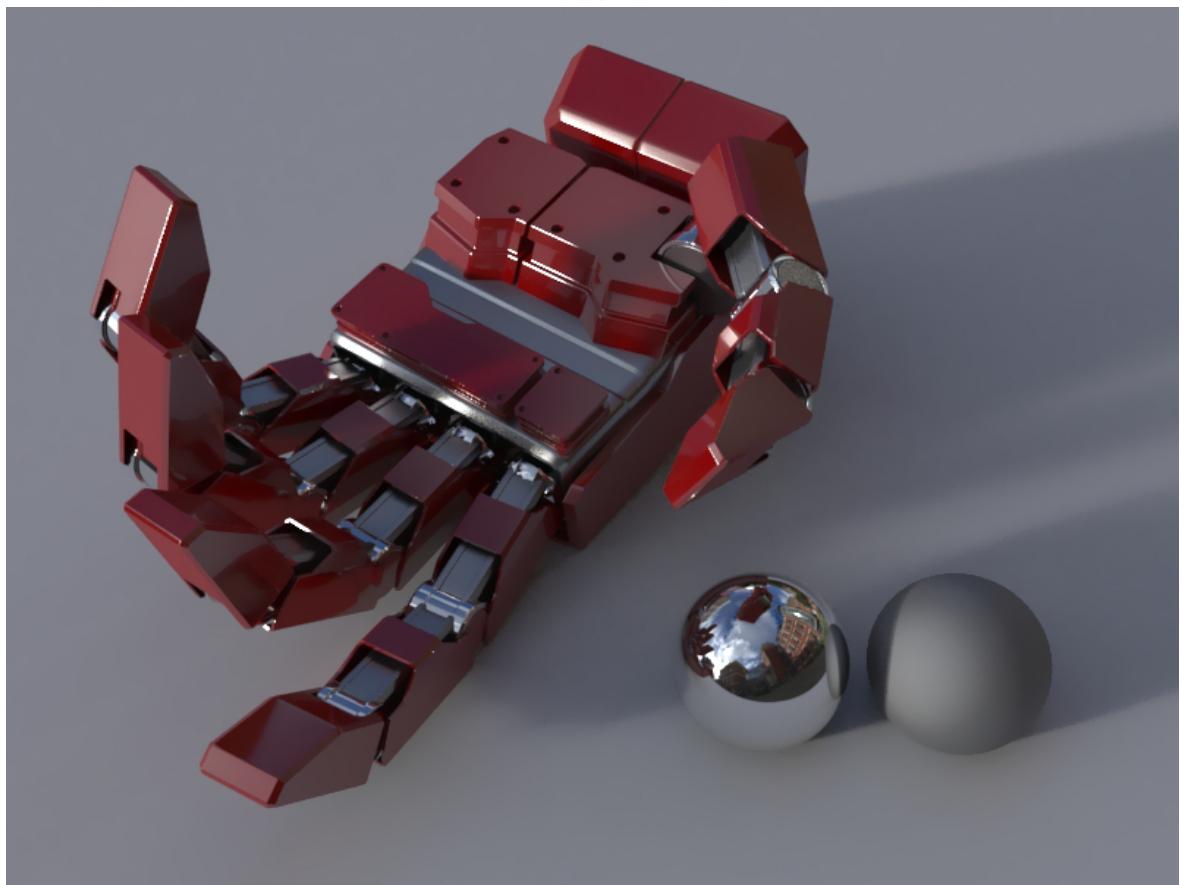
where L_o is the outgoing light at point x in the direction w , L_e is the emitted light from point x , integrating all incoming light w_i over the hemisphere Ω in the direction of \mathbf{n} , fr is the material property, L_i is the incoming light at point x from direction w_i , which is attenuated by the cosine fall-off between w_i and the surface normal \mathbf{n} . Due to the recursive integral, this is a costly procedure, and as a result, research has focused on various methods for optimising GI.

2.1.1 Image Based Lighting

One solution to achieve realistic lighting is to store a large set of light sources L_i in a spherical texture [22]. The spherical texture surrounds the scene, and each pixel in the texture is treated as a light source. In order to store a wide dynamic range of radiant light information, the texture is stored in a high dynamic range (HDR) format, storing floating point values. The spherical texture is referred to as a radiance map (RM), and the rendering technique using the RM for lighting is referred to as image-based lighting (IBL) [23] (see Figure 2.2 for an example of IBL). The spherical texture can also be referred to more generically as an environment map, where it does not specifically act as a source with radiant light information, but is also used as a backdrop to a scene. The term *radiance map* is used throughout



(a)



(b)

Figure 2.2: Image based lighting. The RM (a) is used to illuminate the scene (b), which consists of a glossy robot hand, reflective chrome ball, and a diffuse ball.

this thesis unless specified otherwise.

This kind of lighting is often used for composition of synthetic objects into a photograph, where real-world lighting can be captured and used for illuminating synthetic objects into the scene [24, 25, 26]. This is commonly seen in live-action films, virtual reality and augmented reality. While particularly useful for composition, they are also simple alternatives for GI effects in completely synthetic scenes. The simplification comes from removing the recursive component of the rendering equation by storing light transport values directly into the RM. For example, the path of light from the sun, reflected off a building and into the camera is stored as a single pixel value in the direction of the building. However, storing lighting information as an image impacts memory usage as well as the performance of sampling rates in high resolution RMs. This has implications for further research in sampling and compression algorithms.

2.1.2 Radiance Map Creation

There are three methods for creating RMs: measurement (photography), analytical functions (simulation), or drawing/painting [23] (see Figure 2.3). Photography became a common approach since the introduction of recovering HDR RMs from images with conventional cameras [27]. This research has been further extended to accurately capture direct sunlight [28] and image calibration [29].

Artist created RMs produce the least physically accurate RMs, though they provide stylistic freedom. These are created by image editing programs where the artist paints light sources directly into the image [16, 17].

Analytical models allow artists to create new RMs without the need to manually photograph or draw the environment. The Preetham skylight model [30] is a widely used analytical solution. This model gives artists accurate control over atmospheric simulation with relative ease of use for controlling conditions such as haze. Research in this domain continues to produce accurate, predictive and user-friendly models [31, 32, 33], with the Hosek-Wilkie model [9] as the state of the art. The analytical solutions are derived from measured and simulated spectral radiance of the sun and

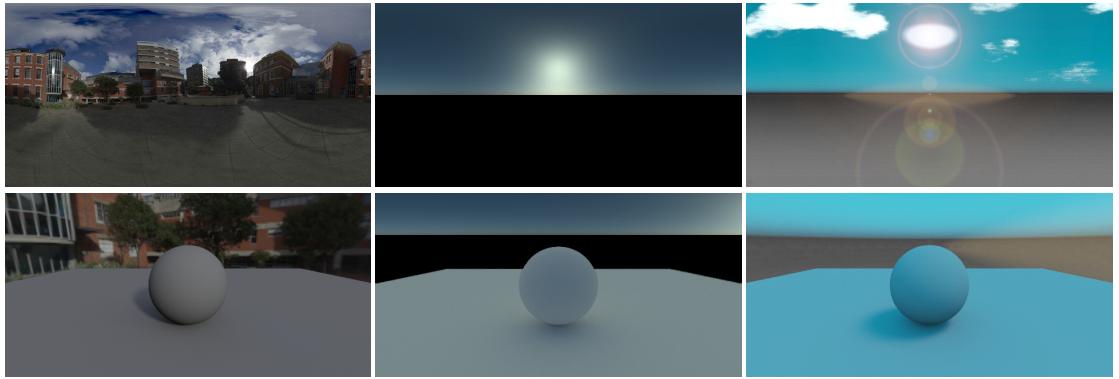


Figure 2.3: RMs (top row) and their corresponding rendered result (bottom row). From left: measured (photographed), analytical (Hosek-Wilkie model), and artist created RM.

sky. However, creating a new RM is limited to the model. For example, the model does not account for overcast clouds, trees, buildings, the night sky or indoor scenes. The sky model is limited to defining clear skies with varying solar elevation and atmospheric scattering effects, allowing for rendered scenes with sharp shadows and varying shadow length based on the solar elevation.

There are large databases of RMs comprised of all three of the above methods of creation. While the simulated database of RMs can be organised based on the sky model parameters, the wider set of RMs encompassing measured and artistically created RMs still requires a set of parameters to organise it by. The parameters for describing the illumination properties of the entire database is the core goal of this thesis.

2.2 Feature selection

A way of organising a database of RMs is to parameterise each RM. The parametrisation allows for a method of computing how similar or different a set of RMs are. These parameters are referred to as features, as each one represents a key feature of the RM. Determining what the important elements of RMs are is one of the challenges addressed by this thesis. A concise set of features allows for intuitive and interactive browsing of the database. Features based on the textural

properties of RMs also need to be considered to allow artists to find RMs with the desired backdrop to the scene. The reflections on the virtual objects, such as a chrome ball, should also be similar to the photographed scene that the virtual object is composited into. As such, textural properties will also ensure that the reflections on the virtual objects exhibit similar textural properties as the background. Furthermore, the illumination properties (such as the cast shadows or colour of the lighting) of the virtual objects should be similar to the photographed scene. For a high quality composition, the lighting, reflections and backdrop need to be accounted for. Therefore, the features defined in this thesis describe both illumination and textural properties.

2.2.1 Texture based features

This thesis is focusing on the illumination aspect of RMs, and accounts for specular materials (such as the mirror chrome ball in Figure 2.2) through textural features. A common component in IBL are clouds in outdoor scenes, thus detecting these elements may be crucial in the proposed illumination space system. Furthermore, for backdrop purposes, clouds are important for setting the mood of the scene. As such, sky and cloud feature extraction is presented in Chapter 4. Textural properties of indoor scenes is outside the scope of this thesis.

2.2.2 Illumination based features

An important element of the RM is how it illuminates the scene. This includes how objects cast shadows, the glossy highlights, as well as the tone of the light source. The research in this domain is limited due to the lack of available resources for HDR RM databases. Because RMs have been used extensively in production for films for the past two decades, production companies have developed a large set of RMs. Few online resources [34, 35] provide a small amount (\sim 100 RMs) of data which can be used. A few other sites [36, 37, 38, 39, 40, 41, 42, 43] offer a moderate number of RMs. Most recently, a large collection of indoor RMs have been made available and is used in recent work [44]. Karsch et al. [25] has also shown interest

in using a database of RMs. In this case, they utilise low-dynamic range (LDR) panoramic databases such as SUN360 [45] which are tone-mapped to HDR. Such applications can be improved with features specifically focusing on the illumination properties of RMs. As previously described, analytical sky models are used to simulate skies with a set of parameters. The parameters contain terms for describing the fall-off lighting above the horizon and directly around the sun. However, these models are specifically designed for sunny clear skies. This is not practical for typical RMs which contain image content such as clouds and buildings. This is especially problematic if the image content obscures the horizon or the sun. The ideal set of features should not be limited to such a model, and ideally capture properties on a diverse range of RMs (including indoor and overcast lighting).

Another method for sorting RMs is the use of Smart IBL (sIBL) [34]. sIBL is a format which organises the data using tags such as light position, colour, alternate light sources, and other details. However, sIBL is only a specification, and providing tags such as the light position is not straight forward. One issue is the detection of the light position in the RM. Because we are dealing with large databases of RMs, it is undesirable to manually specify such tags. Furthermore, the specification is missing key components, such as the size of the light source and if there are multiple light sources, how they relate to one another. However, the sIBL specification is a good starting point for identifying basic properties which are desirable by artists.

The illumination space proposed in this thesis (detailed in Chapter 5) provides a stronger framework for finding appropriate RMs than the current state of the art. The key idea behind the illumination space is that there is a direct correspondence between the RM and the rendered scene. Furthermore, the features allow for changes in the parameters such that it is possible to naturally move between RMs continuously.

To illustrate the problem of organising a RM database, a naive solution is demonstrated. On a dataset of 496 RMs, the ℓ_2 -norm is computed between each RM, forming a distance matrix with a size of 496x496. Dimensionality reduction is computed on the distance matrix, clustering RMs together in two dimensions. It is shown that the clustering is not accurate, nor is the manifold intuitive to

navigate (see Figure 2.4).

The RMs are similar in space, yet their shadow profiles vary considerably. (a) shows a darker image with a sharp shadow, (b) has a long soft shadow, (c) has a shorter soft shadow, (d) has a very short and very soft shadow, and (e) has a medium length, fairly sharp shadow. Furthermore, the overall intensity and hue of the rendered scene varies. Given how close the RMs are, it is expected that the RMs demonstrate similar illumination properties. Furthermore, the points (a)-(e) are sampled linearly, yet their rendered images do not demonstrate any continuous change in illumination.

This thesis solves these problems by automatically detecting and extracting features which cluster RMs intuitively. The clustering is continuous on each feature, allowing for continuous changes in illumination. The features make up a low dimensional feature space allowing for real-time browsing of the RM database.

In order to determine the important elements of RMs, a perceptual user study is conducted in Chapter 3. From the user study, the features are defined and applied to the RM database. Texture based features that account for reflective and backdrop properties of RMs are described in Chapter 4. Following this, light features that account for the lighting properties of RMs are described in Chapter 5. Finally, the thesis is summarised, discussed and concluded in Chapter 6.

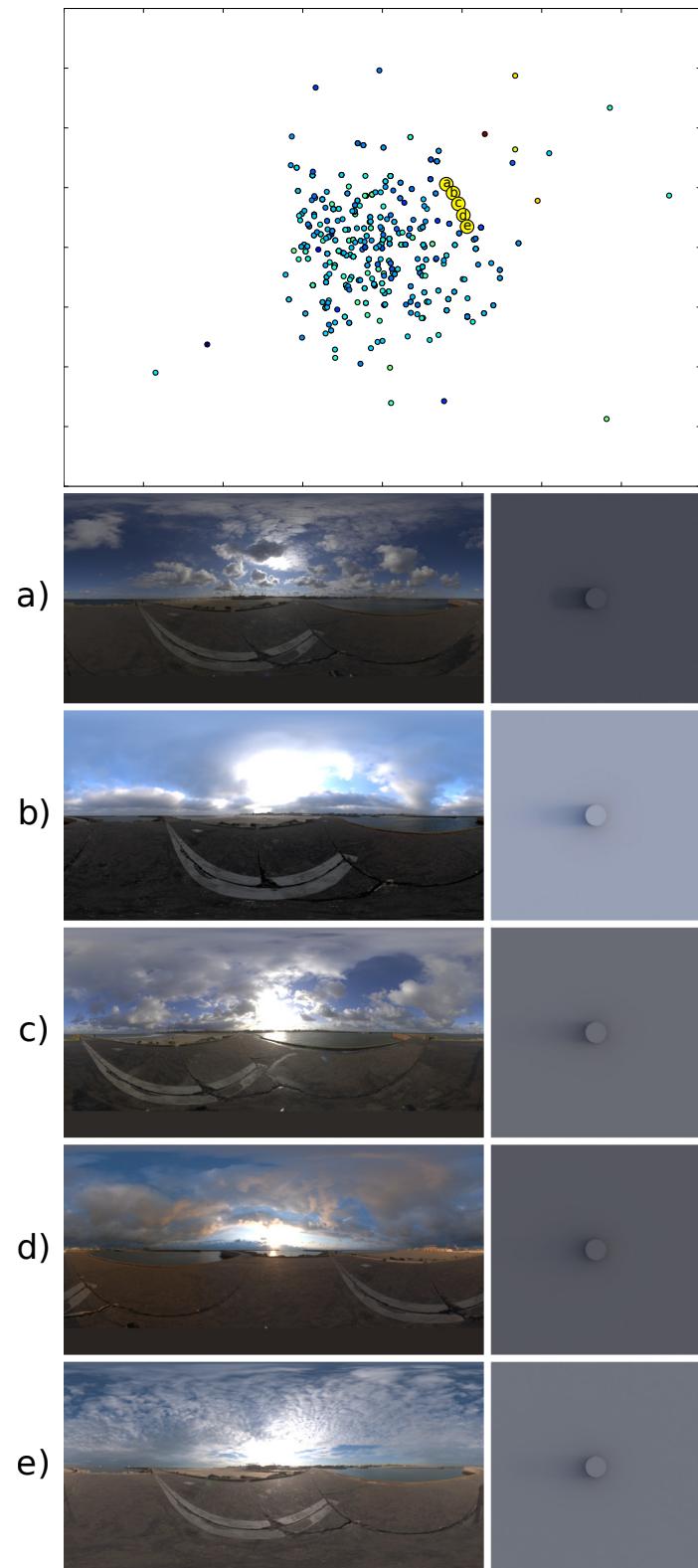


Figure 2.4: Visualisation of the feature space using the ℓ_2 -norm and dimensionality reduction on RMs down to three dimensions (visualised with the x-axis, y-axis, and colour). Points a, b, c, d and e are sampled in the graph, and their corresponding RMs (left) and rendered images (right) are shown ([a-e, top-down]). There is no clear similarity between each RM, nor is there any sense of spatial continuity between RMs.

Chapter 3

The Human Visual System and Illumination from Radiance Maps

3.1 Introduction

The aim of this thesis is to describe the illumination properties of RMs, therefore, the features should describe properties that affect the illumination as observed in the rendered image. In order to determine the important illumination properties to measure, a user study is conducted. This user study not only contributes to the selection of key features, but also introduces a novel user study framework for evaluating the illumination quality of a rendered object. Other contributions in this chapter include a perceptually optimised RM resolution, as well as verifying conversion of LDR to HDR RMs. The observations of the user study guide the selection of key features, and the optimal RM resolution provides a performance increase for image processing time.

The realistic mixture of synthetic objects with the photographed scene is one of the primary goals in mixed reality and visual effects in films. The overall aim is to generate a seamless composition between the rendered 3D objects and the photographed real world it occupies. The seamlessness is the measure of how perceptually apparent the synthetic object is in the final composition, and how indistinguishable it is from the photographed scene. Given that the ultimate

receiver of image information is the human eye, this measure is determined by the human visual system (HVS).

In order to provide real-world lighting information, RMs are commonly used for IBL. This requires the extra effort for capturing 360° panoramic images, which involves photographing and stitching together images at multiple viewing directions. Furthermore, each viewing direction is photographed multiple times with varying exposure levels to obtain the sufficient dynamic range. This process often requires extra time and effort or professional capturing devices. The final quality of the rendered output can vary greatly based on the parameter set considering the resolution and dynamic range of RMs, the complexity of geometry and materials, and parameters for rendering. Optimising these parameters is an active and challenging goal in computer graphics research.

Previous work optimise the parameters for rendering while maintaining image quality defined by the HVS [46, 47]. The general framework of these studies is the following: a reference image is rendered by a sophisticated rendering algorithm with unoptimised parameters to capture a photo-realistic result. A test image is then rendered with perceptually optimised algorithms or parameters. The test image is then compared with the reference image to evaluate whether the optimisations have maintained image quality. The measure here is the *visual equivalence* (VE) between the images, where images have VE if they convey the same impression of the scene appearance, even if they are visibly different.

This method can provide rendering parameter thresholds if the reference image can be provided. However, in the case of composition in visual effects and mixed reality, the reference image usually does not exist (e.g. synthesising aliens in a photograph, where aliens do not exist in the photograph as a reference). We present a new perceptual user study to better replicate image composition in mixed reality by removing the reference image. This results in the participant using the scene around the synthetic object as a reference. The aim of the study is different from previous perceptual studies because we measure *seamlessness* of the composition rather than the VE between two images. Using the self-referencing test, we conducted a series of user studies to capture the perceptual range of seamlessness in the HVS. Test images are generated by IBL with HDR RMs that

emulate the photographed real-world scene. By manipulating the RM, we adjust lighting parameters such as light direction, intensity, and bit range in various environments. The directionality and intensity are common parameters in abstract light sources, making them important attributes to test. The bit range is important since it is a common property found in RMs, where some databases of RMs are only distributed in LDR [45], whereas HDR RMs are often used for rendering [27]. The changes are applied to render synthetic objects for composition. The composited output images are then shown to the participants, where they attempt to detect any abnormalities of illumination in a single composite image. Based on the analysis of the user experiments, we investigate which artefacts are salient to the HVS for perceiving seamlessness in composition. Further, we run our tests on different geometry and material properties, but limit them to focus our experiment on illumination changes.

Based on our observations from these tests, we explore the effects of optimising the HDR RMs in two aspects. First, we conducted a perceptual test using RMs of low dynamic range images (LDRIIs), high dynamic range images (HDRIIs), and a synthesised HDRI using inverse tone mapping from the LDRI [48]. Our user study perceptually verified that the inverse tone-mapping algorithm can reconstruct the dynamic range from the LDRI that is required for the IBL within the threshold that the HVS cannot perceive. This may support the argument that an inverse tone-mapping algorithm can save the time and effort that is required to capture a HDR RM in some threshold. Furthermore, we reduce the resolution of the RM based on the fact that the HVS is insensitive to changes in light direction and intensity. Our results show that the size of the RM can be reduced significantly while maintaining the seamlessness of the final composition. This reduction directly results in memory savings, with an average of 99% savings of a general studio setup. Because image processing tasks operate on pixels, a lower resolution RM allows for image processing to increase in performance, which has an implication for processing time in Chapter 4 and 5. From the user study, certain key observations were made on what aspects of the rendered scene participants picked up on. The features of the illumination space capture the important elements described by these observations, such as shadow details and specular highlights.

3.2 Related Work

Illumination composition: Compositing synthetic objects with a real photograph has been an active research topic in computer graphics and augmented reality. Although previous works related to tracking and registration need to be fully considered, due to the large amount of literature, we will focus only on the works mainly contributing to illumination and rendering for realistic composition. Composition of a synthetic object into photographs has been a challenging task as research proposed by Nakamae et al. [49]. The general procedure for the task is a differential rendering technique presented by Fournier et al. [50]. In order to render 3D objects for image composition, we need to know the real-world light information. There are a variety of ways to capture real-world lighting in an environment for rendering, all ranging in cost of setup time and complexity [51, 52].Debevec [53] demonstrated a method of image composition that gave the synthetic object natural illumination. This method made use of HDRIs [27] to capture the radiance of the environment, giving realistic lighting and shadows for the synthetic object. Advancements in this area can be found in [24] for composition, and [26] for HDR capture.

Perceptual study in this domain is limited and has recently gained interest. Lopez-Moreno et al. [54] measured thresholds of the HVS where an object under different illumination was recognised as being incorrectly illuminated, with a focus on illumination direction, object orientation and spatial frequency of textured objects. Karsch et al. [55] proposed a user friendly, perceptually plausible composition technique requiring little scene information. They conducted a user study to demonstrate that their compositions are not easily distinguishable from real scenes.

Visual perception for visual equivalence: Perceptually based rendering increases efficiency while maintaining perceivable image quality. This approach exploits the nature of the HVS [56, 57]. Ramanarayanan et al. [58] introduced the notion of visual equivalence (VE), which has been considered in other works [59, 60]. Images have VE if they convey the same impression of the scene appearance, even if they are visibly different. They investigated this idea by testing how materials, geometry

and illumination interact and their relative impact on the HVS. They then presented the visual equivalence predictors (VEPS) as a metric for HVS. Krivanek et al. [47] conducted user experiments to investigate the relationship between the rendering parameters, object properties, and image fidelity.

Perceptually based rendering and optimisation: Perceptually based rendering involves optimising a rendering system while satisfying the HVS. This avoids the one-to-one simulation of the physical world, by only making computations that are necessary for image detail that the HVS is able to perceive. Yu et al. [46] conducted a user study, where their results allowed them to find which approximations are perceptually acceptable for radiosity rendering. Nakano et al. [61] conducted a study for perceptually correct shadows, testing image resolution used for lighting. For their test, the study was limited to interior office lighting and a dodecahedron shape. Their test was with reference against a photograph to identify which scene had more natural illumination between the photograph and synthetic scene. Alternatively, we provide an extensive, robust and generalised study of varying scenes, object shapes, and materials designed for lighting composition - showing that these different properties lead to different approximations.

HVS for seamlessness: Research that conducts perceptual studies without explicit reference images is limited. Vangorp et al. [62] conducted experiments which illuminated two synthetic objects within the same scene. Their research conducted similar experiments to ours in which there is no reference image but with a different focus. The goal of their study was to observe how shape and material variances dictate perception, but was not to investigate the perceptual study of light composition.

HVS for inverse tone mapping: IBL rendering usually makes use of HDRIIs for their RM. Landis [48] outlined a naive inverse tone mapping method of expanding the dynamic range of LDR images using a power function that is suitable for IBL rendering. Banterle et al. [63] presented an inverse tone mapping from LDRIIs to HDRIIs. They showed that the method is suitable for IBL rendering and evaluated it using the HDR visible differences predictor (VDP) metric [64, 65] instead of a user study. An extensive survey on LDR to HDR techniques for display and rendering has been conducted by Banterle et al. [66]. We performed user experiments to show

that expanded dynamic range of the RM is perceptually sufficient for composition.

3.3 HVS in Local Illumination Changes

Previous studies that observe the HVS in composited scenes are limited. A novel experiment setup for user studies is required to measure seamlessness in lighting composition. We make *local illumination changes* to composited objects, where the local illumination change refers to the altered illumination properties of the composited object (as opposed to the unaltered illumination of the other objects in the scene). In addition to the changes of light direction and intensity [67], we ran several extra experiments measuring the visibility of local illumination changes, namely the dynamic range of RMs, and the influence of rendering with a HDR image which has been converted from a LDR image. Furthermore, we applied these results to meaningful parameter optimisations, detailed in Section 3.4 and 3.5.

3.3.1 Experiment Setup

The aim of our experiment is to observe human perception relating to local illumination changes without reference images. The images of the background scene are captured by a fish-eye-lens camera at multiple exposure levels and angles. Then, the images are stitched together to form a 360-degree 16-bit HDR RM with a resolution of 5024 x 2512 (a common resolution in film production). Using the HDR RM, the background scenes are rendered by IBL rendering using path tracing [68, 69]. The background scenes shown in Figure 3.1 (left column) mimic the photograph in our test. Two synthetic objects are lit under the modified RM to simulate improper lighting in composition. Then, they are composited into the background scene using differential rendering [27]. In a preliminary user study, we used a large image set and a small number of participants to narrow down the test set. After finding the threshold range to focus on, we increased the number of participants and decreased the number of test images to collect more data on a meaningful range. The rendered images are a resolution of 1280x720. An example

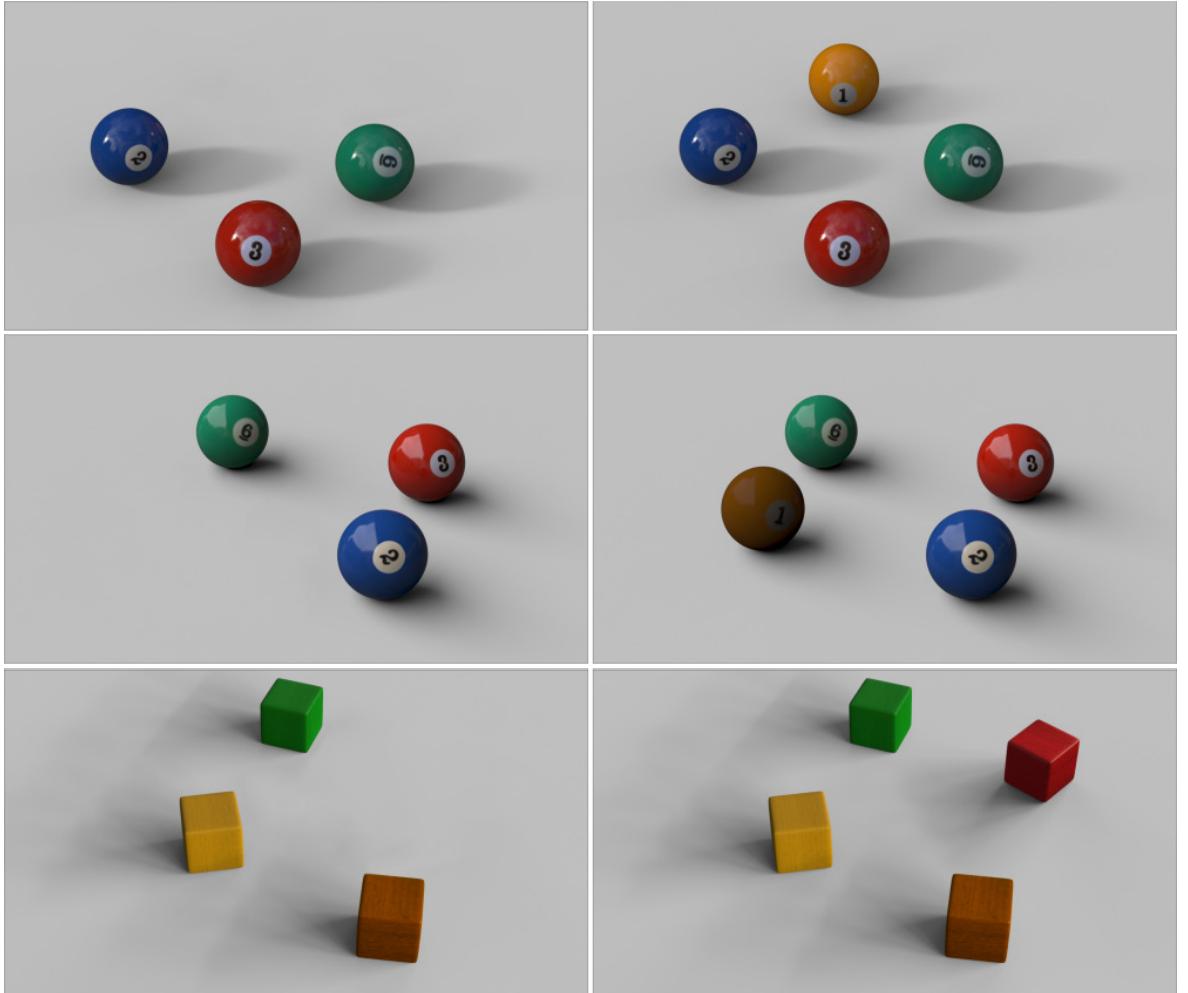


Figure 3.1: Examples of RM modifications in the user study, where the left column is an IBL rendering with the original HDR RM - this acts as a simulation of a photo. The right column has a synthetic object composited into the photo under improper illumination. The synthetic object of the first row is illuminated with an 8-bit RM, the second row's RM has its intensity reduced, and the third row modifies the direction of the RM.

of a composition with a modified parameter of an IBL can be seen in Figure 3.1 (right column).

