

Automatické pořízení a optimální aplikace světelných map prostředí

Automatic Acquisition and Optimal Application
of Illumination Environments Maps

Bc. Jaroslav Stančík

Master Thesis
2010



Univerzita Tomáše Bati ve Zlíně
Fakulta aplikované informatiky

ZADÁNÍ DIPLOMOVÉ PRÁCE

(PROJEKTU, UMĚLECKÉHO DÍLA, UMĚLECKÉHO VÝKONU)

Jméno a příjmení: **Bc. Jaroslav STANČÍK**
Studijní program: **N 3902 Inženýrská informatika**
Studijní obor: **Automatické řízení a informatika**

Téma práce: **Automatické pořízení a optimální aplikace světelných map prostředí**

Zásady pro vypracování:

1. Nastudujete metody pořízení HDR snímků a jejich zobrazení na běžných zobrazovacích zařízeních.
2. Nastudujte principy pořízení světelných map prostředí (skládáním panorama, snímáním odrazu na kouli, širokoúhlým objektivem).
3. Použijte některý z principů pro rychlé/automatické pořízení mapy prostředí v LDR případně