3.3.2 Stimuli

In order to focus our test on the illumination changes, we limit the object geometry and materials. Each test image includes four synthetic objects having similar shape and topology. Objects include billiard balls and wooden blocks. The cubes were modelled as wooden blocks, with mostly diffuse material. Billiard balls were modelled with a glossy-specular material, since specular components are also an important source of information for visual perception [70]. Previous research has shown that spheres are the least discriminating shape with respect to material properties [62], which influenced our choice for using spheres in conjunction with a shiny material type. Our choice of cube is based on the fact that it contributes a more complex shadow than the sphere.

The tests are subdivided into different lighting environments. Three HDR RMs were created to represent real-world environments:

- **Sunny RM:** high intensity, outdoor scene with dominant sun light.
- **Overcast RM:** outdoors, under partial shelter during wet weather and an overcast sky.
- **Indoor RM:** indoor environment with multiple light sources emitting from the windows and indoor artificial lights.

The RMs are shown in Figure 3.4. The tests are further split into categories based on the lighting parameters. The variation of the lighting parameters of the virtual object are simulated by the manipulation of the RMs:

- **8-Bit LDR:** The test objects are rendered with a LDR RM (Figure 3.1, row 1).
- **Illumination intensity:** The test objects are rendered with the different values of the intensity of the HDR RM, where intensity is defined as a scalar of the light intensity from the RM (Figure 3.1, row 2).

- **Illumination direction:** The test objects are rendered with the lighting direction changed in increments of five degrees across the azimuth angle, from the zero to forty-five degrees (Figure 3.1, row 3). The range is filtered based on the pilot user studies. Most people observe artefacts over the range.
- **LDR to HDR:** The test objects are rendered with an extended LDR RM into HDR. The details are explained in Section 3.4.
- **Multiple resolution:** The test objects are rendered with a lower resolution RM. The details are explained in Section 3.5.

3.3.3 Procedure

We produce the set of IBL adjustments for each category, which is a combination of two shapes, two materials, and three lighting environments. Since the diffuse material is tied to the cube shape, and the specular tied to the sphere shape, this is a total of 6 combinations of shape, material and lighting categories. For the RM itself, there were 10 direction, 6 intensity and 7 resolution variations. Furthermore, we tested an 8-bit LDR case as well as an LDR to HDR case. This is a total of 25 RM variations. Combining the RM variations and categories, this is a total of 150 images in the user study (see table 3.1). We also submitted one image per category that had no adjustments.

Table 3.1: Groups of stimuli in the perceptual study

Test type	# variations
Direction	10
Intensity	6
LDR	1
LDR to HDR	1
Resolution	7

The test images are shown to the participants without an explicit reference image. Because there are signs of consistent degradation between images (such as an increasing shadow angle), the order of images are randomised and include images

with no adjustments. The randomness and range of scenes and object types reduce the likelihood that participants can detect patterns. The participants were asked if they could identify the areas which have been adjusted within the test image, and how noticeable this adjustment is. The categories are labelled as follows: 1 is *not noticeable - cannot identify any adjustments*, 2 is *slightly noticeable*, 3 is *moderately noticeable*, 4 is *very noticeable*, and 5 is *extremely noticeable*, adapted from [71], and participants were asked which object has the modification. If the participants cannot correctly identify the adjustment in the correct part of the image, then we consider that they cannot distinguish the change in illumination. The images are viewed on a calibrated Dell U2212HMC monitor. Our user study is made up of 20 participants, all of whom are shown all images. All participants have normal or corrected-to-normal vision.

3.3.4 HVS Observation and Analysis

In this section, we discuss our observation of the basic characteristics of the HVS in image composition while varying light direction, intensity, and the dynamic range of the RM. The results motivate our next step to optimise the RM, as discussed in Section 3.4 and 3.5.

Illumination direction: We tested the HVS's ability to perceive light direction changes. Between the preliminary and final survey, light direction was investigated by rotating the RM from 0 to 45 degrees in 5-degree increments. The results are shown in Figure 3.2a, showing the range between 5 and 30 degrees. Based on the results, we observe that the HVS is insensitive to localised changes in illumination direction in a composited image within a certain range. For the light direction test, it increasingly becomes noticeable from 5 degrees for sunny scenes, regardless of the geometry or material. It is particularly noticeable from 15 degrees. The ball object is generally more noticeable than the block. For most scene types, it becomes increasingly noticeable from 30 degrees onward. These abnormalities are typically recognised by the angle of the specular highlight or shadow area. The experiments found by Lopez-Moreno et al. [54] is similar to our experiment, but their measure differs to ours in that they use the number of correct guesses instead of a numerical

quality value. Our results show a lower tolerance, where 30 degrees of rotation is the upper limit before becoming too noticeable, though the different measure makes this difficult to compare. The focus in Lopez-Moreno et al.'s experiments are lighting direction without a plane to capture shadow, which could explain the difference in tolerance.

Illumination intensity: We tested the ability of the HVS to perceive light intensity changes. It was simulated by reducing the overall intensity values of the RM in increments of half the intensity value. The results are shown in Figure 3.2b. Based on the experiment, participants could not observe any abnormality while decreasing intensity by 50% (we skip this range and display the rest in Figure 3.2b). After 50% reduction, intensity changes show a steady increase in noticeability and are more apparent than the shadow changes. The intensity changes are predominantly recognised from the expected texture colour of an object, and is slightly more noticeable in the diffuse material.

8-bit LDR: We perform additional tests to perceive local illumination changes while altering the bit range of the RM from HDR to LDR. The original 16-bit HDR RMs are converted to 8-bit LDR RMs (with consideration for gamma correction). The LDR RMs are then used to render test objects. The results are shown in figure 3.3. Many participants perceived slight artefacts in the sunny and indoor scenes. The majority of people could not notice artefacts in local illumination changes rendered by the LDR overcast RM. In the overcast RM, the distribution of the intensity in the original HDR RM is mostly below the 8-bit range (see Figure 3.4). Therefore, in the case of overcast scenes such as cloudy and rainy skies, it is perceptually acceptable to use a LDR RM.

Summary: The user study shows the perceptually acceptable range of 3 different parameters; directionality, intensity and bit range. The parameters of the RMs are either typical in abstract light sources (directionality and intensity) or are commonly found in RMs (bit range). The results show that users were most sensitive to the intensity user study, whereas users were insensitive to the directionality and bit range of the RM. However, relative within each test, users were able to distinguish certain limits of noticeability. The bit range of a RM showed interesting results, where LDR overcast skies are deemed acceptable.

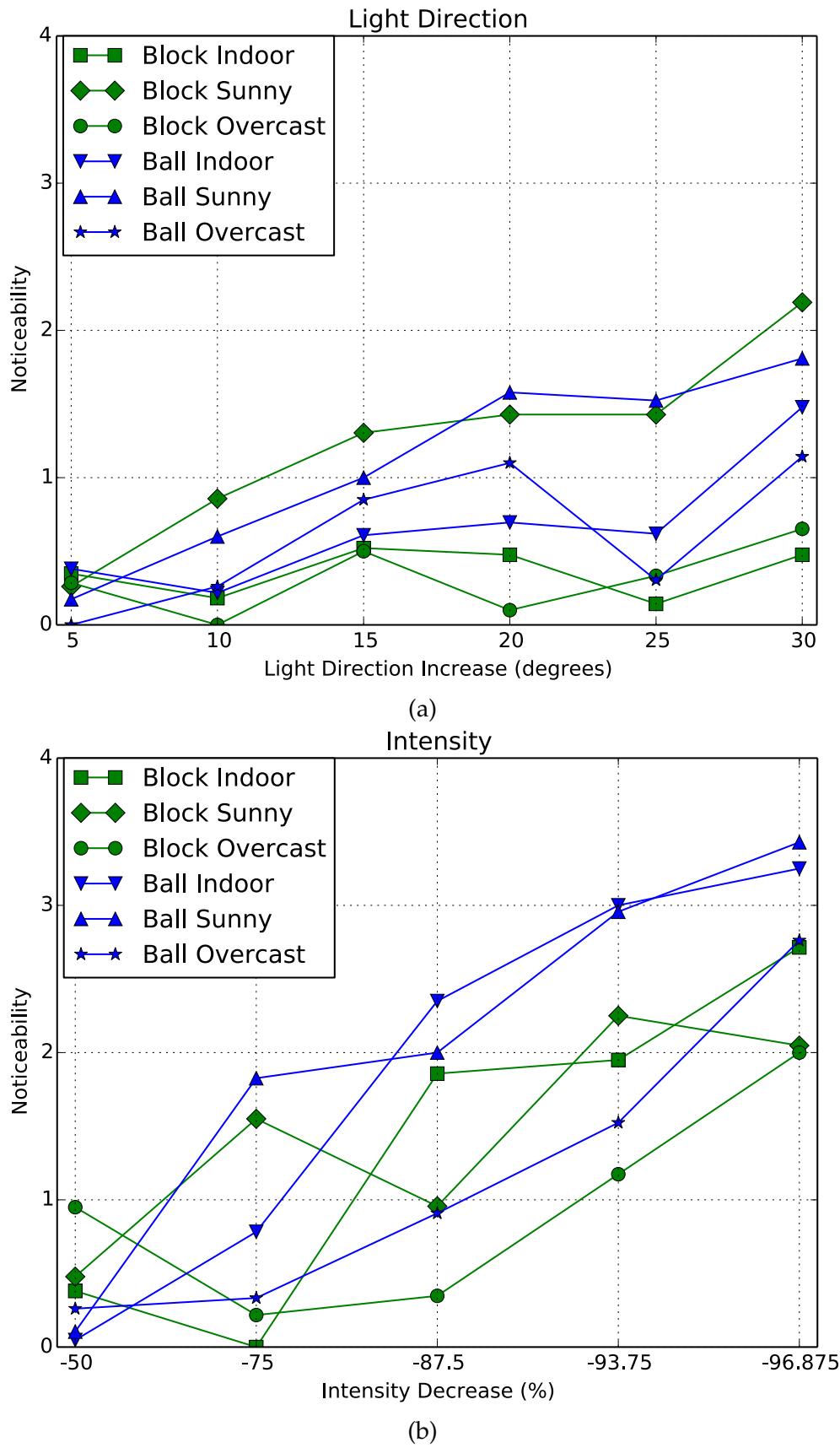


Figure 3.2: Observation of the HVS based on (a) illumination direction and (b) intensity. The noticeability scale is the mean score of the participants. The graph is focused in the range of sensitivity, beginning at the moment that the illumination details were detected by the HVS.

However, indoor and sunny scenes need to maintain HDR values in order to increase the contrast in intensity of the shadows and unoccluded scene. Furthermore, the HDR values ensure that the sharpness of the shadows remains intact.

3.4 HVS for Inverse Tone-Mapped Radiance Map

It is common to obtain LDR RMs from cameras, especially for video or live streaming where there is not enough time to capture the exposure time required for HDR images. While it is difficult to obtain HDR RMs, the previous experiments have shown that the HDR values of sunny and indoor scenes is required to maintain high quality lighting. However, there are algorithms for converting LDR to HDR, therefore, we can evaluate the possibility of using LDR data by first processing it into HDR and use the LDR to HDR converted RM for rendering. Similar to the previous user studies, we evaluate the quality of LDR to HDR converted RMs using our user study framework.

Procedure: We conduct a series user studies to evaluate an inverse tone-mapping operator (iTMO) that expands the dynamic range of an LDR RM. In this study, we utilised iTMO of [48], where the expanded luminance L_o is defined as

$$L_o(\mathbf{x}) = \begin{cases} L_\gamma(\mathbf{x})(1 - i) + L_\gamma(\mathbf{x})\epsilon i, & \text{if } L_\gamma(\mathbf{x}) \geq v \\ L_\gamma(\mathbf{x}), & \text{otherwise} \end{cases}$$

$$i = \left(\frac{L_\gamma(\mathbf{x}) - v}{1 - v} \right)^\alpha$$

where L_γ is the inverse gamma corrected function of the LDR gamma-encoded luminance value, ϵ is the expanded range, α is the exponent falloff, and v is a threshold value defining the starting luminance for expansion.

In this evaluation, we render the background scene with the original HDR RM. The original HDR is then converted into a gamma encoded LDR, which is converted back into a HDR using the iTMO. We render synthetic objects using the iTMO

converted RM. Then they are composited into the background scene. All RMs are of high resolution (5024x2512). The parameter ϵ is used to boost the dynamic range of the LDR RM. This value is approximated experimentally.

Observation and analysis: Results of the test are shown in Figure 3.3, which shows that the participants in the survey were generally unable to distinguish which object had been illuminated with the alternate RM. This shows that the iTMO produced a sufficient composite that seamlessly blends with the rest of the scene.

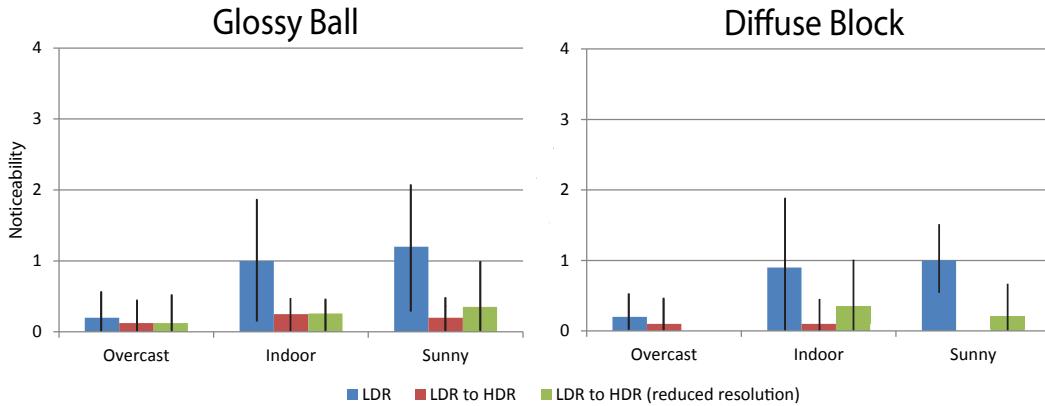


Figure 3.3: The user study results of IBL using RM of an 8-bit LDR, extended LDR to HDR, and the extended LDR to HDR of optimized resolution.

Figure 3.4 shows the variation of dynamic ranges of LDR, LDR to HDR using the iTMO, and the ground truth HDR IBL. We observe that the dynamic range in the overcast scene is quite limited, thus the LDR image is sufficient in capturing the lighting information. The indoor scene has multiple high intensity light sources, thus making it necessary to capture the illumination properties using a wider dynamic range. Further, we see differences in illumination from the various light sources (the immediate light bulbs contrast with the dim light coming through the window). The iTMO matches similarly to the ground truth in this case. Finally, the outdoor scene has a clear sky with a distinct light source (the sun). The dynamic range is very wide, and the iTMO closely matches this range - though some of the detail is lost, where the peak is flattened compared to the HDR image which has a distinct peak. This is due to the fact that the LDR image has a wide area of the sun capped at 255, thus expanding this range maintains the flattened peak. The reflection in the window has the dynamic range boosted as well, causing a second

peak (which does not match the ground truth). This is because of the same problem - the values in the reflection are capped at 255. Whereas the wider dynamic range can differentiate between high intensity values in the reflection of the window, and the high intensity values in the sun. Though these artefacts are not perceivable in the final rendering (as shown by the results in Figure 3.3), they should be carefully considered when working with LDR RMs.

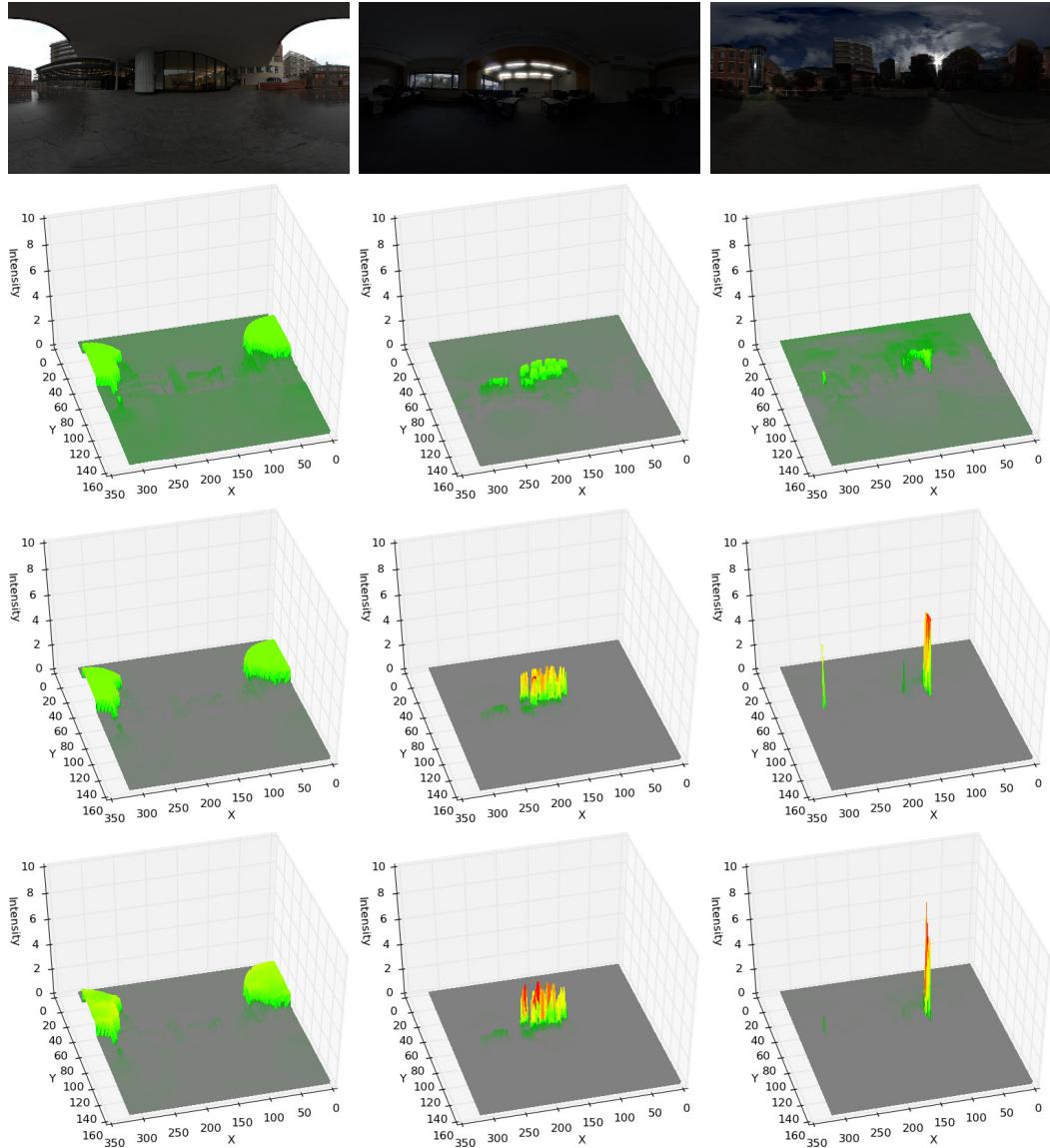


Figure 3.4: The dynamic range of each scene. Column 1, 2 and 3 is the overcast, indoor and sunny scene respectively. Row 1 is the RM, and rows 2, 3 and 4 is the dynamic range of the LDR, LDR to HDR, and HDR respectively. The LDR luminance values have been normalised between 0 and 1.

3.5 Perceptually Optimised Radiance Maps

Based on our observation in Section 3.3 and 3.4, we conduct more studies to find the perceptual threshold to optimise the resolution of a RM. Since we found that there is some threshold that can be exploited with direction and intensity changes, reducing the RM has similar effects. For example, a small RM will create offset shadows or narrow shadows. We conduct the same experiment as outlined in Section 3.3 to find optimal thresholds for the resolution of the RMs. We produce 7 variations of each category with a reduced RM resolution, and one extra image with no adjustments in each category. This is a total of 48 images to run the experiment with. We had a total of 30 participants in the survey for the test. Based on the results shown in Figure 3.5, we chose the RM resolution that was optimised before the artefacts became noticeable. The resolution 320x160 still maintains the same visual quality as a RM with a resolution of 5024x2512. This is a saving from approximately 43mb to 183kb. We also find that diffuse objects except the sunny day can be further reduced down to 80x40 resolution.

As shown in Section 3.3, reducing dynamic ranges of a RM from 16 to 8-bit cause significant down grading in the seamlessness of image composition. However, the iTMO nicely reconstructs the dynamic ranges of LDR RM as shown in Section 3.4. Based on the results, we performed a final user study to find the optimal resolution ranges for the RMs which is reconstructed by the iTMO. As shown in Figure 3.3, most participants could not notice abnormalities in the final IBL rendering and composition using the reconstructed LDR RMs by the iTMO for an RM resolution of 320x160. The results will contribute by saving effort in capturing iTMO RMs instead of HDR RMs, as well as saving memory to store the final RM.

3.6 Composition with a Photograph

The results are tested with real photographs. We have applied our observations and thresholds to optimise the RMs. The RMs were then used to illuminate the synthetic objects, which we then composite with the real photographs as shown in

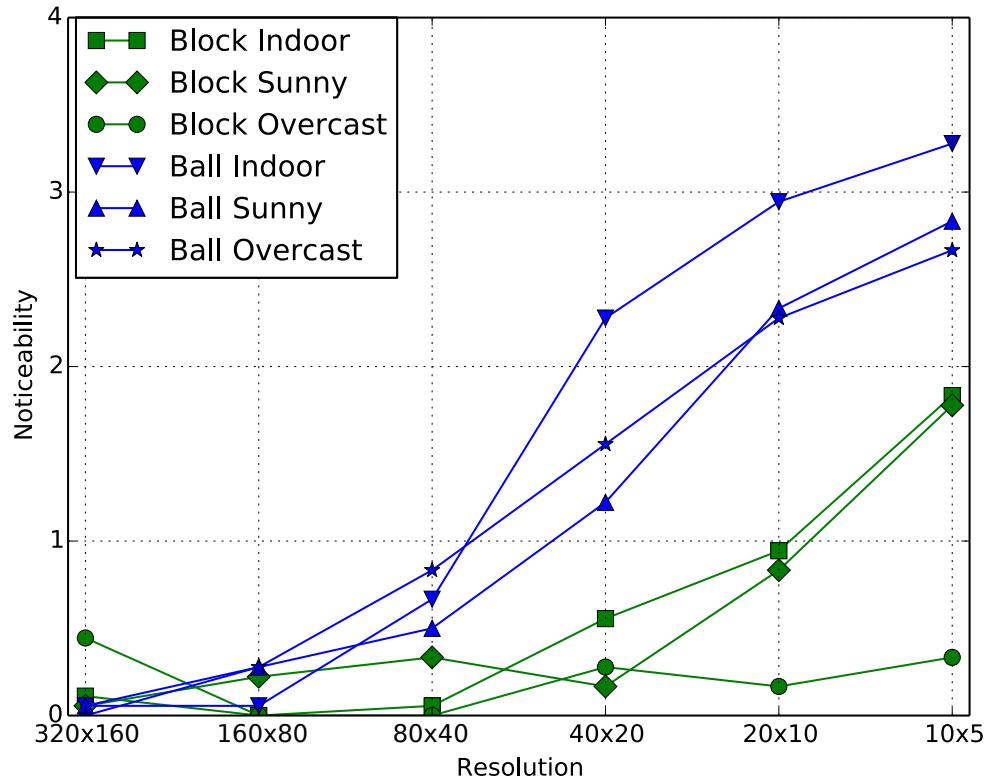


Figure 3.5: Results of the survey with respect to RM resolution. The noticeability scale is the mean score of the participants, and each category represents a decrease in resolution.

Figure 3.6, 3.7 and 3.8. An example with other synthetic objects is also shown in Figure 3.9. The background scene consists of a mixture of wooden blocks and billiard balls spanning diffuse to glossy-specular objects. The photographs of the background scenes are taken in the same lighting conditions as the HDR RMs. The synthetic objects are rendered using the original HDR RMs, HDR RM in the optimised resolution, and LDR RM created by iTMO. Then, the local illumination changes are composited into the photograph using differential rendering. The HDR RM is the only light source in the scene. This makes the task more challenging, because artificial light sources would usually be added to match shadows and highlights in production setup. Therefore, even in IBL rendering with the original full resolution HDR RM, some artefacts are slightly noticeable in the composition with strong sunlight (for example, Figure 3.8). In most cases, the result is visually equivalent to the IBL rendering using the original HDR RM.

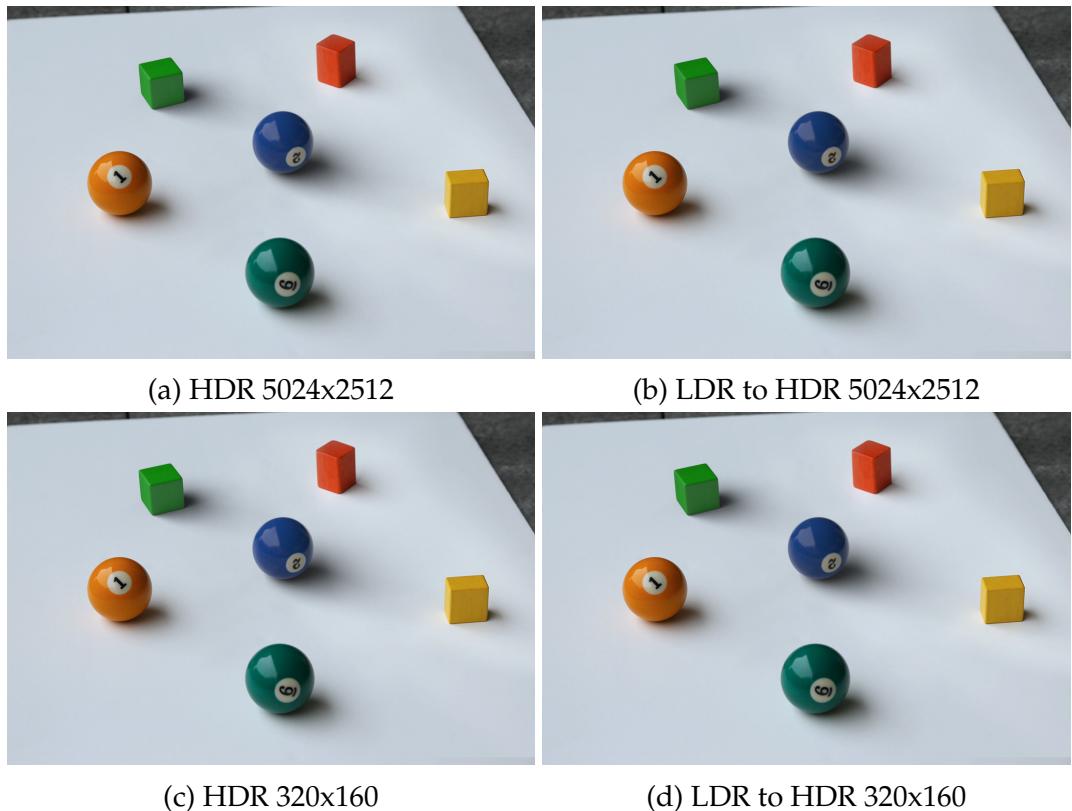
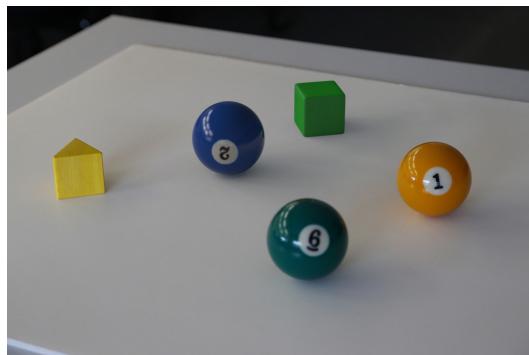
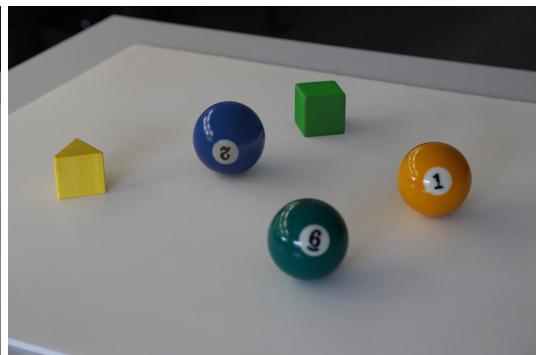


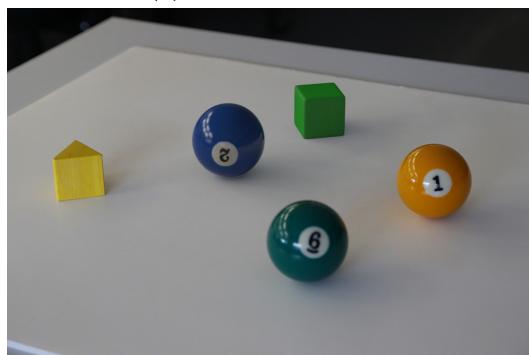
Figure 3.6: Overcast RM lighting synthetic objects (green block, blue ball) composited into a photo with varying resolution and bit range. A seamless composition is achieved even with a low resolution LDR to HDR RM.



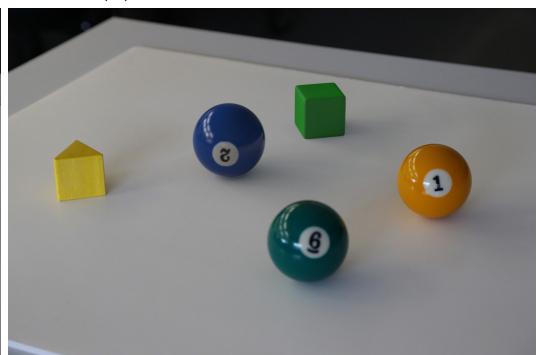
(a) HDR 5024x2512



(b) LDR to HDR 5024x2512



(c) HDR 320x160



(d) LDR to HDR 320x160

Figure 3.7: Indoor RM lighting synthetic objects (green block, blue ball) composited into a photo with varying resolution and bit range. A seamless composition is achieved even with a low resolution LDR to HDR RM.

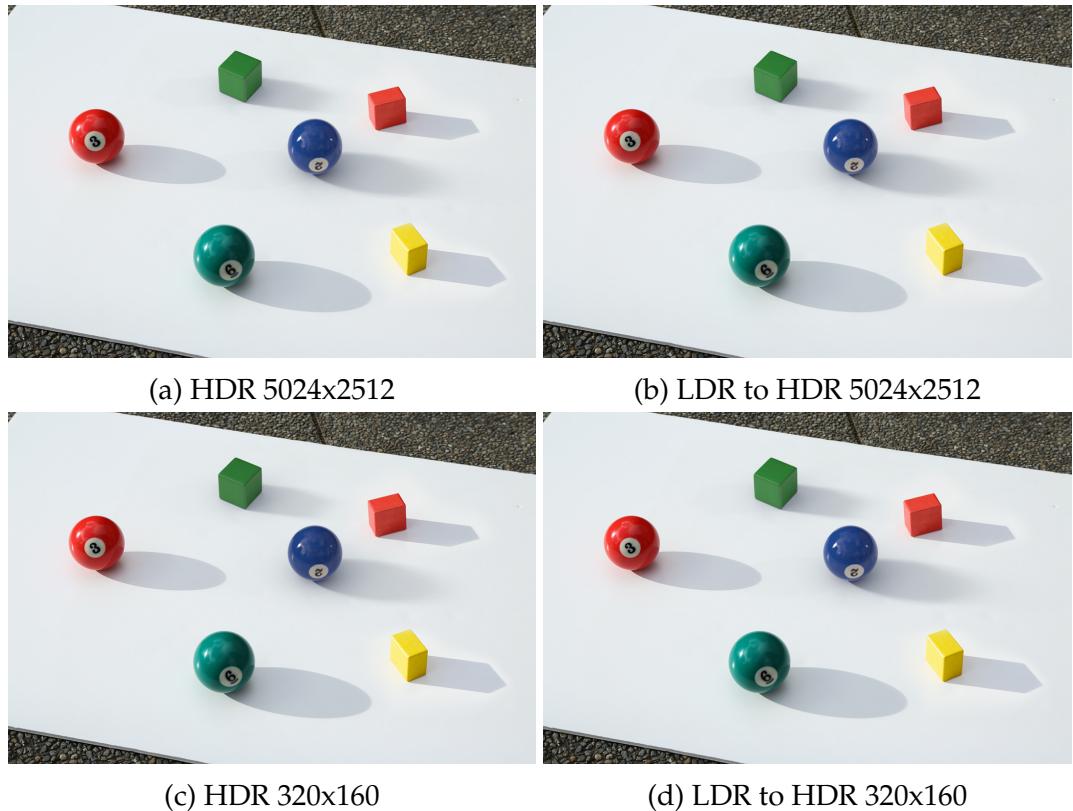


Figure 3.8: Sunny RM lighting a synthetic objects (green block, blue ball) composed into a photo with varying resolution and bit range. The artefacts in the shadow is due to the fact that directly capturing the sun is difficult, even with many exposure levels. However, the lighting remains consistent in quality between each image.

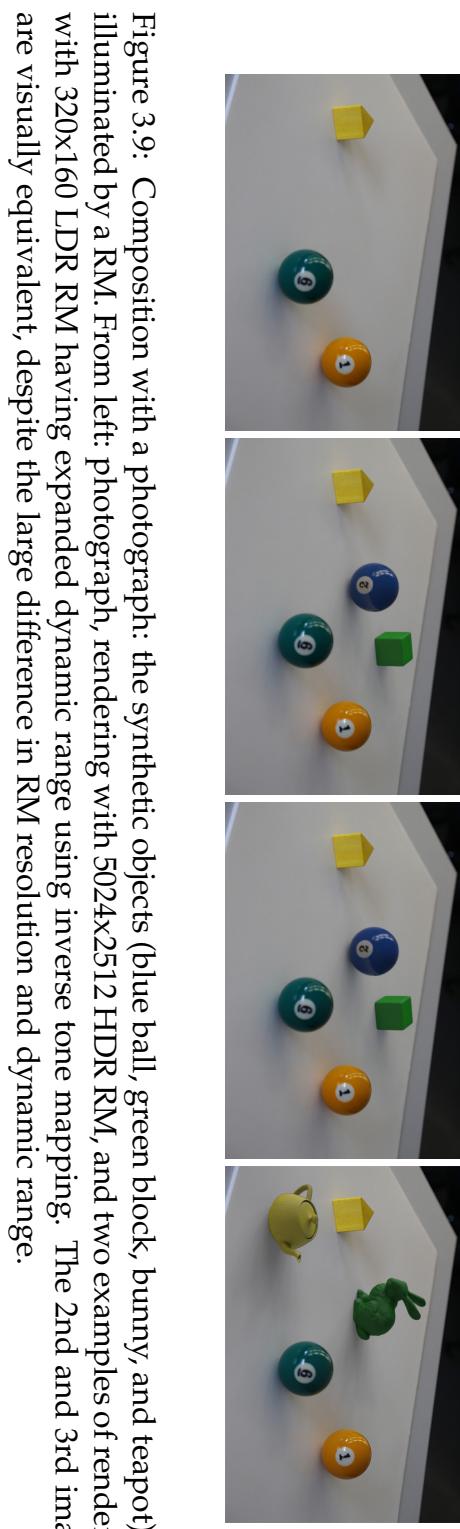


Figure 3.9: Composition with a photograph: the synthetic objects (blue ball, green block, bunny, and teapot) are illuminated by a RM. From left: photograph, rendering with 5024x2512 HDR RM, and two examples of rendering with 320x160 LDR RM having expanded dynamic range using inverse tone mapping. The 2nd and 3rd images are visually equivalent, despite the large difference in RM resolution and dynamic range.

3.7 Conclusion

The aim of this chapter is to observe the HVS when perceiving the seamlessness of image composition. This is achieved by performing a series of user studies without a reference image. This observation lead to further experiments to quantify perceptually optimised thresholds for reducing resources of IBLs with respect to both dynamic range and image resolution. Our results show that the inverse tone-mapping algorithm can accurately reconstruct the dynamic range required for IBL rendering in image composition in the range of diffuse to glossy-specular material types. Furthermore, we quantify the threshold of the optimal resolution of the RM. Our results contribute to the time and effort it takes to capture HDR RMs as well as the size of it while maintaining the visual quality of the final composition. We approximately save 99% of memory for a RM used in the general production setup, which is important in areas such as mobile graphics and image processing of RMs. The new perceptual test framework can be extended to identify optimal ranges of other parameters in future research. Future research can also explore a wider set of shape, material and lighting conditions, as well as vary the camera distance from the objects.

Our results are tested with real photographs and showed that the quality of the final IBL rendering using our optimised RM is visually equivalent to the IBL rendering using the original HDR RM. In the test, we explicitly used the RM as the only light source in the scene. Therefore, even in the IBL rendering with the original HDR RM, some artefacts (e.g. shadow edge and specular highlight) may be noticeable in the composition with strong sunlight (Figure 3.8). In film production, this is often solved by artists by adding artificial lights and tuning them until they achieve satisfactory results. However, the goal of this thesis is to automatically extract features on large databases of RMs. In this thesis, we focus on HDR databases which do not have such problems. However, this chapter illustrates that it is possible to process LDR databases as well, though some error may be introduced. This thesis will focus on HDR databases, but LDR databases show promising results, and is left for future work.

The results of this chapter motivate our research direction in Chapter 4 and 5. From

the user study, participants often reported artefacts by looking at changes in the shadow detail of the scene, especially in sunny scenes. This is as opposed to using the specular highlight as a guide. Therefore, the illumination space is constructed by focusing on features which describe the shadow properties of RMs. While the participants focused on the shadow details, reflective components should still not be ignored altogether, as there were still some instances of participants using the specular highlights. The reflective components are also more important for mirror like reflections. To begin the feature space, the texture space is explored in Chapter 4. Finally, the features are fully realised in Chapter 5, where the illumination features are dictated by the shadow properties in the rendered image.

Chapter 4

Sky Browser: A Reflectance and Texture Feature Space for Radiance Maps

4.1 Introduction

The RM is commonly used for illuminating synthetic objects. High intensity pixels act as strong light sources whereas dim pixels act as ambient light sources. As such, the image content and textural details in the RM (such as the photographed clouds, buildings, etc.) is not clearly perceived in the diffusely illuminated scene. However, specular reflections capture such details, this is particularly apparent in mirror surfaces which reflect the scene, such as the textural details of RM. The user study from Chapter 3 showed that the reflectivity and textural components of a RM remains an important cue for illumination. Furthermore, the textural details of the RM can be used separate from the rendering process; for backdrop composition. For this reason, it is important to compute features from RMs based on the textural details. While there is a vast literature on image content analysis in image processing and computer vision, there is little in the domain of image content analysis specific to HDR RMs. In this chapter, we extend the domain of image processing to HDR RMs. Since the sky is a common component of HDR RMs, we

focus on outdoor RMs, and leave indoor RMs for future work.

Skies are a well established and long-standing research area in computer graphics. Significant and ongoing research has addressed topics in simulation, rendering, and identification of sky images ([72, 73, 74, 75, 76] and others). RMs not only encode the lighting properties of the environment, but they also contain textural information, which is useful outside of illumination. For this reason, RMs are often interchangeably referred to as high-resolution HDR sky maps (HSMs). HSMs are frequently used in visual effects studio production for live action movies or real-time graphics, where the HSM is used as a backdrop to the scene, or is used to texture highly reflective (mirror-like) surfaces. Three of the common scenarios in which HSMs are needed are:

Scenes filmed on set: most scenes in a live action movie are recorded on a stage, rather than in a pre-existing natural environment. In these scenes, only the actors and foreground objects will remain in the final image. Distant objects are generated with computer graphics, physical models, or matte paintings, and the sky may be obtained from a HSM. VFX artists must find sky images that have various desirable qualities as requested by the artistic supervisor.

Scenes filmed in nature: in this case, the video images already have distant landscapes and sky. Nevertheless, it is often necessary to replace the sky, for various purposes. For example, it may be necessary to provide continuity with a different scene that was shot at the same location. Frequently, the supervisor may request a sky with a different appearance.

Virtual scenes: in this case, the background is a skybox to a completely digital scene. This is can be found film, but is commonly seen in video games since the skybox is an inexpensive way of rendering distant scenes - an important property for real-time graphics. The artist can easily swap out the skybox, but needs to pay consideration on whether the skybox matches the lighting in the virtual scene.

An example of a composition filmed in nature is shown in Figure 4.1. While the illumination space sorts the RM database based on the illumination properties, the high frequency texture information of HSMs still needs to be accounted for in cases where the specular reflection or background details are considered important. In

this chapter, we are focusing on complementing the illumination space with a texture space for HSMs.

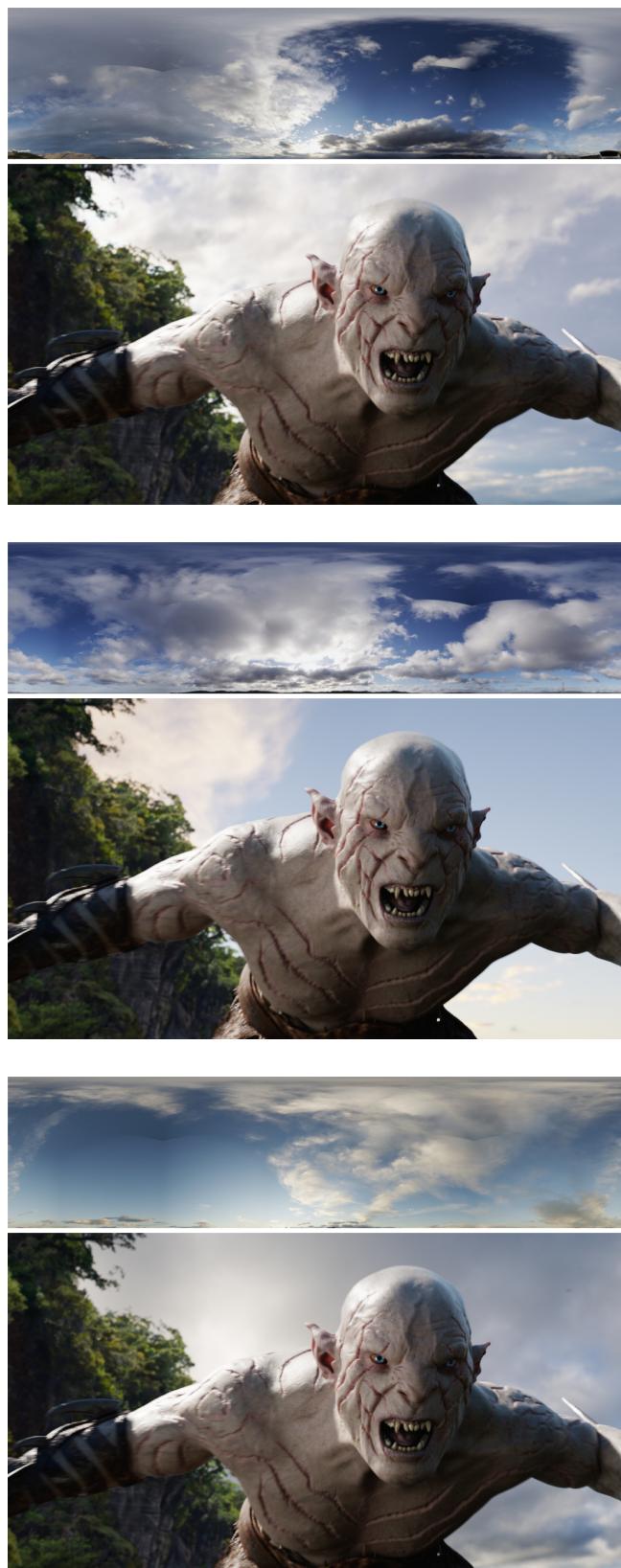


Figure 4.1: Three examples of backdrop composition. The synthetic object (Orc) is composited into a live action background. While maintaining the same illumination of the Orc, the distant backdrop (sky) is changed to adjust the mood of the scene. The sky colour has been modified for the composition.

Given this desire for alternate HSMs, a large database is needed for providing a wide range of images that differ in appearance. The problem arises when the number of HSMs in the database is overwhelmingly high, such that it then becomes highly unlikely that the ideal image is ever seen. If desirable images cannot be found, then it renders the entire database of images redundant. Previous work in this domain typically uses non-spherical low dynamic range (LDR) images of skies and classifies the images using labelled data. We focus our work on defining a set of features without labelled data, as well as including features suitable for HDR images.

We propose a system that allows the artist to intuitively navigate the space of images, assisting in finding the ideal image. Similar to the illumination space, this includes the ability to use a given image as a query to find another image in the database. The search function finds certain qualities of an image, for example, a clear sky with occasional fluffy clouds, or an overcast sky with large, dark clouds. The search function can also take an input HSM, and search the database for images that range in resemblance from very similar to completely different.

The criteria is difficult for a human to verbalise, and similar images are probably rated differently by different people. For example, the supervisor may request clouds that are "ominous" or "peaceful". How these words relate to images is subjective. For this reason, it is not effective to manually tag various images in the sky database with descriptive words such as "wispy, fluffy, peaceful, angry". This also means that we are not able to apply supervised learning methods to the search problem - the labels in a training set would be both subjective and hard to define.

For these reasons we formulate our problem as one of providing a feature space in which distance reflects the visual similarity of the images. The chosen features are low-level image and textural features. This side-steps the issue of defining what various descriptive words mean for different people. Instead, the artist simply navigates across images in the feature space. Our system, *Sky Brower*, is now in use at a visual effects facility.

4.2 Related Work

The sky is a common component of many images, and selecting the right sky is important to suggest the time of day, weather and mood. Since Klassen [77] presented his work on sky visualisation, skies have been an important subject in computer graphics with many associated research publications [72, 73]. Research in these areas is ongoing, and we refer to [75, 76] for an entrance to this literature, and refer to [28, 22] on methods for capturing HDR environment maps.

Given an outdoor scene, changing the background with a better sky image is a common task in 2D image processing as well as visual effects for live action movies. The main task is to search for alternate images using an appropriate query. Generic content-based image retrieval (CBIR) methods can be used, but these systems rely on features (shape descriptors and interest points) that are not appropriate for clouds [78]. As well, many of these systems require supervised learning. Therefore, we only focus our survey on papers which are highly related to our main topic, the sky.

Proper labelling of an image or a part of it can guide searching in many applications [79, 80, 81]. However, verbalising the criteria is difficult for a human due to inconsistent meaning of subjective terms such as "peaceful" or "smooth". With this in mind, we formulate our problem as feature extraction and searching the feature space, where relative distance reflects the similarity of the images. Other methods to get around subjective labelling include crowd-sourcing, as found in [81].

The general method of the search function investigates machine learning and texture classification to define feature extraction techniques. Haralick et al. [82] describes easily computable textural features based on grey tone spatial dependencies. This is often referred to as the grey level co-occurrence matrix (GLCM). Gu et al. [83] compares techniques for measuring cloud textures. They use GLCM features to measure spatial properties, where they found that entropy based features gave good results for frequency properties. Chethan et al. [84] consider textural features based on the Gabor transform to classify clouds, as well

as using a support vector machine (SVM) as their method of classification. Mazzoni et al. [85] label parts of images as a clear sky, or a type of cloud. They use the Multi-angle Imaging Spectro Radiometer (MISR), an instrument used by NASA, to study clouds and aerosols. Heinle et al. [86] classifies skies into seven different categories. They used the k-nearest neighbour method for classification, and the colour and tonal variation of an image as features.

Recently, Tao et al. [74] developed an interactive search system for finding sky photographs using supervised learning techniques. Their method allows for offline computation as well as an interactive user interface. Most similar to our work is [87], which extracts four features to characterise the images in the database. Other recent works include [88, 81]. Our method fundamentally differs from these papers on four points: We focus on HSMs (an industry standard image category), we remove the ambiguity and manual labour of labelling data required for supervised learning, we explicitly define two textural properties for more artistic control over clouds, and we use the spherical harmonics as a novel tonal feature. These four components target a specific and important area of computer graphics and the movie industry.

4.3 Overview

The search is based on observable but subjective image properties. For example, such properties may include how blue the sky is, or how patchy the clouds are. There are tonal properties such as the contrast or intensity of an image, and there are textural properties, such as the bumpiness of a cloud. Further, skies tend to have strong properties relating to how blue, white or red the sky is, such as clear blue skies, large bright white clouds, or a red sunset.

Clouds come in a wide variety of forms that can differ depending on the atmosphere and temperature. We considered the possibility of using the scientific names of clouds as their class labels, and attempted to find features that could categorise them so. For our application, a major disadvantage of supervised learning is the requirement of labelled data. Human labelling of the sky images is

not only expensive but conceptually difficult as well. It is difficult to identify a set of labels (such as “wispy”, “romantic”, etc.) that are useful and consistently interpreted. Further, in our experience the desired labelling is simply not done in some cases.

Instead, we define a search space that does not require a labelled dataset, but that can be visually traversed with no prior training. This requires a set of features that capture perceptually relevant information while ignoring information that is not important or even imperceivable. In addition, we require a minimal set of features, in order to avoid the curse of dimensionality.

Unfortunately the number of possible features is large (it is some fraction of the number of possible programs that take an image patch and output a number), and choosing a best subset is not feasible due to the combinatorics. Choosing a small set of features therefore requires intuition and experience with the problem. We discarded keypoint features such as SIFT [89] because skies are more appropriately considered as random textures than as images of objects with common and reproducible parts. Instead we explored features such as GLCMs that have proven successful for texture modelling and classification.

After several months of experimentation on actual datasets, we chose the following four-dimensional feature space: *GLCM correlation*, the *entropy of the Laplacian* (EL), a *ratio between red and blue spherical harmonic coefficients*, and the *mean intensity*. The result is a small set of features that define a visually searchable space.

There are numerous advantages to this approach: Firstly, it does not require manual labelling of the data (or finding a readily available labelled dataset). Secondly, the classification method is not defined by scientific labels that would have to be learned. Instead the search space simply relies on visual perception of the images. Finally, images often fall between scientific labels (such as a single sky having two cloud types), so removing scientific labelled data gives more artistic freedom for defining a continuous feature space for skies. For example, the ability to move from “very patchy” clouds to “somewhat patchy”, to “not patchy at all”, while maintaining other key features, for example, “a very blue sky” in conjunction with the varying levels of patchy clouds.

4.4 Features

The following are the set of features we use to define our search space. The HDR images are in a linear colour space. The images can be extremely high resolution (up to 20,000x10,000 pixels). To reduce computation time, we introduce a pre-processing step where we scale the images to a standard resolution of 360x160. For the spherical harmonics feature, we reduced the input image to 512x256.

4.4.1 GLCM Correlation

The GLCM is a commonly used technique in texture classification [82]. The method involves finding a co-occurrence histogram of an image, and running various formulas across the histogram. The histogram counts how often two intensity values in a grey-scale image co-occur with some spatial relation ((dx,dy)), for example, the number of times that a pixel with value 5 is to the right of a pixel with value 20.

The GLCM histogram is a lot of information – potentially much more than that image itself, depending on how many spatial relations are considered. For this reason, various summary statistics are often used [90]. After experimenting with several GLCM summary statistics, we selected correlation since it is minimally correlated with a second textural feature (described in the next section). For a particular spatial relation, the correlation is

$$\sum_{i,j=0}^{255} P_{i,j} \left(\frac{(i - \mu_i)(j - \mu_j)}{\sqrt{\sigma_i^2 \sigma_j^2}} \right)$$

where i, j are indices of the GLCM corresponding to pixel values on an 8-bit scale, μ is the mean, σ is the variance, and $P_{i,j}$ is the GLCM histogram normalised to serve as a probability. The correlation characterises an approximate smoothness in a particular spatial direction. We used the eight nearest-neighbour directions as the spatial relations. The computation is accelerated restricting the computation to a number of windows evenly placed across the sky area in the HSM. Twenty-

seven windows are selected by our experimental tests on images with a resolution of 2048x1024. The windows are uniformly placed across the image (9x3).

4.4.2 Entropy of the Laplacian (EL)

While the GLCM correlation captures a type of roughness or smoothness, it does not say anything about the distribution of changes. To account for this, we introduce the EL. The Laplacian ∇^2 is an approximate scalar curvature measure. We form the histogram of the Laplacian values at all pixels. The entropy of the normalised histogram distinguishes whether the curvature is concentrated in a few values (low entropy) or takes on many possible values. The feature is computed as

$$-\sum_{i=0}^N P_i(\ln P_i)$$

where P_i are values of the normalised histogram of Laplacian values. The latter are computed by a standard finite difference stencil.

The GLCM correlation measure in combination with the EL can differentiate between clouds with difficult to describe textural qualities. Figure 4.2 shows an example of images distributed in the feature space. The images in row 3 are all coherent images in a sense, as described by their low EL measure. Yet the GLCM correlation measure separates the images; the image in row 3, column 3 is very smooth, and the image in row 3, column 1 has a lumpy texture.

4.4.3 Spherical Harmonics Ratio of Red and Blue

The two features defined above capture textural properties of an image. However, an artist also looks for images with certain tonal properties. Red and blue values are salient among skies, for example, it is often the case that skies consist of vibrant blue skies, red tainted clouds, or red sunsets.

As a starting point, we consider the ratio of the amount of red and blue in the sky image. As discussed in [86], the ratio of red and blue defines how much cloud is in

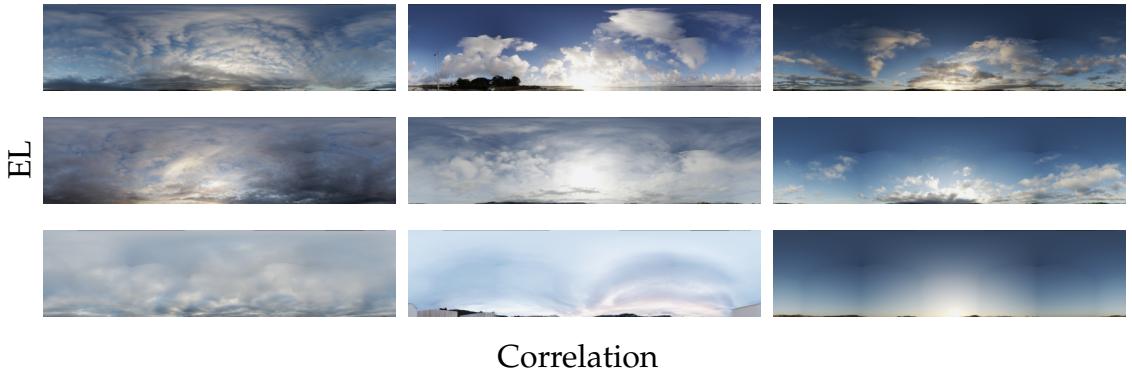


Figure 4.2: Textural feature space examples for clouds.

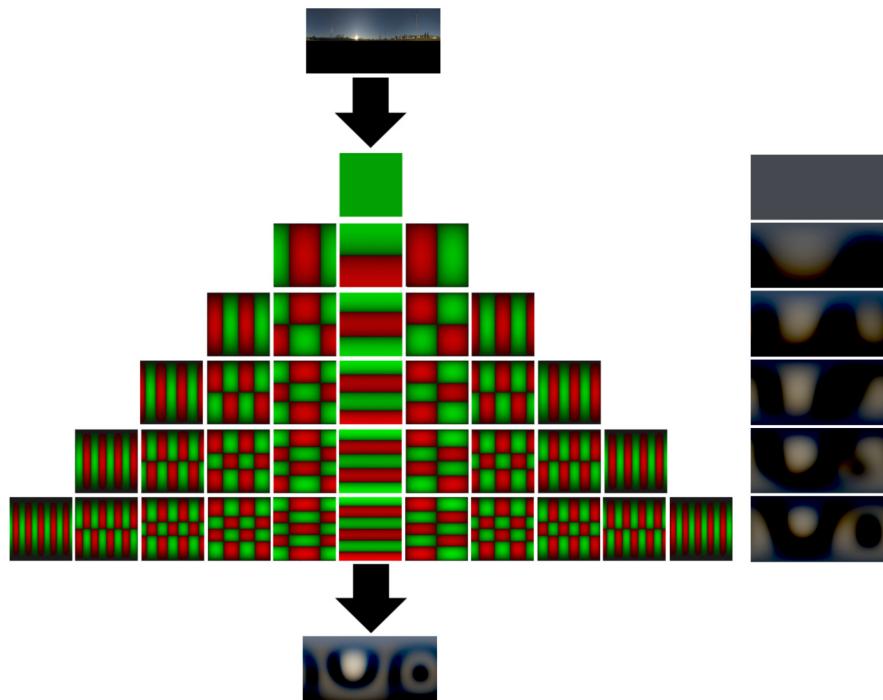


Figure 4.3: The spherical harmonics expansion. Above the expansion is the input image, and below is the approximation of the input. To the right of each band is the corresponding approximation of the input image.

the sky, so this ratio has the additional effect of defining cloud cover. Artists can increase the amount of red to find more clouds in the image, as well as increasing it further to find red skies or clouds. We found that green is correlated with the red coefficient in sky images, thus it did not add any useful information in the search

function. Furthermore, the ratio is independent of intensity. This is a desirable property as an artist can freely change colour regardless of intensity, and vice versa.

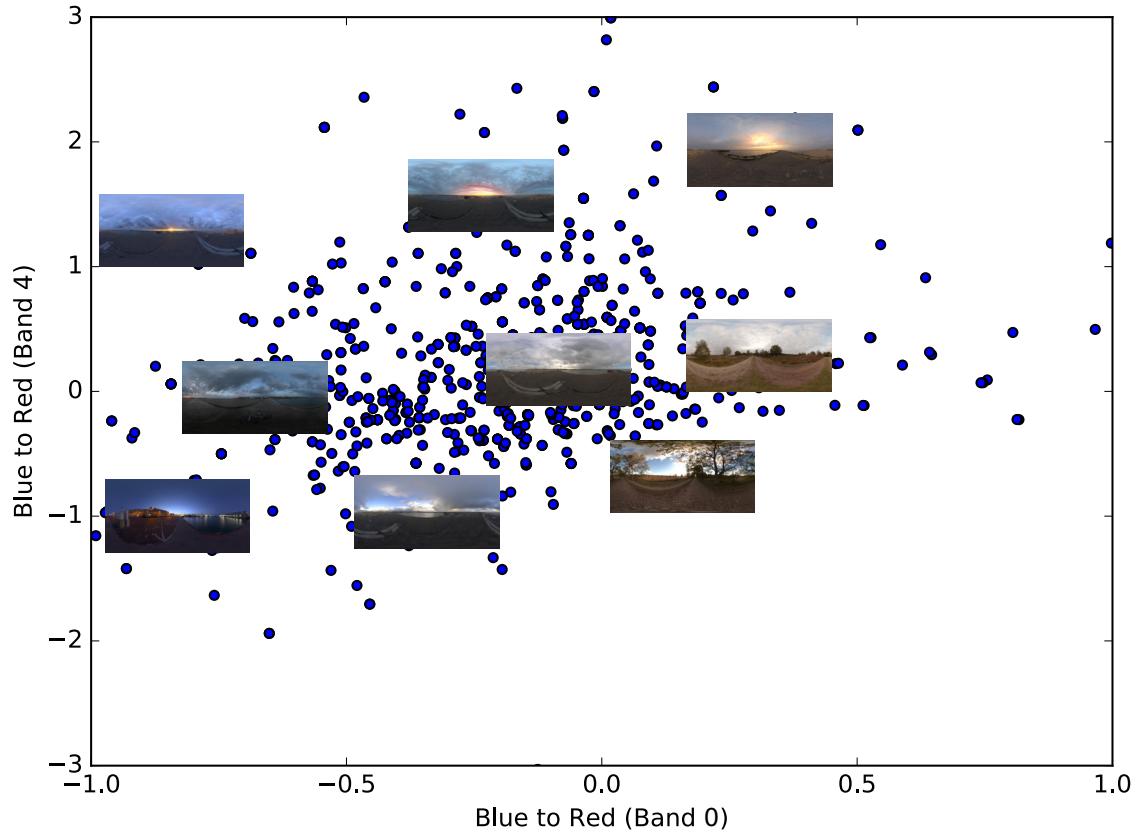


Figure 4.4: Textural feature space examples for tone.

The redness of an image is somewhat subjective however, as redness can be interpreted for the entire image, or around the sun. To handle this choice, we consider the problem in the spherical harmonic frequency domain, and allow the artist to freely choose low or high frequency tone (Figure 4.3 illustrates how the basis decomposes the RM into low and high frequency components). The spherical harmonics are an orthogonal basis on the sphere, where the real basis functions are

defined as

$$y_l^m(\theta, \phi) = \begin{cases} \sqrt{2}K_l^m \cos(m\phi)P_l^m(\cos\theta), & m > 0 \\ \sqrt{2}K_l^m \sin(-m\phi)P_l^{-m}(\cos\theta), & m < 0 \\ K_l^0 P_l^0(\cos\theta), & m = 0 \end{cases} \quad (4.1)$$

where P_l^m are the associated Legendre polynomials which returns a real value number in the range $(-1, 1)$, and K_l^m are the normalization constants

$$K_l^m = \sqrt{\frac{(2l+1)(l-|m|)!}{4\pi(l+|m|)!}} \quad (4.2)$$

where l represents the index of the band of polynomials which share the same degree, and m is the index of a specific frequency in that band (where m is in the range of $-l \leq m \leq l$). A RM is then projected into spherical harmonic coefficients

$$\text{coef}_l^m = \int_s f(s)y_l^m(s)ds \quad (4.3)$$

where coef_l^m is computed on each colour channel (RGB). At this point, two sets of coefficients on two RMs could be compared to derive their distance apart. Unfortunately, this comparison would require too many coefficients and would require the spherical harmonics to be rotated to find the best match. We use a similar approach found in [91], which they use summations across spherical harmonic bands for 3D shape matching. In our case, the tone feature is defined using the colour of the RM, and is computed by summing across each frequency band

$$F_l = \sum_{m=-l}^l \text{coef}_l^m \quad (4.4)$$

using a vector of F_l up to some specified band. Allowing the artist to specifying the band gives more control over directional ambience, capturing varying levels of frequency. As shown in [92], $l \leq 3$ is enough to capture diffuse lighting. In our study, we compute 5 bands to allow the artist to capture higher frequency information if they should choose to do so. In equation 4.4, l is defined by the user. This reduces the number of coefficients from l^2 for each colour channel to m for

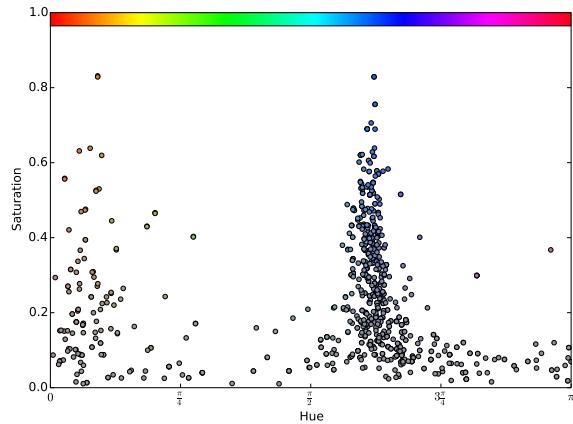


Figure 4.5: Ambient tone colour distribution showing that the HDR database is dense around red, blue and grey (low saturation) values. Highly saturated areas tend to represent bright skies, which often take on red (sunset) or blue (clear sky) values. Low saturation areas tend to represent cloudy skies. Indoor scenes typically show low saturation values with varying hue.

each colour channel. This also has the added benefit of removing the rotational dependency. Using different frequency levels gives results that change depending on the task at hand. For example, the lower order frequencies correspond to the overall ambient tone, whereas higher order frequencies correspond to the tone around high frequency lights (for example, the colour around the sun's falloff or corona). The first order spherical harmonic $l = 0$ corresponds to the arithmetic mean value of the RM.

Figure 4.5 shows a high distributions of skies with a red or blue hue on the first band, corresponding to sunsets and blue skies respectively. Furthermore, a low saturation value makes up the majority of the database, where these skies have bright white skies (white clouds or the sun) or are comprised of indoor scenes. It's often useful to reduce the complexity of a spherical harmonic coefficient from RGB to a scalar value computed as the ratio between red and blue

$$F_l = \ln F_{l_{red}} - \ln F_{l_{blue}} \quad (4.5)$$

However, this specifically applies to most natural scenes. If the database is comprised mostly of greenery in the upper hemisphere (e.g. forests), then a

different ratio will apply.

The first five bands gives the artist control over whether they are seeking the colour of the overall sky, or the colour of the bright region of the sky adjacent to the sun (the sun's falloff or corona). This result can be seen in Figure 4.4.

4.4.4 Average Intensity

Because the ratio of red and blue is invariant to intensity, we use the mean intensity of the image as the fourth feature. These four features (correlation, entropy, red/blue ratio, mean intensity) define the texture space that is navigated by artists.

4.4.5 Feature Correlation

Pearson's Correlation Coefficient shows that our features all have low correlation with one another (Table 4.1). GLCM Correlation and the EL act as our textural measures, which allows the user to define the variation of cloud types. We observe from Figure 4.2, 4.6 and Table 4.1 that there is weak-negative relationship with the two features, as the lower end of the GLCM correlation measure defines a lack of smoothness, which can be interpreted similarly as a high EL value. Given that the correlation is weak, the combination of these features is useful, as shown in Figure 4.2.

Table 4.1: Feature Correlation

	GLCM	Entropy	Intensity	SH
GLCM	1.00	-	-	-
Entropy	-0.36	1.00	-	-
Intensity	-0.24	0.17	1.00	-
SH	-0.05	0.11	-0.22	1.00
	≥ 0.0	≥ 0.1	≥ 0.3	≥ 0.5

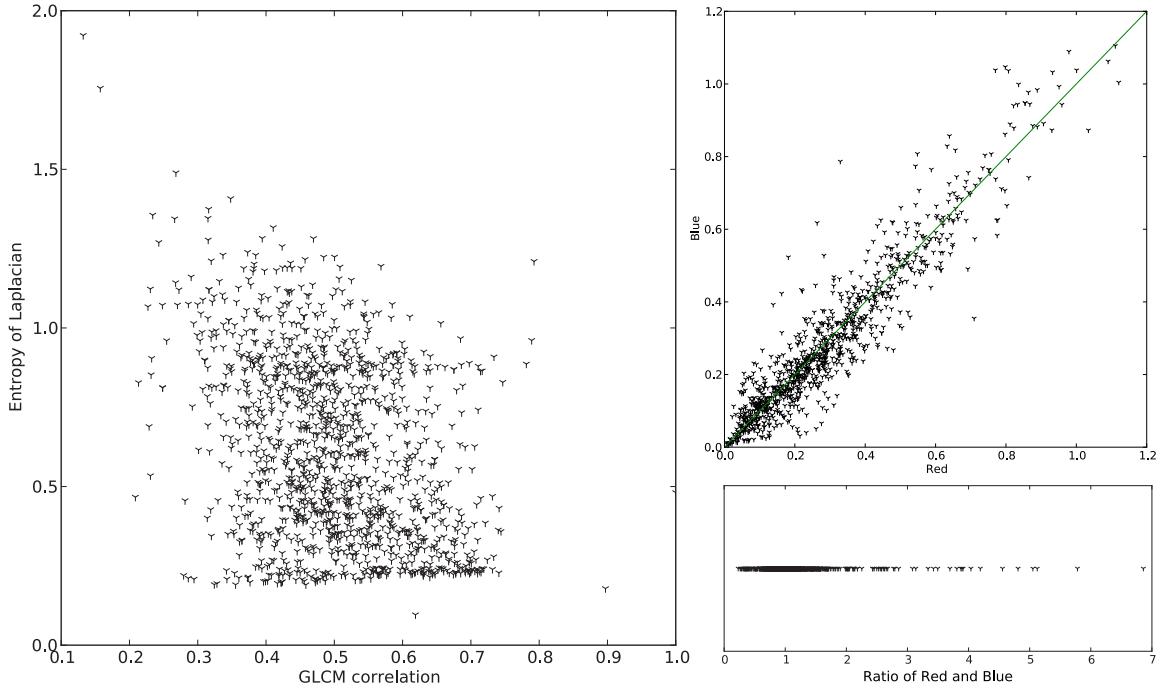


Figure 4.6: Feature space distribution, where each point represents an image. Left: the textural feature space. Right-top: the red and blue spherical harmonic. Right-bottom: the colour of the sky in one dimension by taking the ratio of the red and blue value.

4.5 Results

Figure 4.7 shows the exploration results of *Sky Browser*. It demonstrates a user searching through the feature space using the sliders. Using the parameter input variables from the sliders, we use a k-nearest neighbour (k -NN) approach to query the database. The current system runs at interactive rates with a database of 1300 HSMs. The search function is scalable using parallel processing [93]. The user can begin the search by using an image as a query which has properties that are similar to what they're looking for, as shown in Figure 4.7 (1st column), where the user has used an image to define the starting slider values of the features. Following from this, the user can adjust parameters to move towards their ideal images. For example, in Figure 4.7, the 2nd and 3rd columns show intensity and colour changes respectively, while maintaining the textural properties. The 4th column moves

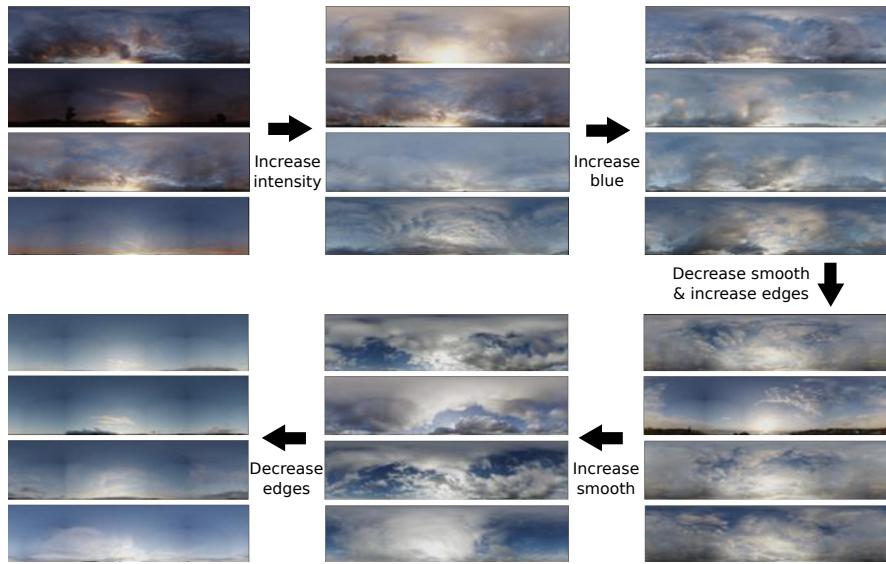


Figure 4.7: The search results (navigating the feature space). The transition sequence is as follows: increase intensity, slight move toward blue, decrease smoothness and increase edges, increase smoothness, and finally a large decrease of edges.

towards patchy images, whereas the 5th column maintains edges (EL feature) but increases the smoothness (correlation feature), returning large and apparent clouds. The 6th column maintains the smoothness, and removes edges, thus producing clearer skies. Figure 4.7 shows just one example of finding clouds. There are many other possibilities, for example moving towards blue in column 5's state can bring in smooth blue skies and smaller distinct clouds, instead of large smooth clouds.

Sky Browser searches HSMs in feature space without labelled images. Therefore, it is difficult to evaluate the result quantitatively. Instead, we conducted a subjective test with 12 professional visual effects studio artists. They were asked to use *Sky Browser* to find suitable images, both by using an existing image as a query and by interactively browsing using the features. Following this test, they were asked following question: "*Given the input, did Sky Browser return similar results?*", where their input would involve searching with an image as well setting the features manually. They answered with a score between 1 and 5, where 1 is *very dissimilar*, 2 is *dissimilar*, 3 is *uncertain*, 4 is *similar*, 5 is *very similar*, and were allowed to score in 0.5 increments. They were asked this question twice, testing the cloud (GLCM and EL) and tone (SH ratio and intensity) features separately.

Table 4.2: User Evaluation Results for *Sky Browser*

	Cloud features	Tone features
Mean	4.0	4.04
Std	0.21	0.50
Max	4.5	5.0
Min	3.5	3.0

Given the low standard deviation and high mean score, the user evaluation results indicates that the artists generally agreed that *Sky Browser* produced similar results to the input query. Along with the score, the artists were able to provide comments on the system. Through these comments, we are able to observe some of the limitations. Some artists noted that slider handles can be quite sensitive:

"It's like driving a race car, these sliders need a really soft touch, kinda fun clicking around on the images and seeing where the sliders fit in."

This may be mitigated by scaling the data around dense regions. Others noted that sparse regions of the database can produce unexpected results as well:

"The tool works pretty awesome, some stuff like very red sunrises etc. appear to give poorer results but I reckon that's more to do with the database size."

As such, a visual cue for sparsity in the database can improve expectations when browsing. Another limitation is that other cloud-like structures may appear in the search results, such as trees:

"Sometimes the browser would confuse trees for clouds or larger detail. I would expect it to find images with increased detail/contrast but without trees when the source image was without trees. Overall pretty good with keeping the general idea of the look even when moving toward more red/sunset or sunrise looks."

4.6 Conclusion

The focus of this chapter is on defining a minimal set of features capable of unsupervised classification. The main contribution are the set of features that describes the appearance properties that make up the *Sky Browser* application. The features are minimally correlated and thus define a search space that an artist can intuitively navigate. To navigate the search space, the artist has control over the features as parameters. These parameters are useful to describe images based on tonal and textural properties. A limitation of the present system is that larger slider movement is needed to navigate in areas that are sparsely populated with data, and fine slider movements can be too sensitive in densely populated areas. Future work may investigate user interface designs or scaling factors to address this issue. Our feature space for *Sky Browser* is defined based on the evaluation by professional visual effects artists, and additional features can be applied for particular purposes. While this chapter covers the textural component of RMs, it does not describe how the RM illuminates the virtual scene. In the next chapter, we will define the illumination space, which will capture these details.

Chapter 5

Illumination Space: An Illumination Feature Space for Radiance Maps

5.1 Introduction

The user study from Chapter 3 showed that the textural components of a RM remains an important cue for illumination, which is addressed in Chapter 4. Texture based features can provide concise semantic descriptions of images [81], but such methods are limited to describe what images look like rather than their illumination properties. As shown in the user study in Chapter 3, the illumination properties are important for defining perceivable shadows as well as the tonal ambience in the rendered image. Rather than clustering RMs which look visually similar, the required solution is to cluster and search RMs based on how similar their rendered images look. However, highly reflective materials are part of the rendered image, and are particularly important when matching the reflection to the background scene. Chapter 4 addressed this issue, whereas this chapter focuses on a compact representation of shadow information and tone of RMs, which can effectively cover diffuse to glossy materials.

Over time, many studios and individuals have collected a large number of RMs. LDR databases, such as the SUN360 database [45], have been used in previous studies for composition [25]. HDR databases are now readily available in post

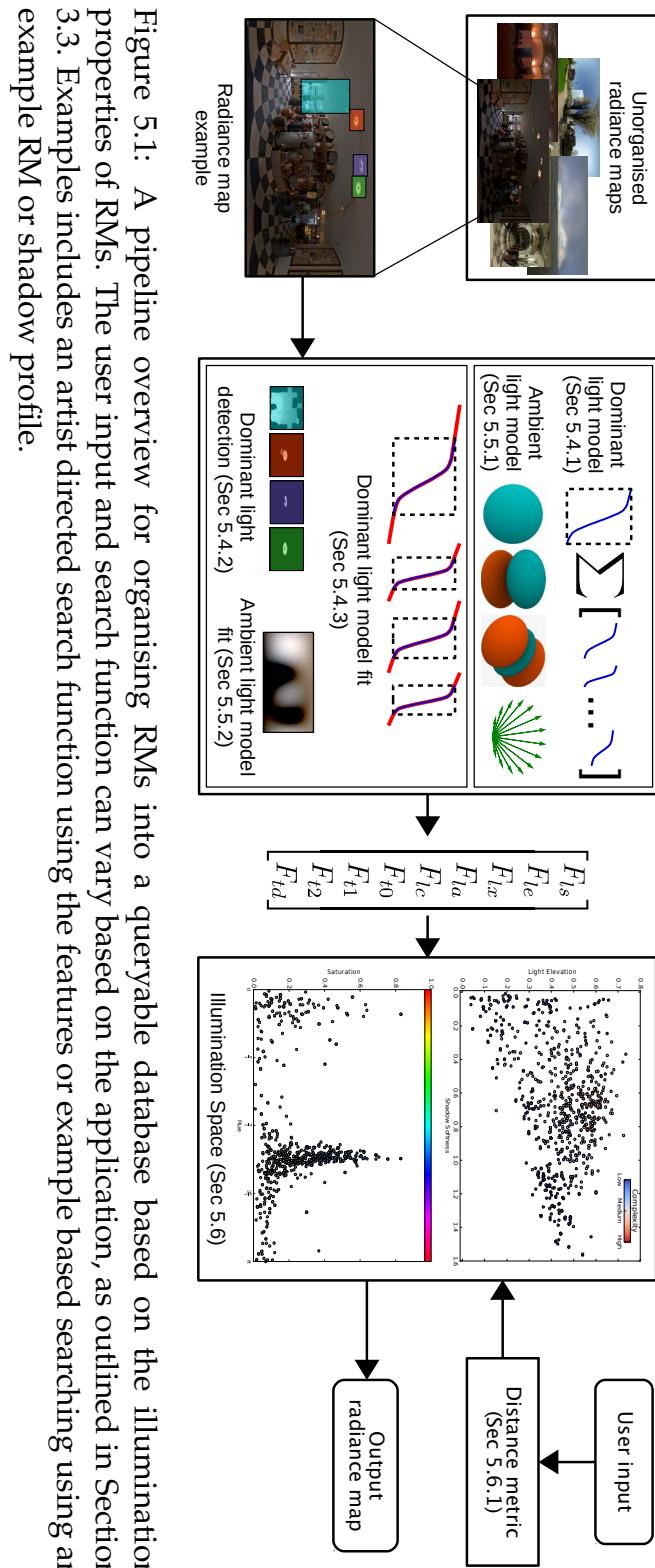


Figure 5.1: A pipeline overview for organising RMs into a queryable database based on the illumination properties of RMs. The user input and search function can vary based on the application, as outlined in Section 3.3. Examples includes an artist directed search function using the features or example based searching using an example RM or shadow profile.

production studios [94] and online services [36, 37, 38, 39, 40, 41, 42, 43], providing a large amount of high fidelity illumination data.

Since a large database of RMs likely contains the desired illumination for most given scenes [25], querying the database for finding the proper RM is important but challenging. It involves clustering, browsing, and searching the database using the illumination properties of the RM rather than their textural features. In AR and MR applications, where the captured real-world scene may not have a corresponding RM, there needs to be a method of retrieving a RM from the database which matches the illumination of the real-world scene.

Previous compact descriptions of RMs, such as the spherical harmonics [95, 96], require a large number of coefficients to express high frequency lighting. Therefore, the spherical harmonics have limitations in providing intuitive and semantic descriptions of a RM with high frequency information. Other methods that describe high frequency lighting [97] require a large number of coefficients. A concise and semantically intuitive set of features is essential for managing large databases.

To address these challenges, this chapter presents a set of novel, semantic and concise '*light features*' to describe a RM's illumination properties. The light features consist of '*dominant light features*' and '*ambient light features*'. Dominant light features describe the dominant light sources in a RM, whereas the ambient light features correspond to the overall colour and tone. To extract the features from the RM, we present a two-step approach. First, a novel light model that mathematically describes light information is developed. Then the light model is automatically fit to the RM using an optimisation process. The light features are then extracted automatically from the fitted light model.

The light model consists of the '*dominant light model*' (DLM) and '*ambient light model*' (ALM). For semantically meaningful features, the models should describe the illumination of the rendered image lit by the RM, rather than the textural property of the RM. An important cue of dominant light sources are the cast shadows. A direct correspondence between the light source and shadow details is obtained by integrating the light source. A shadow edge model (SEM) [98] can be

adapted to the integrated light. However, the original SEM has limitations for fitting real-world light data. Our DLM provides an accurate solution for light model fitting. In order to automate the process, a novel light detection algorithm that automatically extracts regions of dominant lights from a RM is also presented. Furthermore, for a compact expression of ambient colour and tone, a spherical harmonics basis is used as part of the '*ambient light model*'. The ambient light features are obtained from this model. This feature is very similar to the tone feature defined in Chapter 4. For this reason, the majority of this chapter will focus on the dominant light features. However, for clarity, the ambient features will also be explicitly defined in this chapter.

Novel distance measures define the distance between RMs using our light features. These measures separately define the distance of dominant and ambient light, allowing for user flexibility. Using our distance measures and light features, an '*illumination space*' is constructed. This is a low dimensional feature space that efficiently arranges RMs in ways which are intuitive for a user to browse and search.

Based on our survey, this is the first time to present low dimensional features capturing important lighting properties in RMs. Specifically, our light features focus on the semantic illumination properties (as opposed to the texture) of real-world lighting in RMs and the illumination effects on the rendered image. This is also the first time a RM database is semantically organised with a fully automatic pipeline using the light extraction (light detection and model fitting) process. Our illumination space provides a novel approach for browsing and searching the RMs using their illumination properties that has not been provided from the previous methods.

The main contributions are summarised as follows:

- A set of novel '*light features*' are developed to describe the illumination properties of RMs using semantically meaningful parameters. We define dominant light features that express high frequency dominant lights as well as ambient light features for low frequency ambient lights.
- A novel process for automatically extracting the light features is introduced.

We develop the ‘*dominant*’ and ‘*ambient light model*’ that provides a mathematical description of lighting properties. Then the model is fit to the light sources using an automatic optimisation process. This involves a novel light detection algorithm that automatically extracts dominant lights and fits the models to each RM. Laborious manual labelling for a large RM database is avoided.

- Using novel distance measures with the light features, the ‘*illumination space*’ is constructed. This is a low dimensional semantic feature space that arrange RMs by their illumination properties (as opposed to textural properties). This allows for intuitive searching and browsing of RMs.
- Two applications are demonstrated; 1) browsing the RM database using the semantic light features, 2) searching from a given RM to find other similar RMs in the database, and 3) from a given shadow profile, find a RM which produces similar lighting and shadow effects.

5.2 Related Work

Textural Features for RMs: This chapter is focused on features relating to lighting. As such, we refer to the previous chapter of this thesis for an overview on texture based features for RMs. Existing textural features play some role in illumination, but were not designed with this in mind, and as such, have poor results when comparing the rendered image’s illumination. This chapter addresses this by developing a set of features focused on the illumination itself, and describes how the RM affects the rendered scene.

Illumination Features for RMs: Recent work by Karsch et al. [25] has shown interest in using a database of RMs. In this case, they utilise low-dynamic range (LDR) panoramic databases such as SUN360 [45]. Because they are LDR images, they only use directional and positional information of the light sources instead of the LDR intensity values. The method of searching this database is achieved by comparing a mixture of textural and illumination properties of an input image and the LDR panoramic images in the database. If the input image is similar to the

portion of the panoramic image, then the panoramic image itself will be used as the light source for the input image. As part of their optimisation, they also define a loss metric between the LDR RMs with an ℓ_1 -norm between rendered images. This aspect does focus on the illumination, and as such, we compare our method with this approach.

Another method for sorting RMs is the use of Smart IBL (sIBL) [34]. This method organises the data using tags such as sun position, colour, alternate light sources, and other details. The method does not characterise important properties of RMs such as shadow detail and ambient tone. However, recent work has begun to incorporate in view illumination detail such as shadows [55, 99] and the sky [99]. This is improved upon by developing features that describe properties of the RM that affect key components such as the perceivable shadows. Previous work in this domain is also not automatic and requires careful parameter selection for each RM [100], whereas other methods require manually labelled data [25]. Our algorithm is fully automatic, allowing for its use on large databases.

Sampling and Compression for RMs: Storing light information in an image has implications for both memory usage and high sampling rates for high resolution images. Many algorithms sample areas of interest in the RM [101, 102, 103], however, sampling methods aim to oversample the area of interest rather than sample the light source once with a low-dimensional descriptor. Other methods have focused on a frequency space representation, compressing the RM down to as low as nine coefficients [95, 96] for an accurate representation of diffuse surfaces, and hundreds of coefficients for high frequency lighting [97].

5.3 Overview of Light Features

RMs represent real-world lights and are used to illuminate synthetic scenes. Thus, it is desirable to describe a RM based on its illumination properties. Our light features are defined in two parts: *dominant light features* F_l and *ambient light features* F_t . The subscripts l and t are chosen such that F_l is the “*light feature*” and F_t is the “*tone feature*”.

The dominant light features are used to describe high frequency components of the RM, such as the sun, indoor lighting or bright clouds. Whereas the ambient light features describe low frequency components such as the tone of the RM. The key word *dominant* is used to differentiate between groups of pixels which have a high impact on high frequency details (e.g. shadows) as opposed to groups of pixels which don't affect these details very much.

Our light features are concise and have semantic meaning. The conciseness is required to maintain clustering in a low dimensional feature space, and the semantic meaning allows for artists to use the features as parameters to search through the database. F_l is a five dimensional feature vector and F_t is a four dimensional feature vector. The feature vector's semantic meaning in the RM and its corresponding influence in the rendered image is described in Table 5.1.

The light source's size, elevation, azimuth (or x) position, and intensity dictate the shadow's softness, length, direction, and darkness (or amplitude) in the rendered scene, denoted as F_{ls} , F_{le} , F_{lx} and F_{la} respectively (the second subscript represents softness, elevation, x position, and amplitude respectively). The RM's overall ambient colour and intensity is denoted as F_{tl} and F_{td} , where F_{tl} corresponds to the spherical harmonic frequencies, where a combination of low order frequencies is used to describe directional tone. For example, the first three spherical harmonic bands (where the band l is either 0, 1 or 2), then the corresponding feature notation is F_{t0} , F_{t1} and F_{t2} respectively. Finally, F_{td} is a term to describe the diffuse intensity, which describes the intensity of the scene on diffuse surfaces.

A description is still required to account for multiple lights, such as indoor scenes. Due to the infinite number of combinations of lights and shadows (i.e. the number of shadows, their angular disparity as well as the combination of shadow details), it becomes a challenging problem to concisely describe such a quality with a small number of features. We propose to reduce the combination of lights into a single feature based on visual complexity. Previous studies have shown that subjective terms such as complexity are measurable and are important for aesthetic assessment [104]. While the other features (F_{ls} , F_{le} , F_{lx} and F_{la}) are intuitive, the complexity feature may not be straight forward. For this reason, we conduct a user study to guide our decision on deriving this feature. User experiments are conducted to

analyse the light complexity relating to the shadows cast by the RMs. The results of this test are discussed in Section 5.7 and shown in Figure 5.11. Given these results, the feature *complexity* is used to describe RMs with multiple light sources that cause complex shadows (for example, interior lighting), denoted as F_{lc} .

Semantic Table		
Symbol	Light Source	Rendered Image
F_{ls}	Size	Shadow softness
F_{le}	Elevation	Shadow length
F_{lx}	Azimuth	Shadow direction
F_{la}	Light intensity	Shadow darkness
F_{lc}	Light complexity	Shadow Complexity
<hr/>		
F_{t0}	Average colour	Average colour
F_{t1}	Directional colour	Directional colour
F_{t2}	Directional colour	Directional colour
F_{td}	Intensity	Intensity

Table 5.1: Each feature's semantic relationship between the light source and the rendered image. Above the dashed line are the dominant light features, below are the ambient light features.

5.4 Dominant Light Feature Extraction

The dominant light features are extracted in a two-step approach. First, we define a novel light model that describes light and shadow information mathematically. Then, the light model is fit into the RM using an optimisation step. Our novel light detection step makes the whole process automatic.

5.4.1 Dominant Light Model (DLM)

The dominant light features F_l presented in Section 5.3 describe high frequency components in the RM. In the rendered scene, this corresponds to the cast

shadows. A model which describes the light source also inherently describes the shadowing effects of the light source, and vice versa. For example, a light source with a larger area casts soft shadows, and smaller light sources cast sharp shadows. Therefore, if the softness of a shadow is known, it is possible to infer the area of the light. Since we are interested in the illumination effects of a RM, a model based on the shadow information as well as the light is required. The remainder of this section describes the dominant light model (DLM), and the fitting process of the DLM to the RM is described in Section 5.4.3.

In order to define a mathematical model both representing the shadow and light properties, the light source is transformed by integrating across it. Our DLM is derived from the shadow edge model (SEM) by Mohan et al. [98]. We refer to it as a light model (as opposed to a shadow model) since it is describing a light source from the RM. The features F_{ls} , F_{le} , and F_{la} are derived from the parameters of the DLM. Mohan et al.'s SEM is a piecewise quadratic function which represents the umbra and penumbra of a shadow (where umbra and penumbra are the in-shadow and partially in-shadow regions respectively). Mohan et al.'s SEM is defined as follows

$$Q(t) = d + \begin{cases} \frac{-w}{2}, & \text{if } t - t_0 \leq -\sigma \\ \frac{w(t - t_0)^2}{2\sigma^2} + \frac{w(t - t_0)}{\sigma}, & \text{if } -\sigma < t - t_0 \leq 0 \\ \frac{-w(t - t_0)^2}{2\sigma^2} + \frac{w(t - t_0)}{\sigma}, & \text{if } 0 < t - t_0 \leq \sigma \\ \frac{w}{2}, & \text{if } t - t_0 > \sigma \end{cases} \quad (5.1)$$

where d is the vertical offset, σ is the softness of the shadow, w is the shadow's amplitude, t is each point evaluated along the curve, and t_0 is the shadow's center position. The parameters σ , w and t_0 correspond to the light source's size, intensity, and elevation respectively (denoted as F_{ls} , F_{la} , and F_{le}).

The original SEM is specifically designed for the shadow *edge*. The piecewise function has a 0-gradient slope for the first and fourth subfunctions, and the shadow edge is approximated with a parabola in the second and third

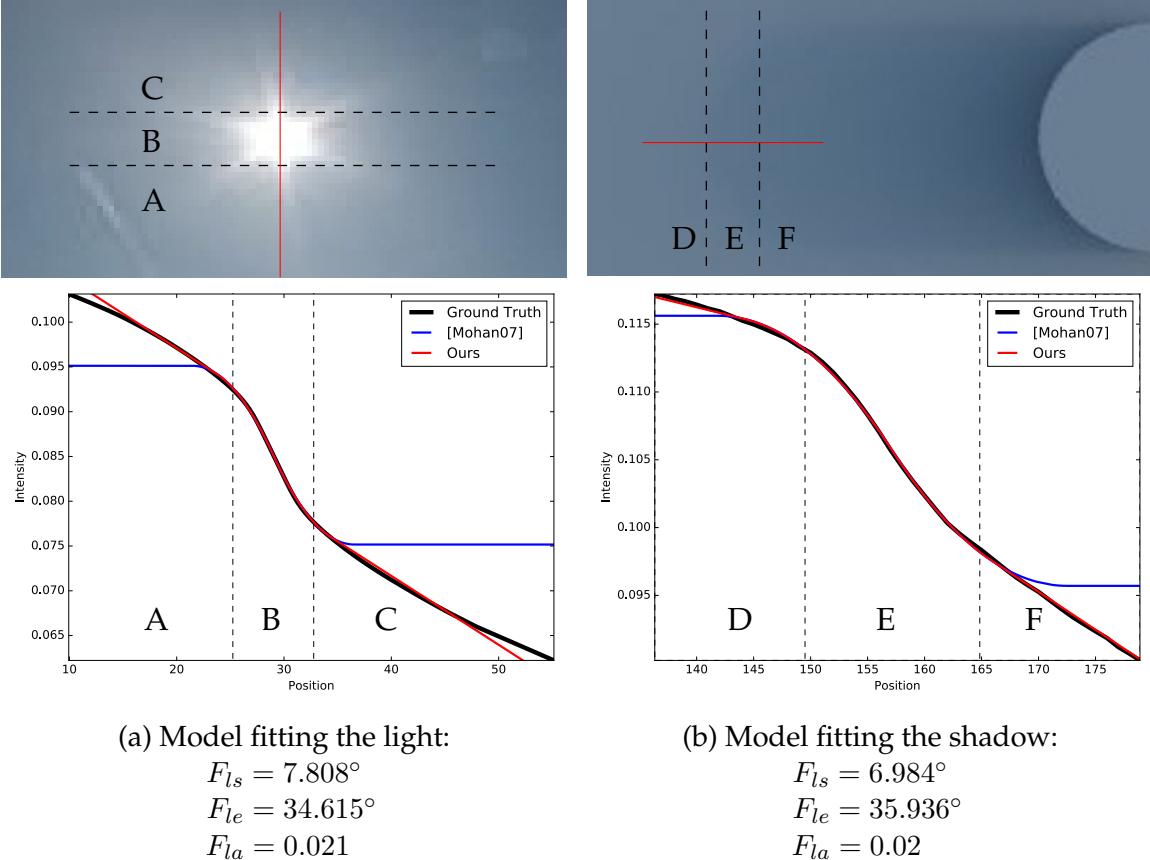


Figure 5.2: (a) is the model fitting the integral of the light source and (b) is the model fitting the raw shadow data. We demonstrate a close correspondence between the two fits. An error of 0.824° , 1.321° and 0.01 for the light size, elevation and amplitude respectively. We also illustrate how our model accurately fits the entire curve compared with Mohan et al.'s shadow edge model.

subfunctions. It was found that while the function works for a well defined shadow edge, it does not fit the overall shadow region (umbra and penumbra) accurately due to the first and fourth subfunctions. This is due to the fact that real-world lighting data does not have a constant umbra region or unoccluded region. A gradual and non-perceivable gradient is very common in such data, as shown in Figure 5.2 (A, C, D and F). Since the SEM does not accommodate for the umbra and unoccluded regions accurately, it has poor fitness for real-world lighting. In order to obtain a better fit, the first and fourth subfunctions of the model are modified to account for slopes found in the unoccluded and umbra

regions. Therefore, our DLM is defined as:

$$Q(t) = d + \begin{cases} (m_a \cdot t) - o_a, & \text{if } t - t_0 \leq -q_a \\ \frac{w(t - t_0)^2}{2\sigma^2} + \frac{w(t - t_0)}{\sigma}, & \text{if } -q_a < t - t_0 \leq 0 \\ \frac{-w(t - t_0)^2}{2\sigma^2} + \frac{w(t - t_0)}{\sigma}, & \text{if } 0 < t - t_0 \leq q_b \\ (m_b \cdot t) + o_b, & \text{if } t - t_0 > q_b \end{cases} \quad (5.2)$$

$$\begin{aligned} q_a &= \sigma - \left(\frac{m_a}{m} \cdot \sigma \right) \\ q_b &= \sigma - \left(\frac{m_b}{m} \cdot \sigma \right) \end{aligned} \quad (5.3)$$

Where m_a and m_b are the gradients on each side of the function, which interpolate between a 0 gradient to the maximum gradient m of the curve dictated by σ . The gradients are offset by o_a and o_b respectively, which are computed by finding the difference between the gradient and evaluating the function at $-q_a$ and q_b respectively. A model which improves the accuracy of the fit is particularly important for fitting entire databases, since manual refinement of the fitting parameters for large amounts of data is not practical. See Figure 5.2 for a comparison between Mohan et al.'s SEM and our light model.

5.4.2 Dominant Light Detection

Our DLM describes each dominant light source separately. Therefore, in order to fit the DLM, we need to extract the dominant light sources from the RM. We develop a novel light detection algorithm which detects the dominant light sources in HDR RMs. Since the features F_l describe high frequency components which corresponds with shadow information in the rendered image, the light detection algorithm aims to reconstruct the rendered image such that the shadow details remain intact. Furthermore, the feature space grows with the number of detected lights, therefore, the light detection algorithm aims to find the *smallest* number of pixels in the RM required to maintain the shadow detail. The features F_{lx} and F_{lc} are also obtained

after the dominant light detection algorithm, which correspond to the detected light's azimuth angle and the overall light complexity dictated by the number of detected dominant lights.

Previous work in this area has focused on LDR RMs [25, 45]. These methods had to work around the problem of differentiating between light sources and reflections within images, whereas other methods detect both [105]. This is due to the fact that the LDR image format does not contain reliable luminance values, and therefore the pixel intensity values for light sources and reflections are identical. In order to differentiate between the two, solutions are based on the textural properties of the image and supervised learning.

Because we are working with HDR images, we do not have this problem. Instead, we face the problem of determining whether a region of pixels contains enough luminous intensity relative to the surrounding pixels to be considered as a dominant light source. Our solution to this problem is to use the corresponding rendered image of a RM as a guide. The shadow information in the rendered image is used as an indication of whether or not pixels in the RM represent a dominant light source which caused the shadows. We propose an iterative algorithm for determining whether or not pixels in a RM are dominant light sources.

Algorithm: For a given RM, we render a ground truth image

$$I(r) = \int_{\Omega} L_i(\omega_i) f(\omega_i) |\cos \theta| \, d\omega_i \quad (5.4)$$

where I is the rendered image given a RM r which computes the rendering equation. L_i is a light sampled from RM r . We use a top-down orthographic camera of a cylinder on a ground plane, as shown in Figure 5.3. The cylinder casts shadows in each direction, and the ground plane captures the shadow information. We use a cylinder because it is symmetrical and casts shadows evenly in each direction. The material property f is a diffuse Lambertian material for both the cylinder and ground plane.

From the ground truth rendering, we then extract the shadow information. To do this, we adopt a differential rendering approach [22] and compute the difference

between the rendered image of a cylinder on the plane with a rendered image of only the plane. This in effect is subtracting the diffuse value of the scene from the image, producing shadow extracted image in Figure 5.4

$$f(I) = I - I_d \quad (5.5)$$

where I is the rendered image and I_d is the rendering of the Lambertian reflectance model with respect to the normal of the ground plane. The colour of the material is set to white. The result of this operation is a difference image which contains all the shadow detail (see Figure 5.4).

However, there are a lot of shadow details stored in $f(I)$ that we cannot perceive. Our goal is to minimise the complexity of a RM by detecting the smallest number of light sources required to reconstruct the ground truth rendered image. Therefore, we modify the difference image by removing shadow information which we cannot perceive. This has an effect on the optimisation such that we are not fitting the ground truth of all the shadow detail, but instead, we are fitting shadows that we can only perceive. This in effect will discriminate against non-dominant light sources which do not cast shadows in the perceived shadow region.

We use the just noticeable difference (JND) and Weber's Law [106] to determine which shadows are perceivable to the human eye

$$\frac{\Delta I_d}{I_d} = k \quad (5.6)$$

where I_d is the diffuse unoccluded intensity value, and ΔI_d is the change in intensity which is perceivable. It is shown that the ratio between the diffuse light and the perceived additional light is constant k . Therefore, we use this law to act as a threshold value to eliminate shadow regions which are not perceivable to the human visual system. For the human visual system, k is approximately in the range of $\frac{1}{12}$ to $\frac{1}{100}$ [107]. Experimentally we found that $\frac{1}{30}$ worked well. In practice, this is a threshold value on the difference image. Therefore, the JND threshold is k scaled by the diffuse intensity

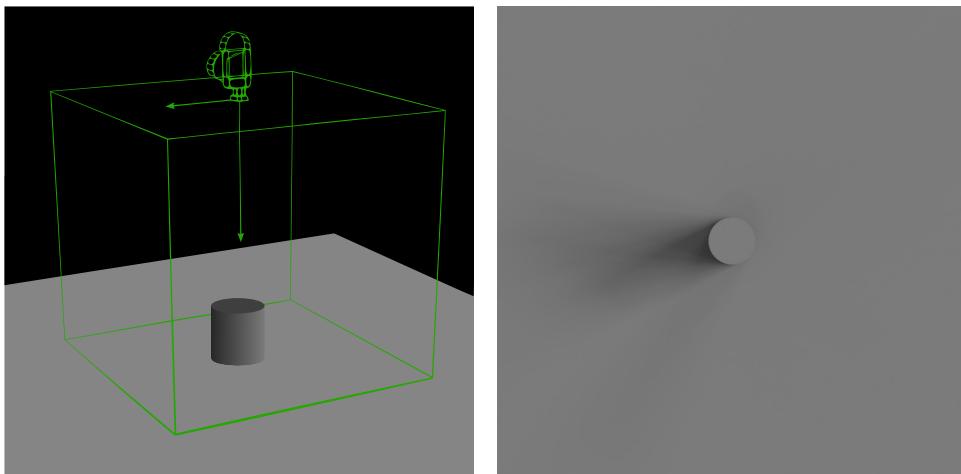


Figure 5.3: On the left is a top-down orthographic camera of a cylinder on a plane. This setup is used to obtain a clear view of the shadow information cast in all directions on the horizontal plane, and on the right is an example rendering using an IBL as a light source.

$$\text{JND} = k \cdot I_d \quad (5.7)$$

which we apply to $f(I)$

$$g(f(I)) = \begin{cases} 0, & f(I) \leq \text{JND} \\ f(I), & \text{otherwise} \end{cases} \quad (5.8)$$

For brighter or dimmer surfaces, the JND threshold scales appropriately. See Figure 5.4 for an example of the rendered images and the visible difference of the shadow region.

Once the ground truth data is established, we then compute superpixels using SLIC [108] on the RM we aim to detect the lights from. We consider each superpixel as a potential dominant light source. From this set, we need to determine if the superpixel is a shadow-casting or shadow-altering light source. We use a variation of an l_0 optimisation to obtain the optimal set. We find the smallest set of superpixels such that we reconstruct the JND filtered difference image

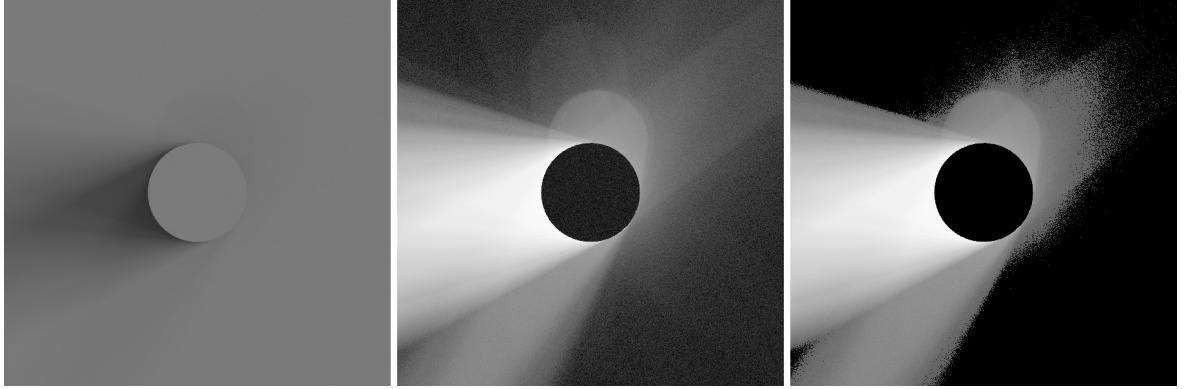


Figure 5.4: From the left: rendered image, difference image showing shadow detail of the cylinder on the ground plane, and thresholding to remove shadows not perceptible to the human eye using Weber’s Law. These show a zoomed in version of the data that was used.

$$\begin{aligned} \min_x \quad & \|x\|_0 \\ \text{subject to} \quad & f(I(Ax)) = g(f(I(b))) \end{aligned} \tag{5.9}$$

where A is a matrix where each column is a RM for each superpixel, b is the rendered image by the unaltered input RM, and x is a binary activation function of the superpixels in A . l_0 -optimisation is NP-hard as there is a large number of combinations in A to approximate b . However, to improve the iterative algorithm, we sort the superpixels in descending order, from the highest intensity superpixel to the least. From this, we use a brute-force matching pursuit algorithm and enable each superpixel sequentially. We compute the difference image of the rendering at each step of the iteration, and compare it with the ground truth shadow difference image. The error is computed using STSIM-2 [109, 110], which is a metric that corresponds with the human visual system and has shown good results compared with other perceptual metrics [111]. If the similarity score improves on the previous iteration, the current superpixel is kept, if it does not improve, it is rejected. Because we are discriminating against shadow information that we cannot perceive using Weber’s Law, we are in effect discriminating against light sources that casts shadows we cannot perceive with the human visual system. This has the desirable effect of omitting light sources in the RM that do not emit enough

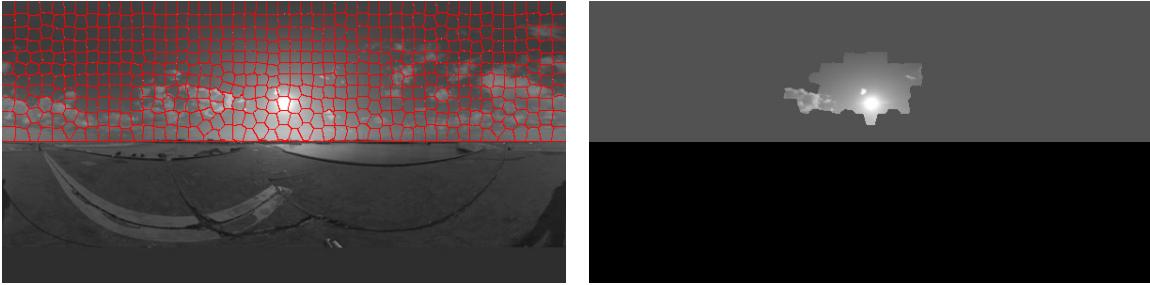


Figure 5.5: Our light detection algorithm computes superpixels on the RM (left) and finds the set of superpixels which contribute as a dominant light source (right).

radiant intensity to cast perceivable shadows. There is a free parameter which can be added to the error to discriminate against superpixels which do not increase the error, but also do not improve the error by very much. In practice, we use a small value 0.0001.

Once we iterate through each superpixel, a set of superpixels are found as the dominant light sources (see Figure 5.5 for an example). These are the bare minimum superpixels required to produce a rendered image with shadow information which looks very similar to rendered image by the original RM. Every superpixel adjacent to another is combined into a single patch. Each patch area is considered a dominant light source. It is possible for multiple dominant light sources to share the same patch region. To avoid such cases, local maximas are computed on the light detected RM. Instead of assigning superpixels to adjacent superpixels, the superpixels are assigned to the nearest local maxima. To improve the local maxima detection, the noise is reduced using a Gaussian blur filter. Each feature for each light source having the azimuth position F_{lx} is set to the x position of the local maximas. See Section 5.7 for the results of our dominant light detection algorithm.

5.4.3 DLM Fitting

The DLM describes the lighting properties of a dominant light source. A RM can be comprised of pixel regions which act as a dominant light source (e.g. the sun, light bulb, window, etc.). Each of these dominant light sources can be described by a

DLM separately. In order to fit the model to the dominant light source in a RM, the pixel region must be extracted from the RM either manually or automatically (an automatic method is detailed in Section 5.4.2). Since a RM may contain multiple dominant light sources (e.g. multiple light bulbs), each dominant light source will be fitted separately. For example, a RM with three light bulbs will have a DLM applied to each one. Four of the five features from the dominant light feature vector are extracted from the DLM directly (F_{ls} , F_{le} , F_{lx} , F_{la}). However, this raises a complication if there are multiple light sources within a RM.

Initialisation

In order to obtain an accurate fit of the model to the data, the optimisation algorithm is initialised with a guess. Specifically, the parameters σ , w , t_0 , d , m_a and m_b from equation (5.2) need to be initialised with boundary conditions.

For a given dominant light source, a 1D curve C is produced by integrating across the light source vertically (Figure 5.6a.). This transforms the light data into shadow data, which is the data that the DLM fits. During the integration, the elevation which had the highest intensity is stored as l_e . This elevation is likely to be similar to the optimal t_0 (from equation (5.2)), therefore t_0 is initialised to l_e , and the vertical offset d is set to the integrated value at the position l_e . The DLM is designed to correspond to the penumbra region with a linear falloff on the umbra and unoccluded regions. However, the curve C may contain a large umbra or unoccluded regions with a non-linear gradient. To focus on the point of interest, C is resampled at a higher rate within the penumbra. This is achieved by first oversampling C with cubic interpolation, producing a new curve C_i (as shown in Figure 5.6a.). The curve C_i is then resampled with a high concentration rate around l_e using Gaussian random sampling. To determine the density of the concentration, the Gaussian kernel has three parameters to estimate. The width, amplitude and position. The position is constrained by l_e and the amplitude is constrained by sampling C_i at the point l_e . Because the model approximately follows the shape of a cumulative distribution function (CDF) of a Gaussian, it is approximated by taking the derivative of C_i and fitting a Gaussian around it with constraints on the

position and amplitude (see Figure 5.6b). C_i is then resampled using the Gaussian kernel giving C_s (Figure 5.6c.).

With the resampled curve C_s , the boundary conditions of the optimisation are estimated. First, σ is the width of the shadow edge and cannot exceed the total height of the set of superpixels in the light source. Because the superpixels have been resampled with a higher concentration rate, this tightens the boundary further. The boundary for σ is set halfway between 0 and the difference between the maximum and minimum value of the sampled points from C_s , and the guess is initialised to the centre of the boundary. The amplitude w is bound between 0 and the difference between the unoccluded diffuse intensity and the total integrated intensity of the light source patch, which corresponds to the minimum umbra intensity. This parameter is also initialised to the centre of the boundary. The gradients m_a and m_b can be estimated by the curve fitting process. However, the parameters can be accurately estimated by sampling the first and last 10% of the shadow curve respectively, estimating the line, and set the gradients as constants in the fitting process.

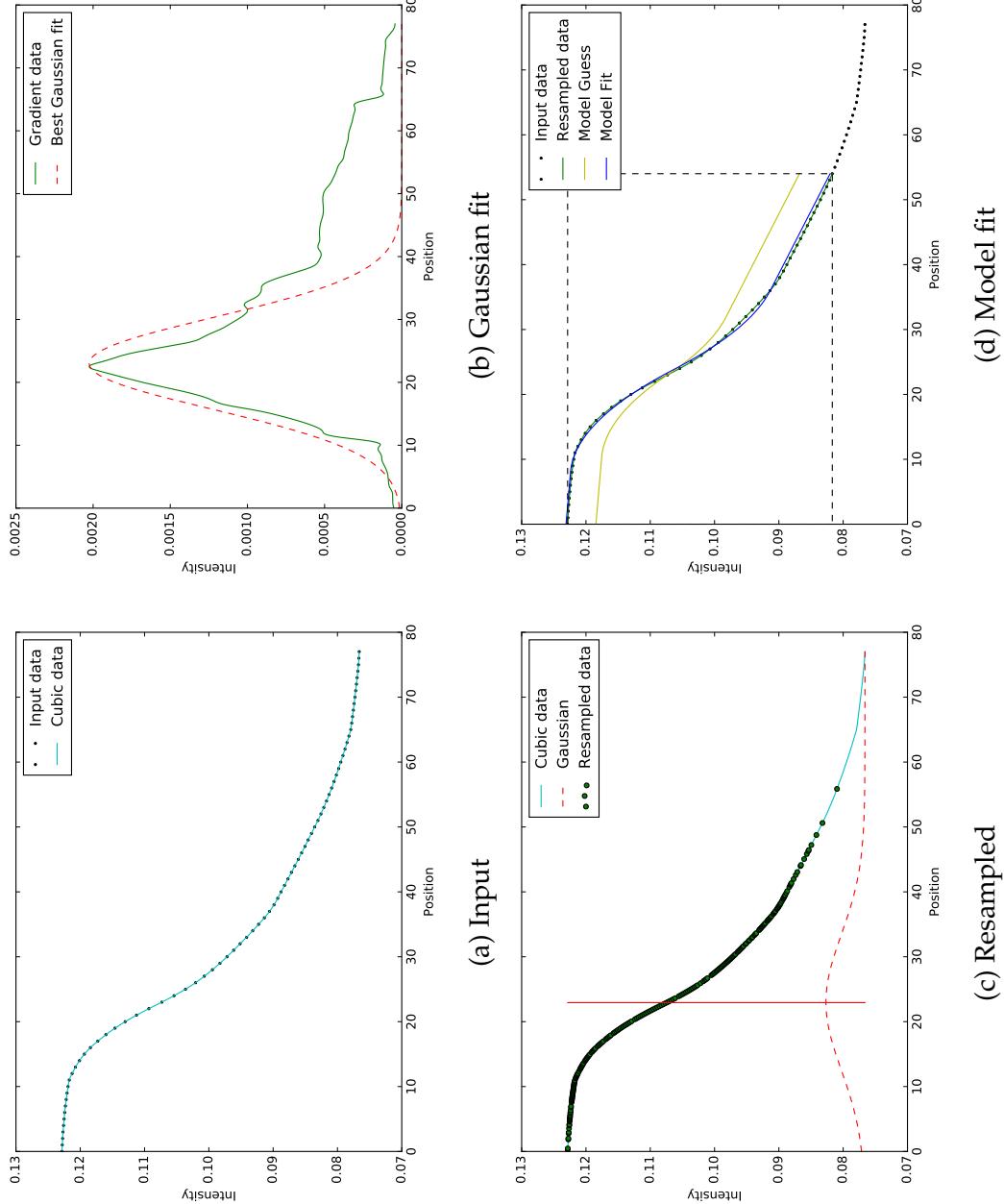


Figure 5.6: The fitting pre-process. (a) The input data is oversampled using cubic interpolation. (b) The derivative of the cubic curve (or a 1D representation of the original light source) is computed and a Gaussian distribution is approximated around the peak. (c) Using the Gaussian, the oversampled curve is resampled, producing a higher density around the penumbra. (d) The DLM approximates the original curve using the oversampled and high density points.

Optimisation

Given the initialisation parameters and boundary conditions, a nonlinear dogleg optimisation algorithm is used to approximate the curve [112]. The bounding conditions, initial guesses and the optimal fit are visualised in Figure 5.6d. The optimal fit is also visualised on the RM in Figure 5.7. Of particular interest is the shadow softness parameter σ , the amplitude parameter w which indicates how light or dark the shadow is, and t_0 which is the optimised light elevation. The features are set using the optimised parameters as $F_{ls} = \sigma$ and $F_{le} = t_0$. A possibility is to set $F_{la} = w$, however, this amplitude value specifically measures the shadow edge and not the umbra region. For better results, a ratio is taken between the integrated intensity of the light source and the integrated intensity of the entire hemisphere. This effectively describes the overall impact the light source region has relative to the entire RM. Running the fitting process again with a similar setup on the horizontal axis gives the optimised parameter $F_{lx} = t_0$. Running it a second time also gives another σ . This value is averaged with the vertical σ , improving the accuracy of F_{ls} . The improved benefits of the model is that it does not require manual refinement of the light source. The integration and processing of the light source is all done automatically.

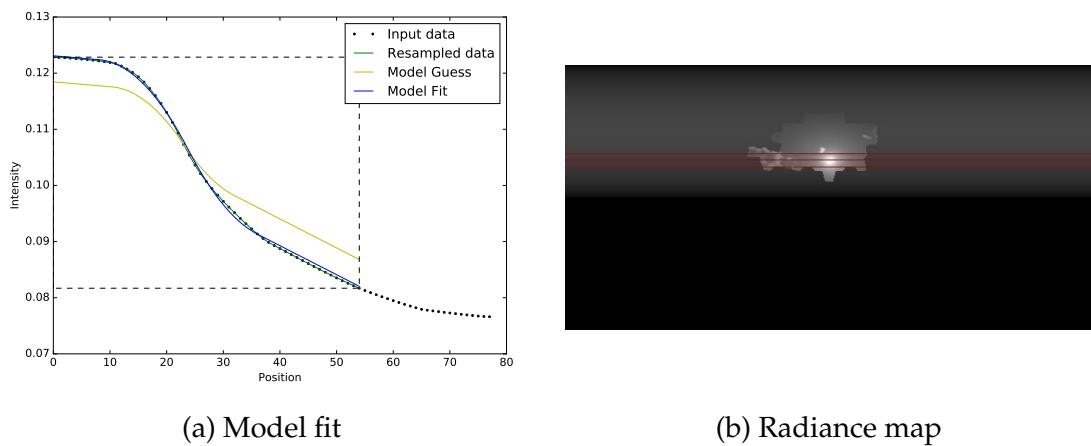


Figure 5.7: (a) Fitting the light model to the (b) dominant light source. The red horizontal lines in (b) visualise the parameters of the model in (a), showing a good fit around the dominant light source (in this case, the sun).

Since the DLM can be fit to multiple dominant light sources for a given RM, the four features (F_{ls} , F_{le} , F_{lx} , F_{la}) of each DLM need to be summarised into a single set of four to store in the final dominant light feature. To aid in our decision on how to summarise the DLM features, we refer to the complexity user study 5.7. It was shown that shadows which were darker were most influential in the users decision. As such, the four features per light source are condensed into a single set of four by averaging each DLM together, weighted by how dark the shadows are cast by each dominant light. Each DLM has an amplitude associated with it. This feature measures how dark the shadow is. Therefore, the amplitude feature is used as a weight for its corresponding DLM. The amplitude features are normalised with respect to all the light sources in the RM

$$F_{ls} = \sum_{i=0}^N F_{ls}^i \cdot F_{la}^i \quad (5.10)$$

where N are the number of dominant light sources in the RM. A similar process is applied to F_{le} and F_{lx} . The final dominant light feature is complexity (F_{lc}). As found in the user study, the number of lights was the strongest factor in determining the complexity of shadows. However, since users were drawn to dark shadows, the amplitudes are also used to weight the complexity feature. Shannon entropy has shown to be a good predictor of complexity [113], and is used on the array of light amplitudes to compute the overall complexity

$$F_{lc} = - \sum_{i=0}^N F_{la}(\ln F_{la}) \quad (5.11)$$

where the light amplitude vector is normalised to a probability distribution. Intuitively, single dominant light source has a low entropy value, whereas a set of lights with even distribution has a high entropy value. The result is that RMs with a large set of dominant light sources is classified as a high complexity light source.

5.5 Ambient Light Feature Extraction

The ambient light feature provides the low frequency information of the RM. This plays an important role in determining the mood and tone of a rendered scene. Four features are used to describe the ambient colour distribution and overall intensity of a RM.

5.5.1 Ambient Light Model (ALM)

The spherical harmonics basis has shown to be an accurate descriptor of low frequency lighting [92] and has been used to describe the visual tone of sky maps in Chapter 4. A similar method is applied for describing RMs, where the RGB values are computed on the RM to parameterise the ambient lighting. The previous work was concerned with the visual tone of the sky map, however, the same features can be used for the tone of the rendered image. The ambient tone features F_{tl} is defined as the summed value across each frequency band

$$F_{tl} = \sum_{m=-l}^l \text{coef}_l^m \quad (5.12)$$

Using a vector of F_{tl} up to some specified band can give more directional ambience, capturing varying levels of frequency. As shown in [92], $l \leq 3$ is enough to capture diffuse lighting. Selecting a coefficient at higher frequencies with relatively higher values can also be used to describe the tone toward stronger light sources.

5.5.2 ALM Fitting

Similar to the previous chapter, l is defined by the user (equation 5.12). This reduces the number of coefficients from l^2 for each colour channel to m for each colour channel. Once l is specified, F_{tl} is then used to fit RMs in the database. Similar to Chapter 4, the features can be stored as ratios

$$F_{tl} = \ln F_{tl_{red}} - \ln F_{tl_{blue}} \quad (5.13)$$

Because the ratio of red and blue is invariant to intensity, the diffuse intensity of the RM with respect to the zenith up angle is also stored as F_{td} .

5.6 Illumination Space

We define an illumination space that describes the illumination properties of RMs using the nine light features. Each light feature has a semantic meaning which has correspondence in both the RM and the rendered image, see Table 5.1 for a clarification on what each feature means to the user.

5.6.1 Distance measure

Given RM^1 and RM^2 as well as their corresponding features F^1 and F^2 , their dominant light distance D_l is computed as follows

$$D_l = \frac{1}{\pi} \arccos(F_{lp}^1 \cdot F_{lp}^2) + \sqrt{(F_{ls}^1 - F_{ls}^2)^2 + (F_{lc}^1 - F_{lc}^2)^2} \quad (5.14)$$

where F_{lp} are the features F_{le} and F_{lx} concatenated into a vector. Each component is normalised except the complexity feature which is clamped between 0 and 1 since its value is unbounded. Alternatively, metric entropy can be used to normalise the complexity feature.

For the colour features F_{tl} , the CIEDE2000 colour distance formula can be used [114] and summed across each band l . Similarly, the distance function is applied to the diffuse intensity F_{td} . In Figure 5.9, an ℓ_2 -norm with F_{tl} is used, and the ambient light distance D_t is defined as follows.

$$D_t = \sqrt{\sum_{l=0}^2 (F_{tl}^1 - F_{tl}^2)^2} \quad (5.15)$$

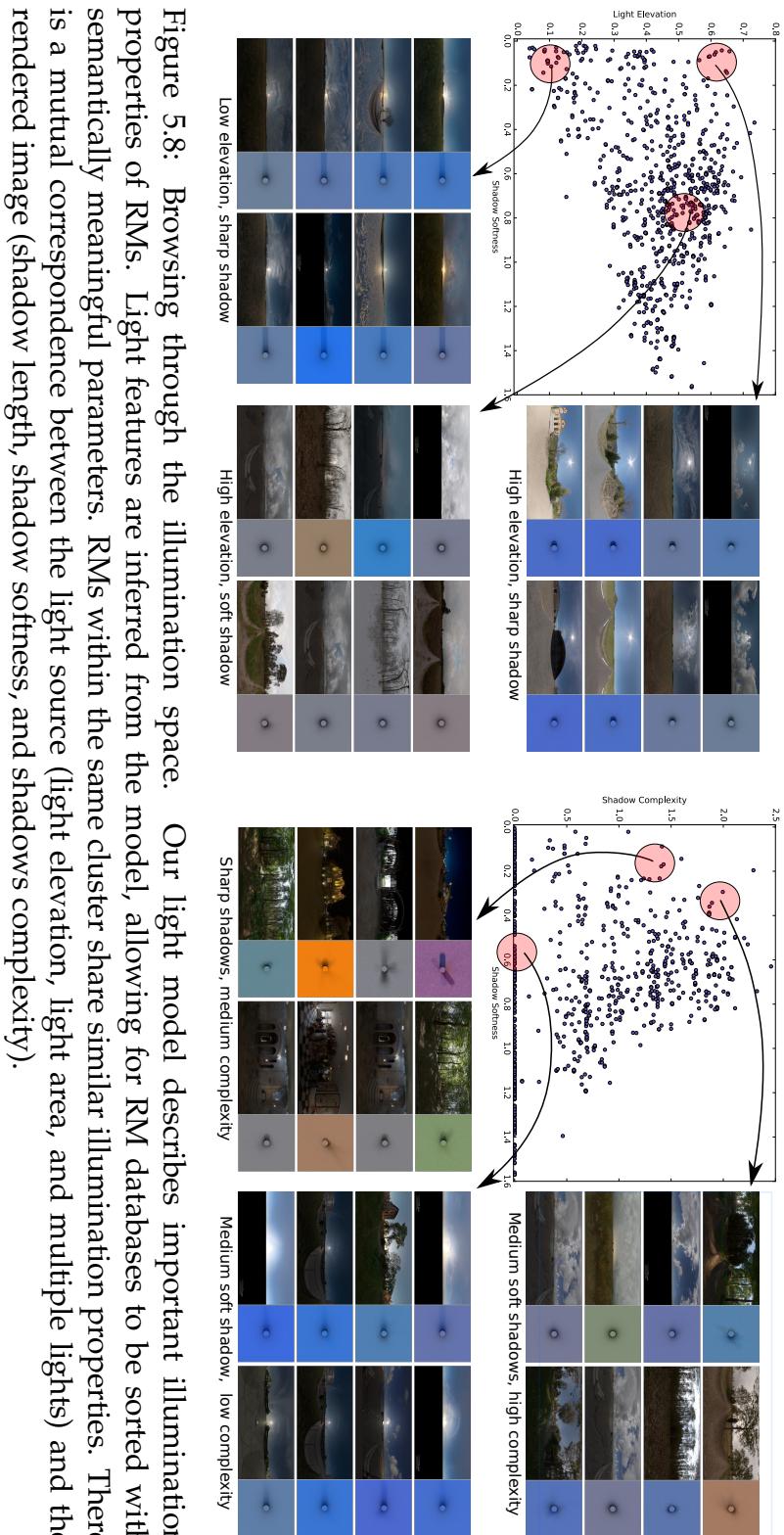


Figure 5.8: Browsing through the illumination space. Our light model describes important illumination properties of RMs. Light features are inferred from the model, allowing for RM databases to be sorted with semantically meaningful parameters. RMs within the same cluster share similar illumination properties. There is a mutual correspondence between the light source (light elevation, light area, and multiple lights) and the rendered image (shadow length, shadow softness, and shadows complexity).

A weighted combination between D_l and D_t can be used to search the database.

$$D = (D_l \cdot w_l) + (D_t \cdot w_t) \quad (5.16)$$

Where the weights w_l and w_t sum to 1.0.

5.6.2 Browsing

Using a combination of these features allows for an intuitive browsing experience, visualising the space on certain axes. Figure 5.8 shows an example of a 2D scatter plot of the RM database where users can freely search for RMs with varying elevation F_{le} and shadow softness F_{ls} . The colour coding aids users in distinguishing skies with single dominant lights or many lights (F_{lc}). Figure 5.9 is another example for visualising the database with varying tone F_{tl} and a fixed elevation F_{le} .

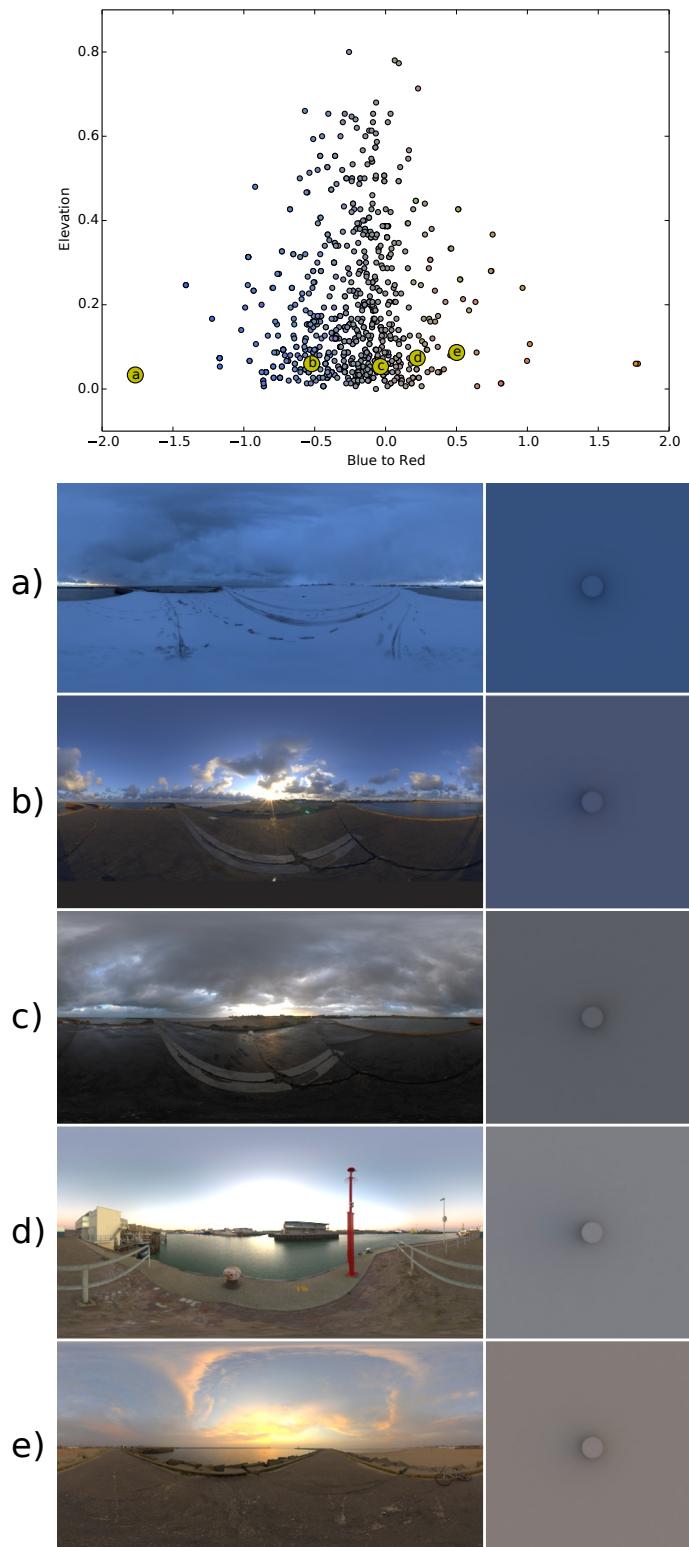


Figure 5.9: Visualisation of the feature space with elevation and tone. Points a, b, c, d and e are sampled at a low elevation, varying from blue to red. The RMs and their corresponding rendered images are shown ([a-e, top-down]).

5.6.3 Searching

Each parameter or a combination of parameters can be used to search using k -nearest neighbour (k -NN). This process can be done in real-time using optimised variants of k -NN algorithms [93]. Aside from k -NN, alternate methods using weighted measures are also feasible. For example, the separation of measures for dominant lights and ambient lights allows for a weighted combination of the two components. From our experiment, we found that users would tend to find RMs based on the prominence of the most dominant shadow. This corresponds to the dominant light with the highest amplitude F_{la} . Therefore, a search space using a single dominant light is used, where additional lights are generalised with F_{lc} such that the user can specify the complexity of the RM they're looking for.

5.7 Results

There are a number of stages to obtain the light features. While the light position (F_{le}, F_{lx}), light intensity (F_{la}) and the light size (F_{ls}) correspond to typical intuitive parameters for abstract light sources (e.g. point light, area light, etc.), the accuracy of obtaining these features should be evaluated. We show this accuracy by comparing with the previous SEM model by Mohan. The ambient features (F_{lt}, F_{ld}) are evaluated through a user study in which professional artists evaluate how well the ambient features search the RM database. Finally, the complexity feature (F_{lc}) is the most unusual parameter as it summarised the overall complexity of a RM into a single feature. To evaluate this feature, two user studies are conducted to show that users showed a clear pattern for determining that RMs with more distinguished shadows are more complex than RMs with single or soft shadows. The light detection algorithm itself is also evaluated by comparing with Karsch et al.'s [25] state of the art method. Finally, we show the results of our features in various applications including browsing, searching and a controlled inverse rendering experiment by searching the RM database using a shadow profile.

5.7.1 Dominant Light Model

Our light model is based on the lights in the RM which cast perceivable shadows in the rendered image, whereas the previous SEM is based on the shadow edge cast by the light source. To evaluate the accuracy of the models, the mean squared error (MSE) is computed on the residuals of the two models with 500 RMs. Since there can be multiple lights per RM, there are in total 2849 light sources to fit. Each light source is integrated vertically to simulate a shadow, and the models aimed to fit the integrated data. Our DLM reduced the residual error from the SEM by 53.8%. There were 3 outliers removed from our model, whereas 16 were removed from the SEM. The outliers are removed from the data analysis to minimise skewing of the results, however, noting the number of outliers is important to show how often each model failed to fit the data. A raw comparison of the 2849 lights is shown in Figure 5.10 as a ratio. This figure illustrates that our model obtained a better fitness in 94.31% of the cases.

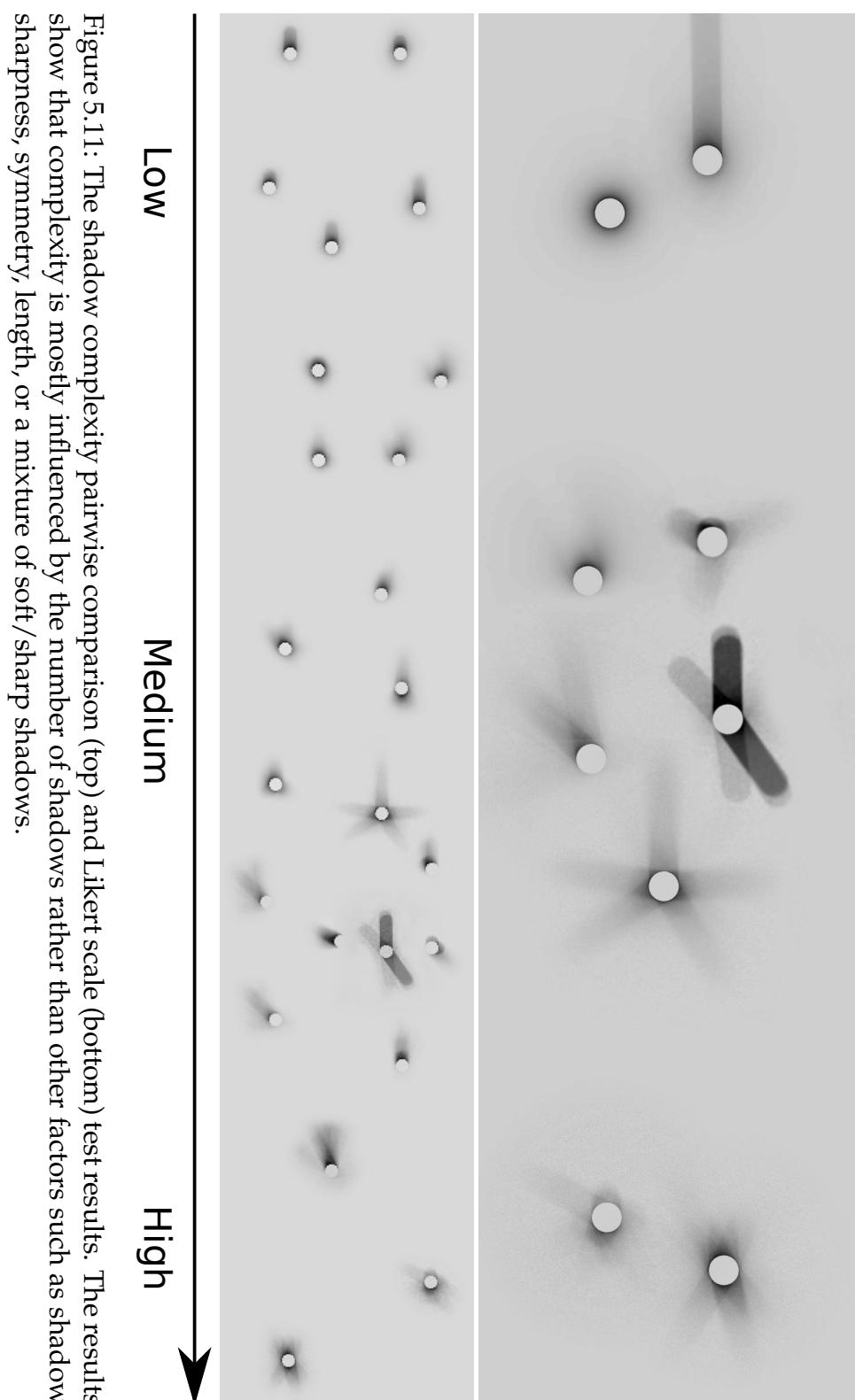


Figure 5.10: Fitness error comparing our DLM (green) with [Mohan07]'s SEM (red). The errors are normalised and compared horizontally (smaller is better). There are 2849 comparisons made (incrementing on the y axis), and the results are sorted in descending order (left to right) for readability.

5.7.2 Complexity User Study

The aim of the complexity user study is to find any particular patterns which are important to parametrise as a feature. From the experiment, we found users determined the complexity score mostly influenced by the number of shadows. Additional factors such as symmetry and shadow softness played a minor role. A test set of images was generated by rendering a cylinder on a plane, identical to the setup in the dominant light detection section (Figure 5.3). The rendered images' intensity values were rescaled for consistency.

Similar to previous studies in visual complexity [115, 104, 116], two studies to measure complexity were conducted. First is a pairwise comparison test and the second is a Likert scale test. There were a total of 40 participants tested. The first test had participants compare nine pairs of images with one another, choosing which image they thought to be more complex than the other. To uncover any underlying patterns in the users' responses, the results of the survey were compressed using multidimensional scaling [117]. While there are many potential variables to consider for complex lighting (shadow length, angular disparity, and variation of softness), the number of shadows was shown to be a strong factor for deciding the complexity of the shadows. The results of this test are shown in Figure 5.11. The pairwise comparison was followed up with a Likert scale test with a new set of participants. The participants were asked to score 50 images ranging from -3 (*Extremely simple*) to 3 (*Extremely complex*). The first 10 images were removed from the analysis to allow participants to understand the range of shadows. The complete results of the Likert scale test are shown in Figure 5.12 and a visualisation is shown in Figure 5.11. The Likert scale test corroborated with the first test (see Figure 5.11), showing that participants were mostly affected by the number of shadows. Furthermore, for images with a similar number of shadows, the results show that softer shadows tend to be deemed more complex.



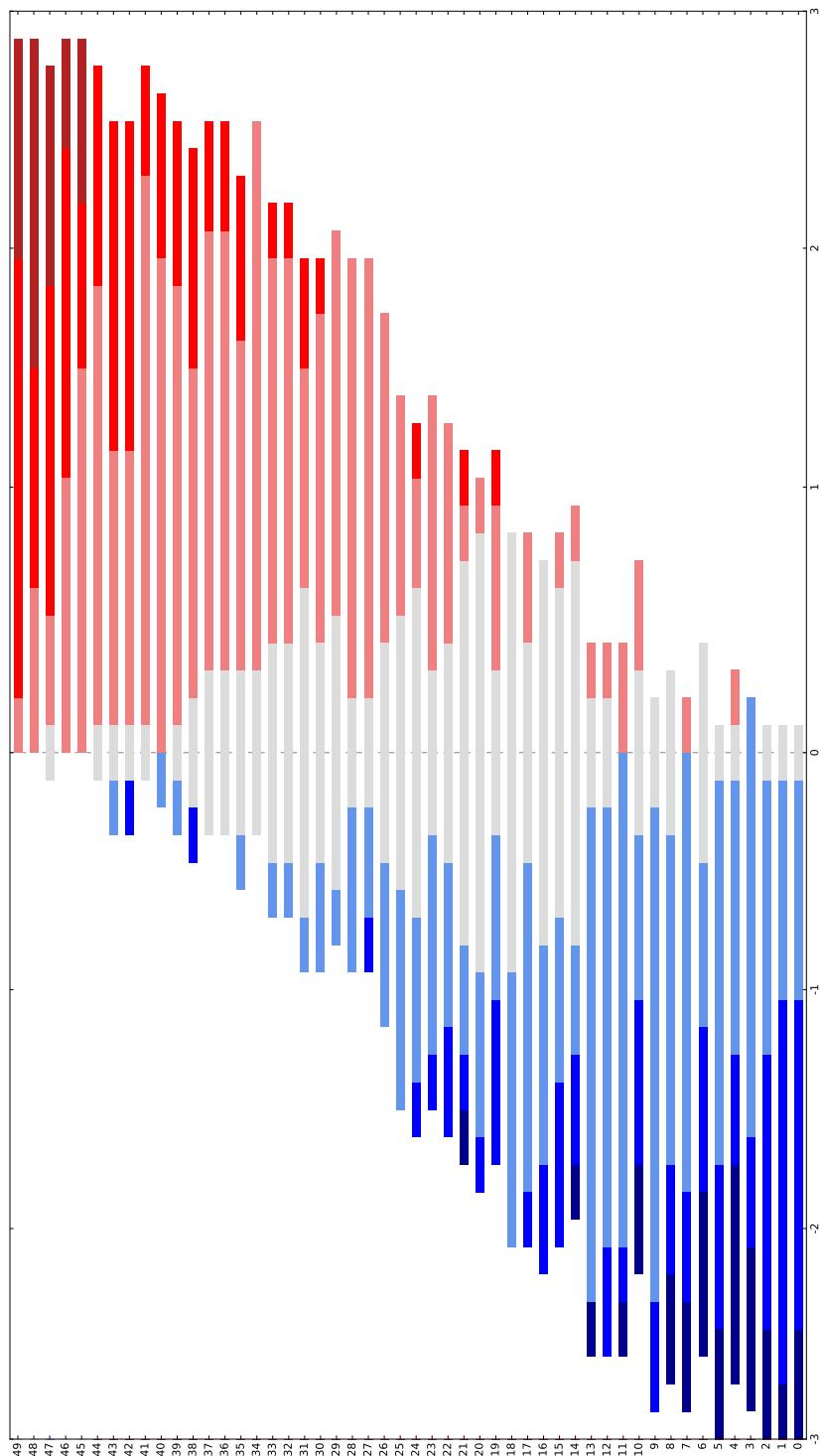


Figure 5.12: The shadow complexity Likert scale test results, from simple (bottom) to complex (top). The results are sorted in ascending order (left to right) for readability.

5.7.3 Dominant Light Detection

We compared our dominant light detection with Karsch et al.'s [25] light detection, as well as naive thresholding. Thresholding is computed by setting a threshold such that the luminous intensity matches the amount of intensity produced by Karsch et al.'s mask. Figure 5.13 illustrates how our method detects only the necessary pixels as the dominant lights to obtain all the perceivable shadow detail. A user study is conducted where 50 participants were asked to compare the ground truth rendered image with the dominant light detected RMs. While Karsch et al.'s method was designed for LDR RMs, for a fair comparison, their mask was used with a HDR RM. The non-dominant light pixels outside of the light mask are filled with a constant ambient term to increase the intensity of the rendered image to match the ground truth. The participants scored on a Likert scale the similarity of the rendered image to the ground truth image. They were then asked to rate how similar the *shadow detail* is between the image rendered by the light detected RM to the image rendered with the unaltered RM. The Likert scale is a 7 point scale of similarity: -3, 'Completely different', -2, 'Very dissimilar', -1, 'Dissimilar', 0, 'Neutral', 1, 'Similar', 2, 'Very similar', and 3, 'Looks the same'. The user study shows that our result maintains the same visual quality as the ground truth yet reduces the number of dominant light pixels more than thresholding and Karsch et al.'s method. Unlike the other methods, our algorithm is fully automatic and does not require labelled data. See Figure 5.14 for the results of the user study.

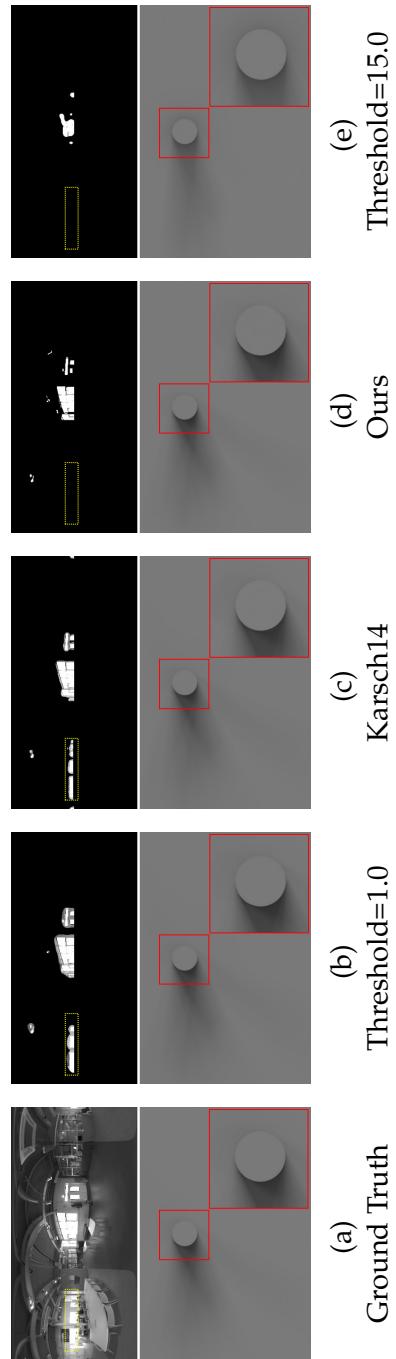


Figure 5.13: Comparison of dominant light detection methods (top row) and their affects on the rendered image (bottom row). (a) is the ground truth, followed by various light detection algorithms. (b) is naive thresholding with a threshold value of $T=1.0$, (c) is Karsch’s method, and (d) is our method. Notice that the rendered images look similar to the ground truth, yet our method has removed unnecessary light sources in the yellow dotted region to extract the dominant lights. (e) is thresholding by increasing the threshold to remove the unnecessary light sources, yet this introduces artefacts in the rendered scene due to removing other important light sources.

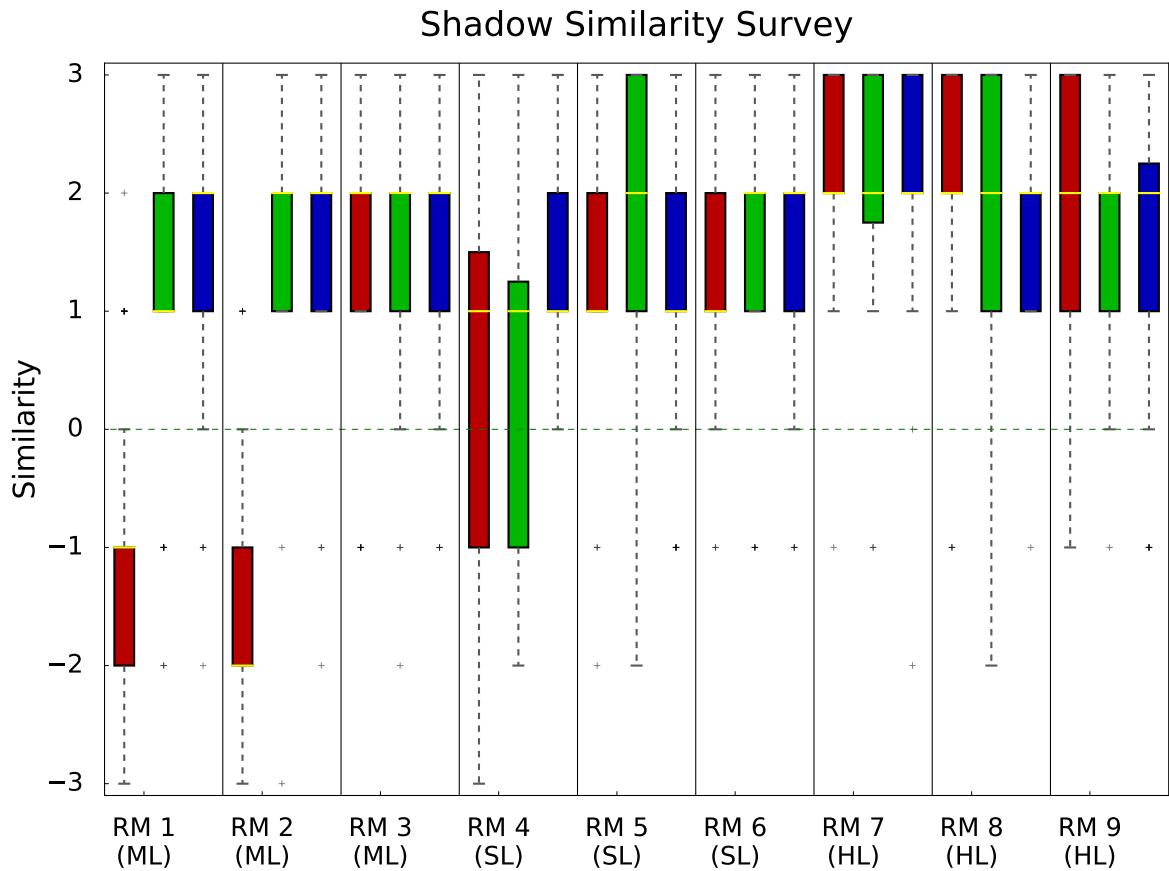


Figure 5.14: User study results comparing [Karsch14]'s light detection (red bars), thresholding (green bars) and our method (blue bars). The reduced amount of dominant light pixels averaged across for each RM is 79.3%, 80.1% and 95.6% for each light detection method respectively. ML are the RMs with multiple lights, SL has one soft light, and HL has one hard light.

5.7.4 Browse using features

Since the features have semantic meaning and correspondence between the RM and the rendered image, users can adjust each parameter freely to browse the database. It was shown that gradually changing individual parameters shows a smooth transition between RMs and their corresponding rendered image. Figure 5.15, 5.16, 5.17 and 5.18 show the example results of browsing along F_{ls} (transition from hard to soft shadows), F_{le} (transition from long to short shadows), F_{la} (transition from light to dark shadows), and F_{lc} (transition from low to high complexity shadows). In each case, some features are constrained while others are not. For example, Figure 5.15 demonstrates modifying the shadow softness feature F_{ls} while the elevation feature F_{le} is constrained. This maintains the shadow length while F_{ls} adjusts the shadow's softness. Other parameters such as the tone features F_t are varying.

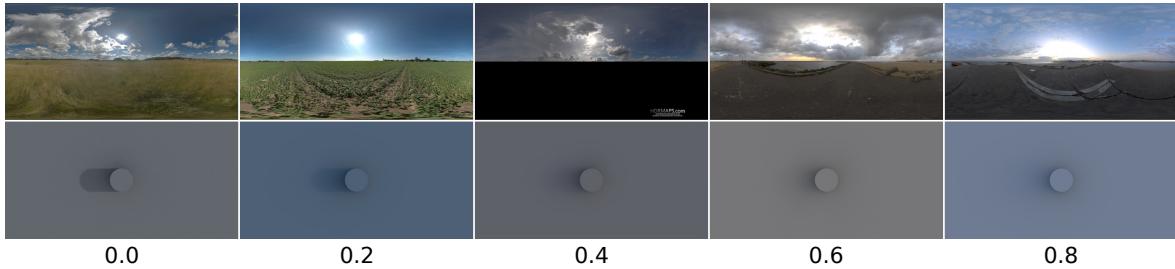


Figure 5.15: Feature F_{ls} , the transition from hard to soft shadows. In the RM, this corresponds to the size of the dominant light source, transitioning from small to large. RMs with similar dominant light elevation F_{le} are chosen.

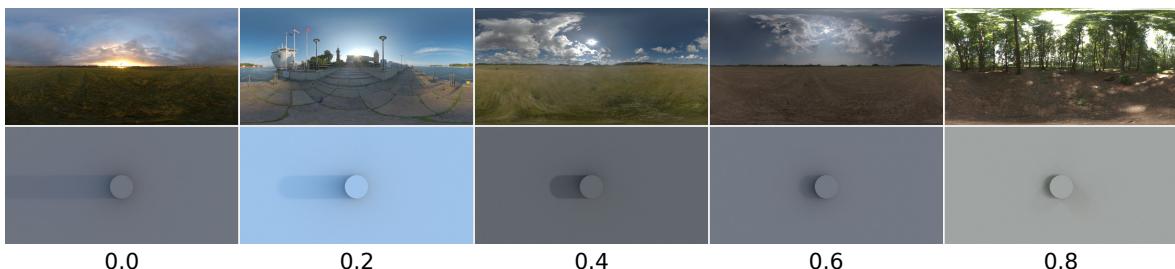


Figure 5.16: Feature F_{le} , the transition from long to short shadows. In the RM, this corresponds to the elevation of the dominant light source, transitioning from low to high. RMs with similar shadow softness F_{ls} are chosen.

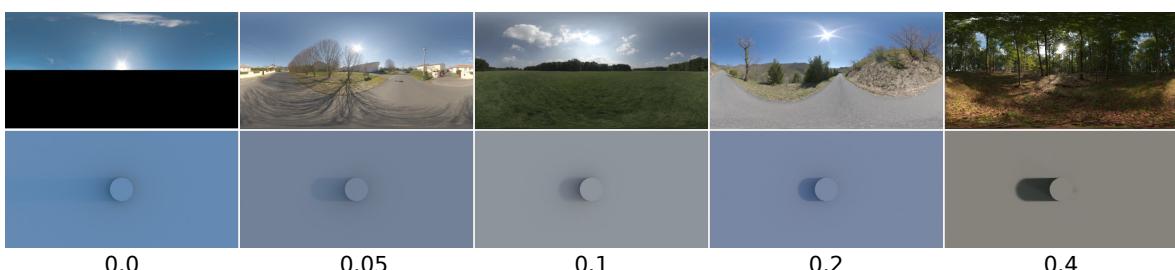


Figure 5.17: Feature F_{la} , the transition from the lightest to the darkest shadow. Specifically, it is the darkest point in the shadow relative to the unoccluded diffuse intensity. In the RM, this corresponds to intensity of the dominant light source relative to the ambient light, transitioning from lowest to highest intensity. RMs with similar shadow softness F_{ls} are chosen.

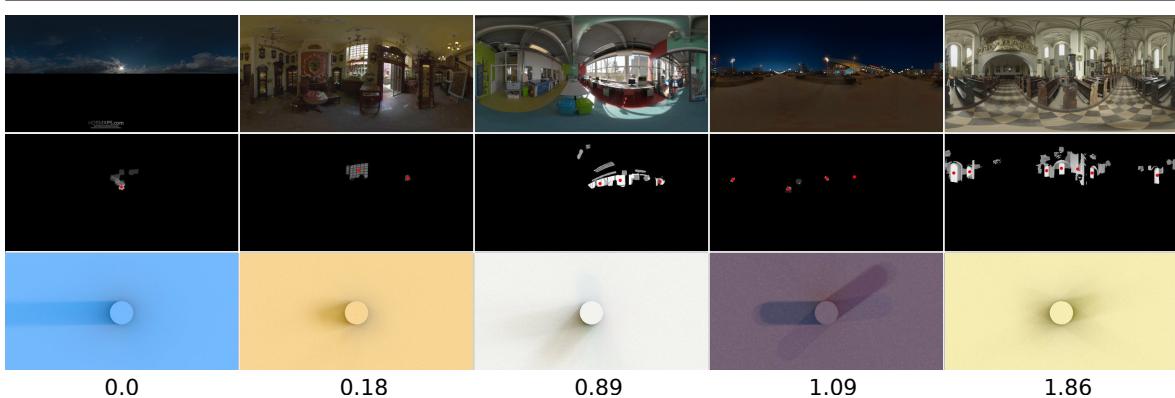


Figure 5.18: Feature F_{lc} , the transition from low to high complexity shadows. The center row shows how the complexity measure corresponds to the number of detected dominant lights (red dots have been added to the center of each detected dominant light source). The rendered images have been normalised to show the detail.

5.7.5 Search using a radiance map

Given a RM, the light features are computed and used as search parameters. This allows users to find RMs having illumination properties similar to the input.

Figure 5.19 shows examples of the illumination search using ambient light features, where the left-most image is the input RM with its corresponding rendered sphere, and the following 4 images are the closest found RMs with similar illumination to the input. The search results are quantitatively evaluated using an objective image measure, structural similarity (SSIM) index [118] on the top 4 results rendered by returned RMs (as shown in Figure 5.19) showing their high similarity. The evaluation can be seen in Table 5.2.

Table 5.2: Quantitative Evaluation: SSIM between input and returned images

	Return 1	Return 2	Return 3	Return 4
Input 1	0.963	0.974	0.970	0.971
Input 2	0.990	0.988	0.977	0.979
Input 3	0.983	0.992	0.982	0.988

As discussed in the previous chapter, a subjective user study with 12 professional visual effects studio artists was also conducted with the ambient tone search. The result is a mean score of 4.0 with a standard deviation of 0.5 for the ambient tone search.

The dominant light search results are compared with Karsch et al.'s model for comparing RMs [25]. Karsch et al. define their distance between two RMs by rendering nine different objects with glossy and diffuse materials. Furthermore, the Wasserstein distance (also known as earth mover's distance (EMD)) is also compared with. The EMD defines a distance between two distributions (using the RM's intensity values). It has shown recent use in computer graphics applications [119, 120], and has shown promise for interpolating RMs [121]. Figure 5.20 results and comparisons of searching using a RM from the RM database.

For our search in Figure 5.20, column 1 demonstrates an automatic search where the weights are evenly distributed for the DML $w_l = 0.5$ and for the AML $w_t = 0.5$. Column 2 and three show artist directed searches. Column 2 aims for more accurate

tone with weights $w_l = 0.03$ and $w_t = 0.97$, and column 3 aims for more accurate shadows with $w_l = 0.8$ and $w_t = 0.2$ respectively. As shown in the Figure 5.20, our method outperforms the others for finding proper RMs having similar illumination properties. We show that our method accurately captures the illumination details in the rendered image.

Both EMD and Karsch et al.'s method are not denoted or explained by features, and as such, don't allow for applications with k-nearest neighbour and clustering algorithms. Therefore, these methods' search speed scale linearly with the size of the database, whereas our light features can be optimised with various different k-nearest neighbour algorithms [93].

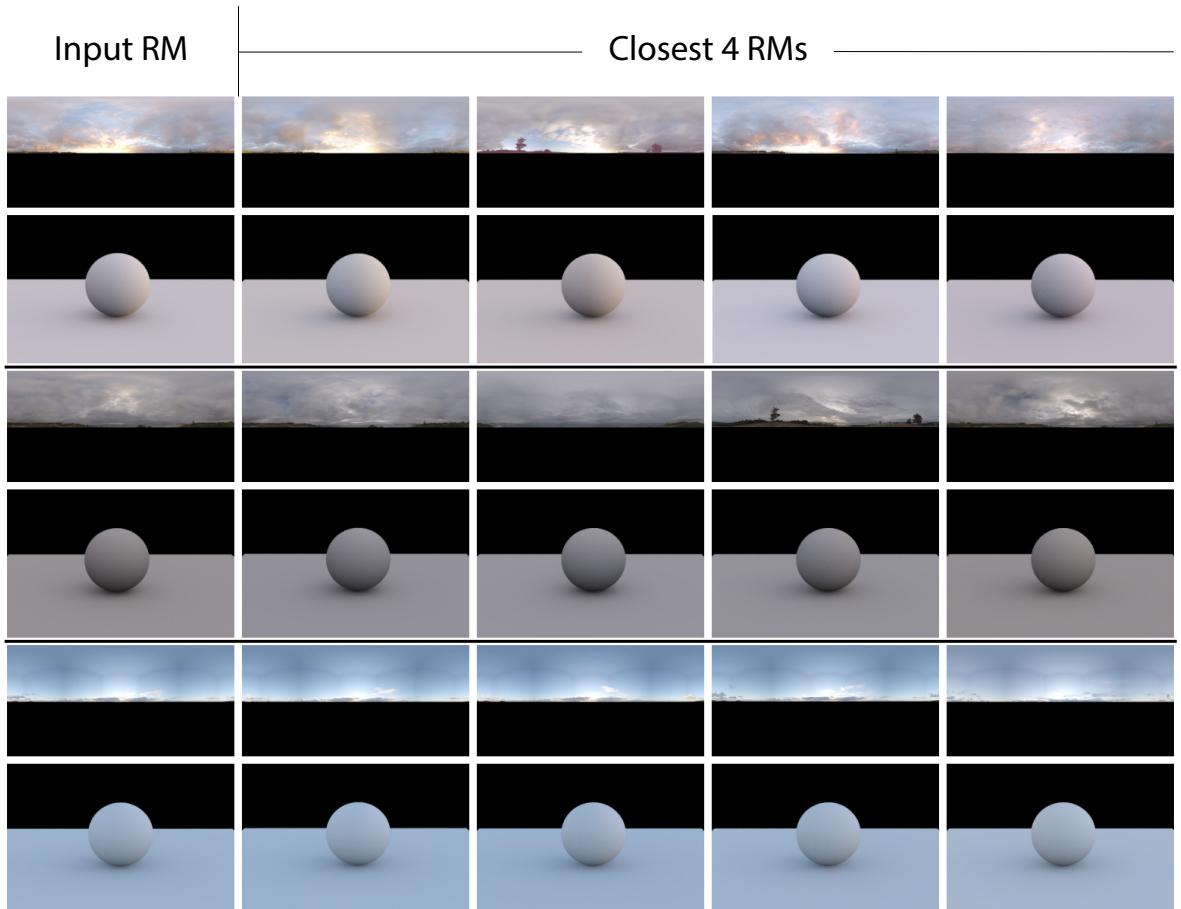


Figure 5.19: The returned results from the ambient tone search. The left-most RM is the input, and the following 4 images are the closest images returned in terms of the ambient tone.

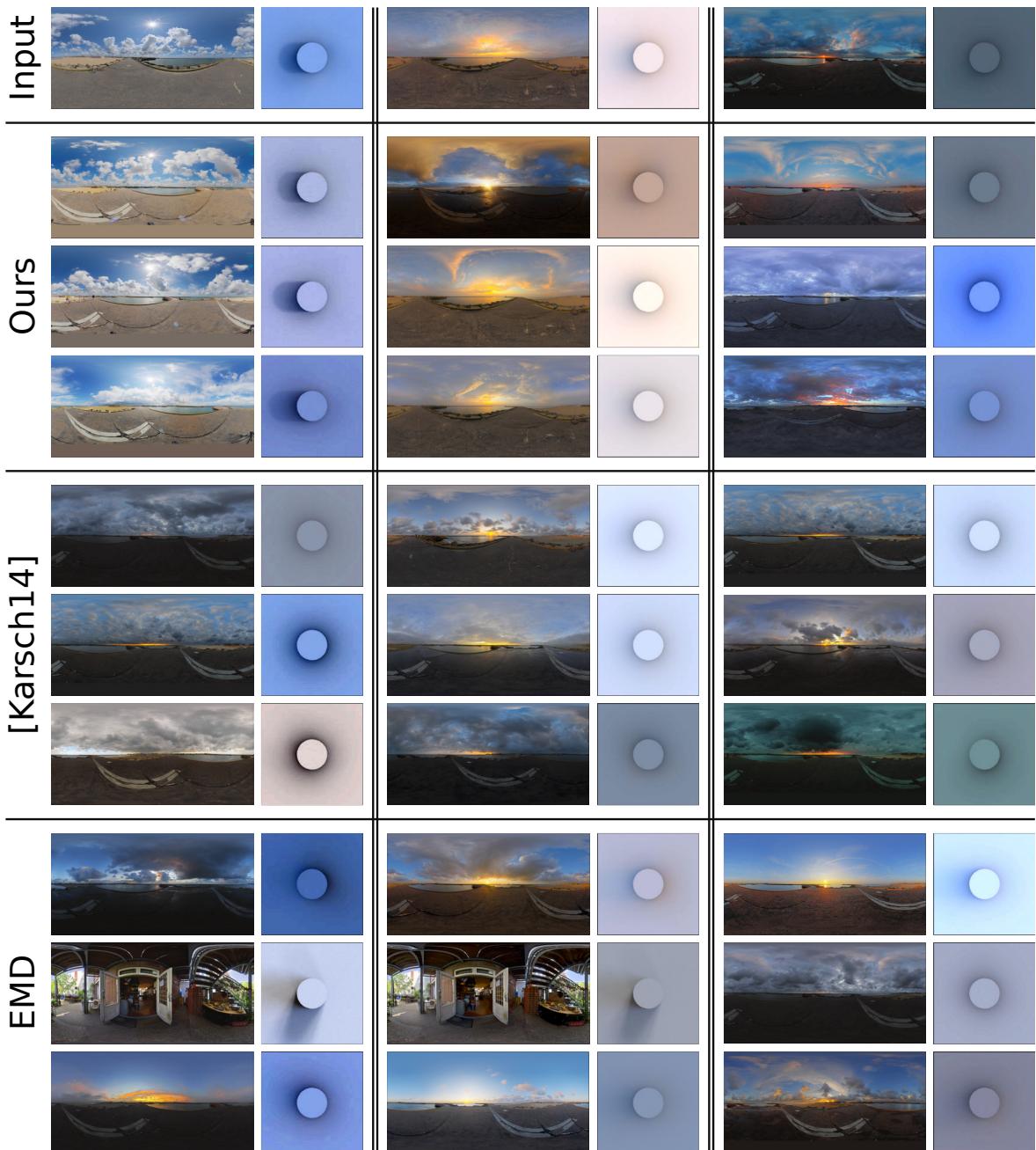


Figure 5.20: The top row is the input RM with its corresponding rendering of a cylinder on a plane. In the same column, the closest 3 RMs to the input RM are shown for each distance measure. Column 1 shows an example with sun light with cloud scattering, column 2 is an example of a red sunset, and column 3 is an overcast blue sky. For our distance measurement, column 2 and 3 are user directed toward tone and shadow detail respectively, whereas column 1 is fully automatic, aiming to find RMs with both similar shadows and tone. The rendered images have been normalised to show the detail.

5.7.6 Search using a shadow profile

Since the DLM provides direct correspondence between the shadow and light source, it is possible to query the database using the light features within a photograph to find a RM which produces similar illumination. A requirement to compute the features from a photograph is to know the shadow profile (i.e. length of the shadow). If this ratio is given, our DLM can be fit to the shadow data and the feature units will have correspondence with the RMs in the illumination space as shown in Figure 5.2. This process can be automated with shadow extraction techniques [122]. See Figure 5.21 and Figure 5.22 for examples of compositing synthetic objects into photographs using the searched RM from the illumination space.

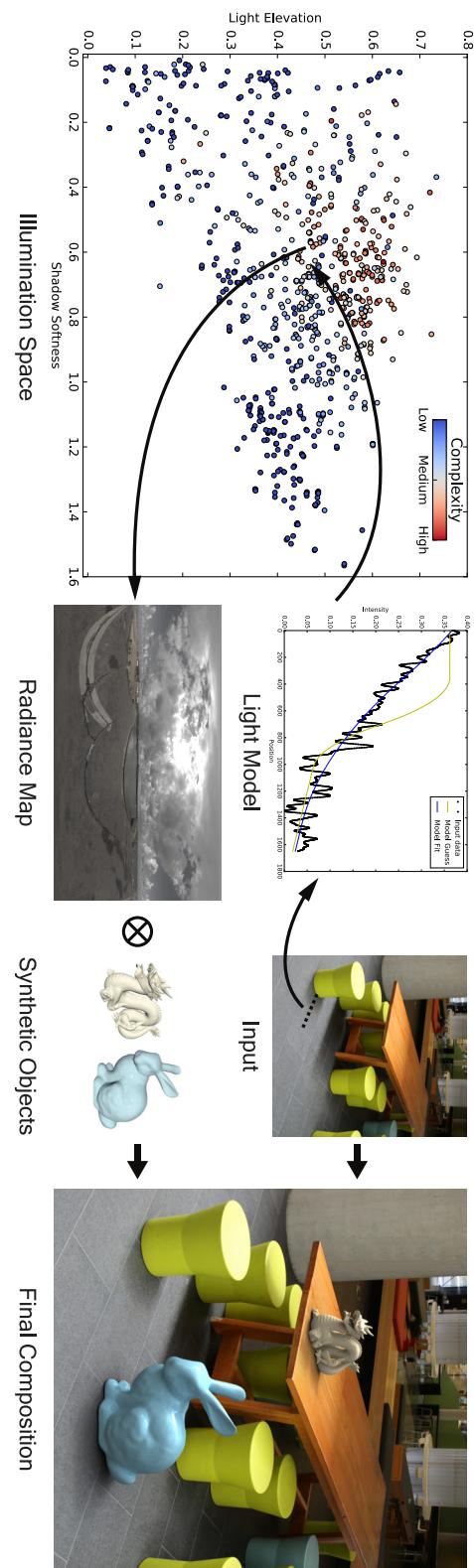


Figure 5.21: Searching through the illumination space with a sunny scene. Given an input photograph and the shadow profile, our light model computes the light features from the shadow information. Using these features, a RM is found in the illumination space. The RM provides similar illumination to the input photograph, allowing for a realistic composition with synthetic objects.

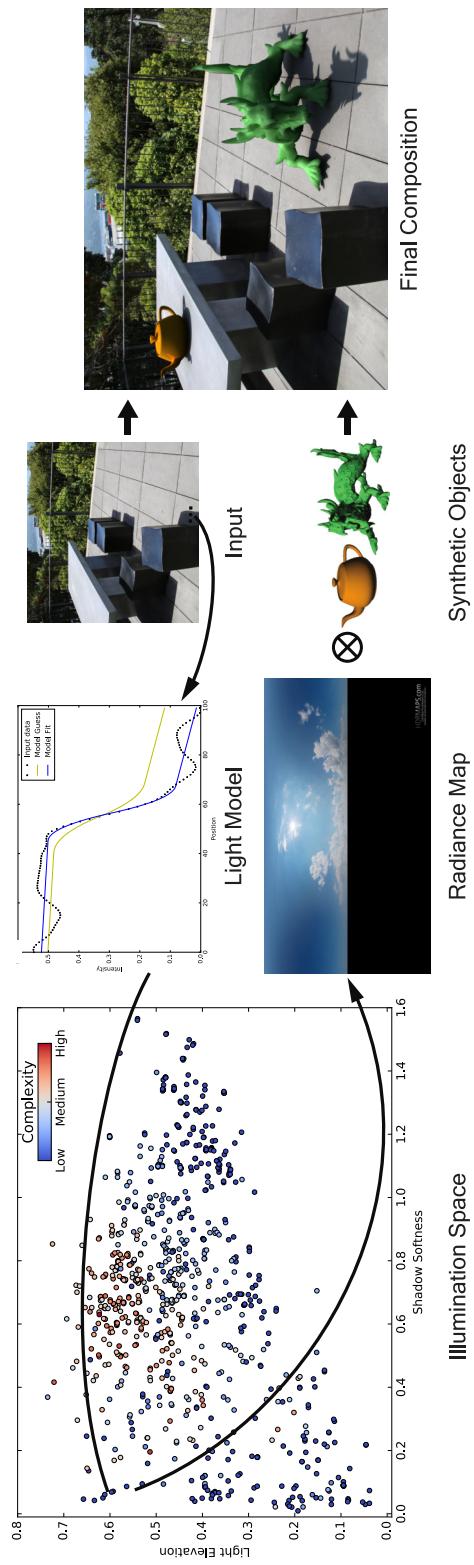


Figure 5.22: Searching through the illumination space with an overcast scene. The RM provides similar illumination to the input photograph, allowing for a realistic composition with synthetic objects.

5.8 Conclusion

This chapter presents a set of novel light features that describe the illumination properties of a RM. The features themselves are concise and intuitive, where there is a direct correspondence between the features in the RM and the rendered image. By introducing a novel dominant light model, light model fitting scheme to a RM, and light detection algorithm, it is possible to automatically and accurately fit the light features to RMs in a large database. Furthermore, an ambient light model describing colour and tone allows for a complete feature descriptor of RMs.

We developed a novel distance measure for measuring the distance between RMs using a weighted combination of our light models. From this, we construct an illumination space that can arrange RM databases using the nine semantic features. The illumination space provides various applications for browsing and searching RMs.

Based on our survey, this is the first formulation of low dimensional features capturing important lighting properties in RMs. Specifically, our light features focus on the semantic illumination properties (as opposed to the texture) of real-world lighting in RMs and the illumination effects on the rendered image. This is also the first time a RM database is semantically organised with a fully automatic pipeline using the light extraction (light model fitting) process. Our illumination space provides a novel approach for browsing and searching the RMs using their illumination properties that has not been provided from the previous methods. As we know, this is the first time a RM database is utilised with low dimensional features capturing important lighting properties. This brings out the full functionality of the RM database, making the rich illumination data set accessible for real-world lighting.

One of the limitations of our method is the use of the local maxima to finalise the light detection algorithm. While the algorithm accurately captures all the necessary pixels, distinguishing between local maximas within a patch area can produce noisy results. This is particularly the case in overcast scenes where many local maximas will be detected for a broad light source, rather than detecting a single

peak. Such error can be mitigated by checking the amplitude of detected peaks. However, improvements on such cases will also improve the overall illumination space and light features. There is some error in the rendered image using the found RM, such as the intensity not always matching the input as shown in Figure 5.20. This is due to a limited number of samples in the database. Obtaining more samples, or interpolating samples in the database can improve these artefacts. The input shadow profiles for Figure 5.21 and Figure 5.22 are obtained manually. Shadow profile extraction is beyond the scope of this thesis. Our contributions will allow for future research to efficiently utilise HDR RM databases for novel applications.

While the features in this chapter describe the shadow and tone properties of the RM, it does not account for textural detail which is an important component for mirror like surfaces. However, the combination of *Sky Browser* from Chapter 4 and the *illumination space* allows for a complete system for searching RMs with varying illumination and textural properties.

Chapter 6

Conclusion

6.1 Summary

The use of IBL and RMs has increased in popularity to mainstream use in visual effects movies, video games, and mixed reality applications. Once a RM is captured, it is stored and ready for repeated use. As such, many studios, vendors and individuals have accumulated RM databases over time. The contributions from this thesis lays down the ground work for leveraging RM databases in various applications. The core contributions include the analysis, definition and processing of RMs. These contributions allow for RM databases to be organised for applications such as browsing, searching and visualising the database.

A number of specific research objectives were stated in the introduction. Here we restate these objectives and relate them to the achievements in the thesis:

- i)* Conduct a perceptual user study to observe the important properties of RMs (Chapter 3).
- ii)* Produce a concise and semantically meaningful set of features for RMs (Chapter 4 and 5 for texture and illumination respectively).
- iii)* Formulate measures and applications for RM databases (Chapter 4 and 5 for texture and illumination respectively).

The user study involved an analysis of RMs and their impact on the rendered image. Given the fact that RMs are commonly used for composition of synthetic objects into real-world scenes, it did not make sense to use traditional pair-wise comparison user studies. A pair-wise study deviates from the fact that an object is being compared with other objects within the same image. In Chapter 3, the perceptual user study is setup to be similar to real-world applications. This involved compositing a synthetic object into a scene with other synthetic objects, where the composited object is illuminated with an altered RM as opposed to the unaltered RM for the other synthetic objects. This style of user study is the foundation used for evaluating the effects of altering RMs as compared with the unaltered "ground truth" RM. From the perceptual user study, various observations of RMs are made. This includes how participants used shadows or highlights as a cue to determine inconsistencies in illumination. These observations not only lead to optimisations to RMs, but also helped indicate the important properties of RMs that need to be represented as features.

The next contribution is to organise a RM database using the observations. Firstly, a traditional approach to organising RMs is adopted. Texture based features are commonly used for standard image databases. However, in this thesis, new features focused on HDR RMs is described. The texture based features allowed users to organise the database purely on the image content. This proved useful for finding RMs with specific reflections or a RM to act as a backdrop to a scene. However, texture based features alone did not characterise all the important elements of a RM; specifically, the lighting properties.

A novel light model and extraction process is used to achieve this task. Since the RMs are HDR, their corresponding rendered images are used to guide this process. The shadows in the rendered image are used to detect the dominant light sources in the RM. From this, an automatic fitting algorithm is used to parameterise each dominant light source. The features are then computed from the parameters of the model.

Distance measures using the features are formulated, allowing for offline or online browsing and searching of the database based on either the texture or illumination properties. Furthermore, since the features are low dimensional, the database is

visualised to improve the user interface and simplicity of browsing.

6.2 Discussion

The organisation of a RM database has been an ongoing problem since RMs have grown in popularity. Previous research used error minimisation techniques to query a RM database. More recently, research has shown interest in using deep convolutional neural networks (CNN) to solve similar tasks. However, these methods use a black-box approach in which a user has little control over the parameters. A semantic and intuitive set of parameters not only achieves the goal of organising the RM database, it also allows for an artist to explore the database. Analogously, this replaces the convolution layers in the CNN with a set of features that directly correspond to important properties of RMs. In order to complete such a task, it was important to understand the background of typical illumination models and how they are presented in software. Chapter 2 explored this area and highlights typical lighting models used by artists. While this is useful to identify how artists handle typical abstract light source parameters (e.g. directionality, position, colour, etc.), it was also clear that there is a large divide in complexity between simple abstract light sources and RMs. This is particularly apparent when a RM is thought of as thousands or millions of individual light sources combined into a single image. Since artists are typically concerned with a few parameters, it was unclear how to describe a RM in a way that was useful for an artist. Thus, the overall complexity of RMs has held back the ways in which the RM database could be organised. Therefore, to achieve the overall goal of this thesis, it was required to find a mapping from highly complex RMs to a low dimensional space.

To gain further insight into the issue, a perceptual user study helped identify how the RM database should be organised. Because the RM is inherently complex, it is much easier to evaluate it based on how it illuminates a scene, rather than observing the image properties of a RM directly. The illumination in an image is comprised of three properties; material models, geometry and lighting. The user study that is conducted in Chapter 3 focused on a specific combination of these three elements. Furthermore, the materials, geometry and lighting conditions are

designed to account for a wide range of combinations, from high frequency illumination (high reflectivity, sharp edges) to low frequency (diffuse reflectivity, soft edges). A preliminary study was also conducted to quickly eliminate unnecessary test conditions. For example, very high resolution RMs (e.g. 20,000x10,000 pixels) did not need to be tested as there were clearly no perceivable differences whatsoever between such high resolutions. The initial setup of the study involved two test cases; the light direction and intensity. Both of these parameters correspond to typical parameters found in abstract light sources. This setup was designed to not only test these two parameters, but to also test the user study setup itself. We found that users were particularly captious in pair-wise comparisons tests. However, the new user study setup showed that users were not able to pick out artefacts as often. Furthermore, the user study setup is ideal for real-world light sources and mixed reality applications since it is common in these applications that the synthetic object is compared with the surrounding scene. While we continued the study for specific optimisations (LDR to HDR and RM resolution), the general feedback and results were indicative of specific properties of RMs that needed to be described.

The features themselves are designed to be as concise as possible. This has two purposes; firstly, it is easier for a user to control a small set of parameters. Furthermore, the curse of dimensionality is a problem in which a higher number of dimensions creates a sparsity in the feature space. The reduced number of clusters will negatively affect both browsing and regression tasks.

However, even if the RM is accurately embedded in a low dimensional space, it may be the case that the desired RM does not exist in the database. To mitigate such issues, two approaches are possible. The first method uses image transfer techniques [123] to combine two RMs into a single RM with both the desirable texture and illumination properties. This can be applied to two different RMs obtained separately from the illumination and texture space.

The second approach synthesises a new RM by interpolating between numerous RMs. However, interpolating between RMs is not a trivial task. Linear interpolation often does not yield desired results when interpolating between distributions. An interpolation scheme that takes into account the entire

distribution needs to be considered. Recent work for interpolating distributions for RMs has shown promising results [121]. This is achieved using ideas from mass transport interpolation and using the earth movers distance as a metric [124].

Outside of image transfer and interpolation, it is common practice for an artist to use two separate passes for rendering. In this setup, an artist can use a RM for shadow rendering, and a separate RM for reflections or backdrops.

6.3 Limitations

Chapter 3 was concluded with examples of compositing into a photograph, whereas the user study was made up of images which were entirely synthetic. Ideally, the entire study is conducted with photographs to accurately replicate real-world situations. We did not conduct such a study because capturing high intensity sunny days is particularly challenging due to limitations of camera hardware. After some trial and error, we concluded that simulating a composition setup would be the best way to minimise the error between the lighting of the composited object and the background scene.

The feature space is dependant on a number of steps in the feature extraction pipeline. The light detection algorithm accurately detects the necessary pixels in the RM, however, the final set of detected lights is based on using the local maxima on the detected pixels. This approach can have artefacts as it does not necessarily detect peaks which correspond to the number of shadows.

6.4 Future Work

The perceptual user study in Chapter 3 focused on a limited number of materials, geometry and illumination models. While we chose specific models to cover a wide range of illumination, the approach can be extended to identify optimal ranges of specific materials, illumination models, or shape variations. Considerations can also be made with video instead of still images. Furthermore, a

comparative study between filtering algorithms for resolution reduction may have effect the image quality. We leave these topics for possible future study. The texture space in Chapter 4 is designed for outdoor skies. Future work to account for indoor HDR RMs may be possible. Using a similar process in this thesis, future work may need to conduct user studies to find important characteristics of indoor scenes to parameterise. As mentioned in the discussion, distribution interpolation can be considered for creating new RMs in the database. However, a current limitation of distribution interpolation is the computation time. Future work in real-time distribution interpolation will allow artists to browse the illumination space continuously. Finally, the limitation of local maxima for determining the final set of lights can be improved by using other techniques which take into account the correct number of shadows in the rendered image.

Appendix

Ethics

There were various user studies conducted in this thesis. Each user study involved image quality or experience evaluation. The following is the generic information sheet, generic consent form, and generic image quality questions used. The exact wording and setup is described in each chapter that has a user study.



Image Quality User Study Information Sheet

Researcher: Andrew Chalmers

School of Engineering and Computer Science, Victoria University of Wellington.

You are invited to take part in a study investigating *image quality*. This study involves one session, during which you will view images on a computer monitor and assess its visual quality. There are three types of questions.

1. You will score an image's quality on a score between 1 and 5, where each score has the following meaning:

- 1** Very bad
- 2** Bad
- 3** Neutral
- 4** Good
- 5** Very good

2. You will be shown pairs of images in which you will determine which image has better image quality.

3. You will be shown a set of images in which you will sort from best to worst based on their visual quality.

Each session will take a maximum of one hour. Data collected during this study will be stored in an anonymous form, in which no personally identifying information about you will be kept. You may withdraw from this study at any time.

This Participant Information Sheet will help you decide if you'd like to take part. It sets out why I am doing the study, what your participation would involve, what the benefits and risks to you might be, and what would happen after the study ends. Feel free to ask the researcher any questions.

If you agree to take part in this study, you will be asked to sign a consent form.

Please make sure you have read and understood both this Information Sheet and the Consent Form before signing.

This research has VUW Human Ethics Committee (HEC) approval.

What are the conditions for my participation in this study?

In order to participate in this study, you must:

- Be over 18 years of age
- Able to view images on a computer screen or on paper
- Not have any serious visual abnormalities (amblyopia, spherical aberrations, etc)

If you have a visual condition and are not sure if it will be an issue, please mention this to the researcher. Note that requiring prescription focal correction (i.e. long or short sightedness) is not an issue. If you are colour blind, you are still welcome to participate in the study, but you need to let the researcher know of your colour blind condition.

What is the purpose of this study?

To determine how well our algorithm works in comparison with user evaluated data.

What will participating in this study involve?

You will be invited to view a series of images for no more than 60 minutes.

You will evaluate the quality of the image through a 1 to 5 scale, through pair-wise comparisons, or by sorting a set of images from highest to lowest image quality.

Again: at any point, you are free to terminate the study. I will ask for a reason, but you do not have to provide one.

What are the possible benefits and/or risks of this study?

There are no risks with this study. The results of this study will benefit the researchers by providing a way of assessing the researcher's algorithms.

What happens once this study ends?

I will remove all identifying information about you from the results, and only store aggregated data: you may request from me these results, but they will not be tailored to you or mention you in any specific manner.

These aggregate results will be stored indefinitely.

The aggregates results of this study will be used in my thesis, and in a proposed research paper to be published this year.

If you have any questions, who can you contact?

Student researcher: **Andrew Chalmers**

email: *chalmeandr@myvuw.ac.nz*

Supervisor: **Taehyun Rhee**

School of Engineering and Computer Science
phone 04 463 5233 x 7088 or email *taehyun.rhee@ecs.vuw.ac.nz*

Human Ethics Committee Information

If you have any concerns about the ethical conduct of the research you may contact the Victoria University HEC Convener: Associate Professor Susan Corbett. Phone 04 463 5480 or email *susan.corbett@vuw.ac.nz*.

Thank you very much for your time and help in making this study possible.



Image Quality User Study Participant Consent Form

Researcher: Andrew Chalmers

School of Engineering and Computer Science, Victoria University of Wellington.

Please read the following notices. If you agree with them, please sign this form.

By signing this sheet: I, the participant, understand and agree that:

- My participation in this research is voluntary and I am aware that I am able to withdraw within 1 week after my participation in the study.
- I am aware that, should I choose to withdraw, that the investigator may ask me why I have chosen to withdraw but I do not have to provide any reasons if I do not want to.
- I confirm that I have been provided, read and understand the Participant Information Sheet.
- I have had the opportunity to ask any questions about this research and had them answered.
- I understand that all personal information will remain confidential and that all efforts will be made to ensure I cannot be identified.
- I agree that data gathered in this study may be stored anonymously and securely, and may be used for future research.
- I agree to take part in this study.
- If I have any further concerns and/or questions, I am aware I can use the following contacts:
 - Student researcher: Andrew Chalmers
email: chalmeandr@myvuw.ac.nz
 - Supervisor: Taehyun Rhee
School of Engineering and Computer Science
phone: 04 463 5233 x 7088 or email: taehyun.rhee@ecs.vuw.ac.nz.
 - Human Ethics Committee Convener: Associate Professor Susan Corbett
phone 04 463 5480 or email susan.corbett@vuw.ac.nz.

.....
Participant Name

.....
Date

.....
Participant Signature

Example survey questions:



Figure 1: What is the visual quality of this image?

Very Bad Bad Neutral Good Very Good

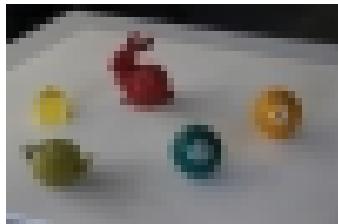
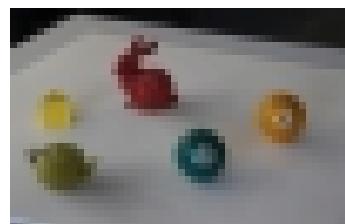


Figure 2: What is the visual quality of this image?

Very Bad Bad Neutral Good Very Good



a



b

Figure 3: Which image has higher visual quality?

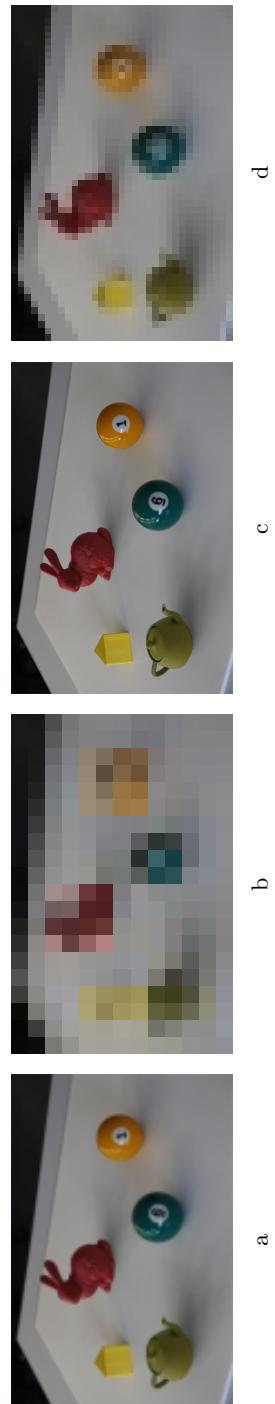


Figure 4: From left to right, sort the images from highest to lowest quality

Bibliography

- [1] "Autodesk [computer software]." Available from <http://www.autodesk.com>. Accessed: 12-12-2014.
- [2] "Blender [computer software]." <http://www.blender.org>. Accessed: 12-12-2014.
- [3] "Photoshop [computer software]." Available from <http://www.photoshop.com>. Accessed: 12-12-2014.
- [4] "The numbers: Movie budget and financial performance records." <http://www.the-numbers.com/movie/budgets/>. Accessed: 12-12-2014.
- [5] Statista, "Video games are getting as expensive as movies." <http://www.statista.com/topics/964/film/>. Accessed: 12-12-2014.
- [6] F. Petrillo, M. Pimenta, F. Trindade, and C. Dietrich, "Houston, we have a problem...: A survey of actual problems in computer games development," in *Proceedings of the 2008 ACM Symposium on Applied Computing*, SAC '08, (New York, NY, USA), pp. 707–711, ACM, 2008.
- [7] A. Brady, J. Lawrence, P. Peers, and W. Weimer, "genbrdf: Discovering new analytic brdfs with genetic programming," *ACM Trans. Graph.*, 2014.
- [8] R. L. Cook and K. E. Torrance, "A reflectance model for computer graphics," *ACM Trans. Graph.*, 1982.
- [9] L. Hosek and A. Wilkie, "An analytic model for full spectral sky-dome radiance," *ACM Trans. Graph.*, 2012.

- [10] B. T. Phong, "Illumination for computer generated pictures," *Commun. ACM*, 1975.
- [11] O. Alexander, M. Rogers, W. Lambeth, M. Chiang, and P. Debevec, "Creating a photoreal digital actor: The digital Emily project," ICT Technical Report ICT TR 04 2009, University of Southern California Institute for Creative Technologies, 2009.
- [12] G. Chen, Y. Dong, P. Peers, J. Zhang, and X. Tong, "Reflectance scanning: Estimating shading frame and brdf with generalized linear light sources," *ACM Trans. Graph.*, 2014.
- [13] K. J. Dana, B. van Ginneken, S. K. Nayar, and J. J. Koenderink, "Reflectance and texture of real-world surfaces," *ACM Trans. Graph.*, 1999.
- [14] P. Graham, B. Tunwattanapong, J. Busch, X. Yu, A. Jones, P. Debevec, and A. Ghosh, "Measurement-based synthesis of facial microgeometry," in *ACM SIGGRAPH 2012 Talks*, SIGGRAPH '12, (New York, NY, USA), pp. 9:1–9:1, ACM, 2012.
- [15] W. Matusik, H. Pfister, M. Brand, and L. McMillan, "A data-driven reflectance model," in *ACM SIGGRAPH 2003 Papers*, ACM, 2003.
- [16] "HDR Light Studio [computer software]." Available from <http://www.hdrlightstudio.com/>. Accessed: 27-12-2014.
- [17] "HDR Shop [computer software]." Available from <http://www.hdrshop.com/>. Accessed: 27-12-2014.
- [18] A. Ben-Artzi, K. Egan, F. Durand, and R. Ramamoorthi, "A precomputed polynomial representation for interactive BRDF editing with global illumination," *ACM Trans. Graph.*, 2008.
- [19] M. Colbert, S. Pattanaik, and J. Krivanek, "Brdf-shop: creating physically correct bidirectional reflectance distribution functions," *Computer Graphics and Applications, IEEE*, 2006.

- [20] A. Forés, S. N. Pattanaik, C. Bosch, and X. Pueyo, “BRDFLab: A general system for designing BRDFs,” in *Proceedings CEIG’09*, Eurographics, 2009.
- [21] J. T. Kajiya, “The rendering equation,” in *Proceedings of the 13th Annual Conference on Computer Graphics and Interactive Techniques*, SIGGRAPH ’86, (New York, NY, USA), pp. 143–150, ACM, 1986.
- [22] P. Debevec, “Rendering synthetic objects into real scenes: Bridging traditional and image-based graphics with global illumination and high dynamic range photography,” in *Proceedings of SIGGRAPH ’98*, pp. 189–198, ACM, 1998.
- [23] G. S. Miller and C. R. Hoffman, “Illumination and reflection maps: Simulated objects in simulated and real environments,” *ACM Trans. Graph.*, 1984.
- [24] T. Grosch, “Differential photon mapping: Consistent augmentation of photographs with correction of all light paths,” in *Eurographics 2005 Short Papers, Trinity College, Dublin, Ireland*, 2005.
- [25] K. Karsch, K. Sunkavalli, S. Hadap, N. Carr, H. Jin, R. Fonte, M. Sittig, and D. Forsyth, “Automatic scene inference for 3d object compositing,” *ACM Trans. Graph.*, 2014.
- [26] M. Knecht, C. Traxler, O. Mattausch, W. Purgathofer, and M. Wimmer, “Differential instant radiosity for mixed reality,” in *Mixed and Augmented Reality (ISMAR), 2010 9th IEEE International Symposium on*, pp. 99–107, IEEE, 2010.
- [27] P. E. Debevec and J. Malik, “Recovering high dynamic range radiance maps from photographs,” in *Proceedings of SIGGRAPH ’97*, pp. 369–378, 1997.
- [28] J. Stumpfel, A. Jones, A. Wenger, C. Tchou, T. Hawkins, and P. Debevec, “Direct HDR capture of the sun and sky,” in *ACM SIGGRAPH 2006 Courses*, SIGGRAPH ’06, (New York, NY, USA), ACM, 2006.
- [29] Y. Gryaditskya, T. Pouli, E. Reinhard, and H.-P. Seidel, “Sky based light metering for high dynamic range images,” *Computer Graphics Forum*, 2014.

- [30] A. J. Preetham, P. Shirley, and B. Smits, "A practical analytic model for daylight," in *Proc. of SIGGRAPH '99*, ACM Press/Addison-Wesley Publishing Co., 1999.
- [31] Y. Uetani, S. Aydinli, A. Joukoff, J. Kendrick, R. Kittler, Y. Koga, *et al.*, "Spatial distribution of daylight-cie standard general sky," *Vienna, Austria*, 2003.
- [32] J. Haber, M. Magnor, and H.-P. Seidel, "Physically-based simulation of twilight phenomena," *ACM Trans. Graph.*, 2005.
- [33] A. Wilkie and L. Hošek, "Predicting sky dome appearance on earth-like extrasolar worlds," in *Proceedings of the 29th Spring Conference on Computer Graphics*, ACM, 2013.
- [34] C. Bloch, C. Bauer, V. Heisterberge, G. Reddick, and chris Huf, "Smart-ibl." <http://www.hdrlabs.com/sibl/>, 2008. Accessed: 09-04-2015.
- [35] "OpenFootage [radiance map database]." <http://www.openfootage.net/>. Accessed: 31-03-2015.
- [36] "CG Skies [radiance map database]." <https://www.cgskies.com>. Accessed: 08-02-2018.
- [37] "Dutch Skies 360 [radiance map database]." <http://www.dutch360hdr.com/>. Accessed: 05-05-2015.
- [38] "HDRI Haven [radiance map database]." <https://hdrihaven.com>. Accessed: 08-02-2018.
- [39] "HDRI Hub [radiance map database]." <http://www.hdri-hub.com/>. Accessed: 26-03-2015.
- [40] "HDRI-Skies [radiance map database]." <http://hdri-skies.com>. Accessed: 08-02-2018.
- [41] "HDR Maps [radiance map database]." [http://hdrmaps.com/](http://hdrmaps.com). Accessed: 25-05-2015.

- [42] "HDR Source [radiance map database]." <http://www.hdrsource.com/>. Accessed: 05-05-2015.
- [43] "Hyper Focal Design [radiance map database]." <http://www.hyperfocaldesign.com/>. Accessed: 08-02-2018.
- [44] M. Gardner, K. Sunkavalli, E. Yumer, X. Shen, E. Gambaretto, C. Gagné, and J. Lalonde, "Learning to predict indoor illumination from a single image," *CoRR*, vol. abs/1704.00090, 2017.
- [45] J. Xiao, K. Ehinger, A. Oliva, and A. Torralba, "Recognizing scene viewpoint using panoramic place representation," in *Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on*, 2012.
- [46] I. Yu, A. Cox, M. H. Kim, T. Ritschel, T. Grosch, C. Dachsbacher, and J. Kautz, "Perceptual influence of approximate visibility in indirect illumination," *ACM Trans. Appl. Percept.*, vol. 6, pp. 24:1–24:14, Oct. 2009.
- [47] J. Křivánek, J. A. Ferwerda, and K. Bala, "Effects of global illumination approximations on material appearance," in *ACM SIGGRAPH 2010 papers*, SIGGRAPH '10, pp. 112:1–112:10, ACM, 2010.
- [48] H. Landis, "Production-ready global illumination," in *Siggraph Course Notes 16*, 2002.
- [49] E. Nakamae, K. Harada, T. Ishizaki, and T. Nishita, "A montage method: the overlaying of the computer generated images onto a background photograph," in *Proceedings of the 13th annual conference on Computer graphics and interactive techniques*, SIGGRAPH '86, pp. 207–214, ACM, 1986.
- [50] A. Fournier, A. S. Gunawan, and C. Romanzin, "Common illumination between real and computer generated scenes," in *Graphics Interface*, pp. 254–254, CANADIAN INFORMATION PROCESSING SOCIETY, 1993.
- [51] A. A. Mury, S. C. Pont, and J. J. Koenderink, "Representing the light field in finite three-dimensional spaces from sparse discrete samples," *Appl. Opt.*, vol. 48, pp. 450–457, Jan 2009.

- [52] J.-F. Lalonde, A. A. Efros, and S. G. Narasimhan, "Webcam clip art: appearance and illuminant transfer from time-lapse sequences," in *ACM SIGGRAPH Asia 2009 papers*, SIGGRAPH Asia '09, pp. 131:1–131:10, ACM, 2009.
- [53] P. Debevec, "Rendering synthetic objects into real scenes: bridging traditional and image-based graphics with global illumination and high dynamic range photography," in *Proceedings of the 25th annual conference on Computer graphics and interactive techniques*, SIGGRAPH '98, pp. 189–198, ACM, 1998.
- [54] J. Lopez-Moreno, V. Sundstedt, F. Sangorrin, and D. Gutierrez, "Measuring the perception of light inconsistencies," in *Proceedings of the 7th Symposium on Applied Perception in Graphics and Visualization*, APGV '10, pp. 25–32, ACM, 2010.
- [55] K. Karsch, V. Hedau, D. Forsyth, and D. Hoiem, "Rendering synthetic objects into legacy photographs," in *Proceedings of the 2011 SIGGRAPH Asia Conference*, SA '11, pp. 157:1–157:12, ACM, 2011.
- [56] C. O'Sullivan, S. Howlett, Y. Morvan, R. McDonnell, and K. O'Conor, "Perceptually Adaptive Graphics," in *Eurographics State of the Art Reports* (EG'04), pp. 141 – 164, 2004.
- [57] D. Bartz, D. Cunningham, J. Fischer, and C. Wallraven, "The Role of Perception for Computer Graphics," in *Eurographics State-of-the-Art-Reports*, pp. 65–86, 2008.
- [58] G. Ramanarayanan, J. Ferwerda, B. Walter, and K. Bala, "Visual equivalence: towards a new standard for image fidelity," in *ACM SIGGRAPH 2007 papers*, SIGGRAPH '07, ACM, 2007.
- [59] J. A. Ferwerda, G. Ramanarayanan, B. Walter, and K. Bala, "Visual equivalence: an object-based approach to image quality," in *Proc. of IS&T 16th Color Imaging Conference*, pp. 347–354, 2008.
- [60] P. Vangorp, T. S. Condon, J. A. Ferwerda, K. Bala, R. Schoukens, and P. Dutré, "Visual equivalence in dynamic scenes," 2009.

- [61] G. Nakano, I. Kitahara, and Y. Ohta, "Generating perceptually-correct shadows for mixed reality," in *Proceedings of the 7th IEEE/ACM International Symposium on Mixed and Augmented Reality*, ISMAR '08, (Washington, DC, USA), pp. 173–174, IEEE Computer Society, 2008.
- [62] P. Vangorp, J. Laurijssen, and P. Dutré, "The influence of shape on the perception of material reflectance," in *ACM SIGGRAPH 2007 papers*, SIGGRAPH '07, ACM, 2007.
- [63] F. Banterle, P. Ledda, K. Debattista, A. Chalmers, and M. Bloj, "A framework for inverse tone mapping," *The Visual Computer*, vol. 23, no. 7, pp. 467–478, 2007.
- [64] S. J. Daly, "Visible differences predictor: an algorithm for the assessment of image fidelity," in *SPIE/IS&T 1992 Symposium on Electronic Imaging: Science and Technology*, pp. 2–15, International Society for Optics and Photonics, 1992.
- [65] "Visual difference predictor." <http://www.mpi-sb.mpg.de/resources/hdr/vdp/>. Accessed: 26-03-2015.
- [66] F. Banterle, K. Debattista, A. Artusi, S. Pattanaik, K. Myszkowski, P. Ledda, and A. Chalmers, "High dynamic range imaging and low dynamic range expansion for generating HDR content," in *Computer Graphics Forum*, 28, 2009.
- [67] A. Chalmers, J. J. Choi, and T. Rhee, "Perceptually based radiance map for realistic composition," in *Proc. of Image and Vision Computing New Zealand (IVCNZ)*, pp. 172 – 177, 2013.
- [68] M. Pharr and G. Humphreys, *Physically Based Rendering, Second Edition: From Theory To Implementation*. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 2nd ed., 2010.
- [69] M. Fajardo, "Arnold software," 2013.
- [70] R. W. Fleming, A. Torralba, and E. H. Adelson, "Specular reflections and the perception of shape," *J. Vis.*, vol. 4, pp. 798–820, 9 2004.

- [71] M. Meilgaard, G. Civille, and B. Carr, *Sensory evaluation techniques*. CRC Press, 1991.
- [72] T. Nishita, T. Sirai, K. Tadamura, and E. Nakamae, "Display of the earth taking into account atmospheric scattering," in *Proceedings of SIGGRAPH '92*, pp. 175–182, ACM, 1993.
- [73] Y. Dobashi, T. Nishita, K. Kaneda, and H. Yamashita, "A fast display method of sky color using basis functions," in *The Journal of Visualization and Computer Graphics*, pp. 115–127, 1997.
- [74] L. Tao, L. Yuan, and J. Sun, "Skyfinder: Attribute-based sky image search," in *ACM SIGGRAPH 2009 Papers*, SIGGRAPH '09, pp. 68:1–68:5, ACM, 2009.
- [75] Y. Yue, K. Iwasaki, B.-Y. Chen, Y. Dobashi, and T. Nishita, "Unbiased, adaptive stochastic sampling for rendering inhomogeneous participating media," *ACM Trans. Graph.*, vol. 29, pp. 177:1–177:8, Dec. 2010.
- [76] Y. Dobashi, W. Iwasaki, A. Ono, T. Yamamoto, Y. Yue, and T. Nishita, "An inverse problem approach for automatically adjusting the parameters for rendering clouds using photographs," *ACM Trans. Graph.*, vol. 31, pp. 145:1–145:10, Nov. 2012.
- [77] R. V. Klassen, "Modeling the effect of the atmosphere on light," *ACM Trans. Graph.*, vol. 6, pp. 215–237, July 1987.
- [78] R. Datta, D. Joshi, J. Li, and J. Z. Wang, "Image retrieval: Ideas, influences, and trends of the new age," *ACM Comput. Surv.*, vol. 40, pp. 5:1–5:60, May 2008.
- [79] B. Johnson, M., S. G. J., A. J., V. O., Kwatra, and R. Cipolla, "Semantic photo synthesis," *Computer Graphics Forum*, vol. 25, no. 2, pp. 407–412, 2006.
- [80] J.-F. Lalonde, D. Hoiem, A. A. Efros, C. Rother, J. Winn, and A. Criminisi, "Photo clip art," *ACM Trans. Graph.*, vol. 26, July 2007.
- [81] P.-Y. Laffont, Z. Ren, X. Tao, C. Qian, and J. Hays, "Transient attributes for high-level understanding and editing of outdoor scenes," *ACM Transactions on Graphics (proceedings of SIGGRAPH 2014)*, vol. 33, no. 4, 2014.

- [82] R. Haralick, K. Shanmugam, and I. Dinstein, "Textural features for image classification," *IEEE Trans. Systems, Man and Cybernetics*, vol. SMC-3, no. 6, pp. 610–621, 1973.
- [83] Z. Gu, C. Duncan, E. Renshaw, M. Mugglestone, C. Cowan, and P. Grant, "Comparison of techniques for measuring cloud texture in remotely sense satellite meteorological image data." IEE Proceedings F (Radar and Signal Processing), Volume 136, Issue 5, 1989.
- [84] H. Chethan, R. Raghavendra, and G. Kumar, "Texture based approach for cloud classification using SVM." Proc. of ARTCom 2009, 2009.
- [85] D. Mazzoni, A. Horvath, M. Garay, B. Tang, and R. Davies, "A MISR cloud-type classifier using reduced support vector machines." Eighth Workshop on Mining Scientific, 2005.
- [86] A. Heinle, A. Macke, and S. Srivastav, "Automatic cloud classification of whole sky images." Atmospheric Measurement Techniques, Volume 3, Issue 3, 2010.
- [87] A. Ono, Y. Dobashi, and T. Yamamoto, "A system for editing sky images using an image database," in *SIGGRAPH Asia 2011 Sketches*, SA '11, pp. 38:1–38:2, ACM, 2011.
- [88] T. Mitani and I. Fujishiro, "Cosmicai: Generating sky backgrounds through content-based search and flexible composition," in *ACM SIGGRAPH 2012 Posters*, SIGGRAPH '12, pp. 52:1–52:1, ACM, 2012.
- [89] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *International Journal of Computer Vision* 2004, vol. 60, no. 2, pp. 91–110, 2004.
- [90] F. Albrechtsen, "Statistical texture measures computed from gray level cooccurrence matrices," 2008. accessed 14 June 2014.
- [91] T. Funkhouser, P. Min, M. Kazhdan, J. Chen, A. Halderman, D. Dobkin, and D. Jacobs, "A search engine for 3d models," *ACM Trans. Graph.*, vol. 22, pp. 83–105, Jan. 2003.

- [92] R. Ramamoorthi and P. Hanrahan, "Frequency space environment map rendering," in *Proc. of SIGGRAPH '02*, ACM, 2002.
- [93] M. Muja and D. G. Lowe, "Scalable nearest neighbor algorithms for high dimensional data," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 36, 2014.
- [94] A. Chalmers, J. Lewis, P. Hillman, C. Tait, and T. Rhee, "Sky Browser: Search for HDR Sky Maps," in *Pacific Graphics*, The Eurographics Association, 2014.
- [95] R. Ramamoorthi and P. Hanrahan, "An efficient representation for irradiance environment maps," in *Proceedings of the 28th annual conference on Computer graphics and interactive techniques*, pp. 497–500, ACM, 2001.
- [96] P.-P. Sloan, J. Kautz, and J. Snyder, "Precomputed radiance transfer for real-time rendering in dynamic, low-frequency lighting environments," in *ACM Transactions on Graphics (TOG)*, vol. 21, pp. 527–536, ACM, 2002.
- [97] R. Ng, R. Ramamoorthi, and P. Hanrahan, "All-frequency shadows using non-linear wavelet lighting approximation," in *ACM SIGGRAPH 2003 Papers*, ACM, 2003.
- [98] A. Mohan, J. Tumblin, and P. Choudhury, "Editing soft shadows in a digital photograph," *IEEE Computer Graphics and Applications*, vol. 27, no. 2, pp. 23–31, 2007.
- [99] J. F. Lalonde and I. Matthews, "Lighting estimation in outdoor image collections," in *2014 2nd International Conference on 3D Vision*, vol. 1, pp. 131–138, Dec 2014.
- [100] M. Laskowski, "Detection of light sources in digital photographs," in *In 11th Central European Seminar on Computer Graphics*, 2007.
- [101] P. Debevec, "A median cut algorithm for light probe sampling," in *ACM SIGGRAPH 2006 Courses*, ACM, 2006.

- [102] S. Agarwal, R. Ramamoorthi, S. Belongie, and H. W. Jensen, "Structured importance sampling of environment maps," *ACM Trans. Graph.*, vol. 22, pp. 605–612, July 2003.
- [103] H. Lu, R. Pacanowski, and X. Granier, "Real-time importance sampling of dynamic environment maps," in *EG 2013 - Short Papers*, 2013.
- [104] L. Sun, T. Yamasaki, and K. Aizawa, "Relationship between visual complexity and aesthetics: Application to beauty prediction of photos," in *Workshop at the European Conference on Computer Vision*, pp. 20–34, Springer, 2014.
- [105] Q. Tian and J. J. Clark, "Real-time specularity detection using unnormalized wiener entropy," in *Computer and Robot Vision (CRV), 2013 International Conference on*, pp. 356–363, IEEE, 2013.
- [106] M. K. Stern and J. H. Johnson, "Just noticeable difference," *Corsini Encyclopedia of Psychology*, 2010.
- [107] G. Mather, *Essentials of sensation and perception*. Routledge, 2014.
- [108] R. Achanta, A. Shaji, K. Smith, A. Lucchi, P. Fua, and S. Susstrunk, "Slic superpixels compared to state-of-the-art superpixel methods," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 34, pp. 2274–2282, Nov. 2012.
- [109] J. Zujovic, T. N. Pappas, and D. L. Neuhoff, "Structural similarity metrics for texture analysis and retrieval," in *2009 16th IEEE International Conference on Image Processing (ICIP)*, pp. 2225–2228, IEEE, 2009.
- [110] X. Zhao, M. G. Reyes, T. N. Pappas, and D. L. Neuhoff, "Structural texture similarity metrics for retrieval applications," in *2008 15th IEEE International Conference on Image Processing*, pp. 1196–1199, IEEE, 2008.
- [111] J. Zujovic, T. N. Pappas, and D. L. Neuhoff, "Perceptual similarity metrics for retrieval of natural textures," in *Multimedia Signal Processing, 2009. MMSP'09. IEEE International Workshop on*, pp. 1–5, IEEE, 2009.

- [112] C. Voglis and I. Lagaris, "A rectangular trust region dogleg approach for unconstrained and bound constrained nonlinear optimization," in *WSEAS Conference*, pp. 17–19, 2004.
- [113] J. Rigau, M. Feixas, and M. Sbert, "Informational aesthetics measures," *IEEE Computer Graphics and Applications*, vol. 28, no. 2, 2008.
- [114] M. R. Luo, G. Cui, and B. Rigg, "The development of the cie 2000 colour-difference formula: Ciede2000," *Color Research & Application*, vol. 26, no. 5, pp. 340–350, 2001.
- [115] A. N. Tuch, J. A. Bargas-Avila, K. Opwis, and F. H. Wilhelm, "Visual complexity of websites: Effects on users' experience, physiology, performance, and memory," *International journal of human-computer studies*, vol. 67, no. 9, pp. 703–715, 2009.
- [116] G. Ramanarayanan, K. Bala, J. A. Ferwerda, and B. Walter, "Dimensionality of visual complexity in computer graphics scenes," in *Electronic Imaging 2008*, pp. 68060E–68060E, International Society for Optics and Photonics, 2008.
- [117] J. B. Kruskal and M. Wish, *Multidimensional scaling*, vol. 11. Sage, 1978.
- [118] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE TRANSACTIONS ON IMAGE PROCESSING*, 2004.
- [119] N. Bonneel, G. Peyré, and M. Cuturi, "Wasserstein barycentric coordinates: Histogram regression using optimal transport," *ACM Transactions on Graphics (Proceedings of SIGGRAPH 2016)*, vol. 35, no. 4, 2016.
- [120] J. Solomon, F. de Goes, G. Peyré, M. Cuturi, A. Butscher, A. Nguyen, T. Du, and L. Guibas, "Convolutional wasserstein distances: Efficient optimal transportation on geometric domains," *ACM Trans. Graph.*, vol. 34, pp. 66:1–66:11, July 2015.
- [121] N. Bonneel, M. van de Panne, S. Paris, and W. Heidrich, "Displacement interpolation using Lagrangian mass transport," in *Proceedings of the 2011 SIGGRAPH Asia Conference*, ACM, 2011.

-
- [122] M. Gryka, M. Terry, and G. J. Brostow, "Learning to remove soft shadows," *ACM Transactions on Graphics (TOG)*, vol. 34, no. 5, p. 153, 2015.
 - [123] A. Hertzmann, C. E. Jacobs, N. Oliver, B. Curless, and D. H. Salesin, "Image analogies," in *Proceedings of the 28th annual conference on Computer graphics and interactive techniques*, pp. 327–340, ACM, 2001.
 - [124] Y. Rubner, C. Tomasi, and L. Guibas, "A metric for distributions with applications to image databases," in *Computer Vision, 1998. Sixth International Conference on*, 1998.

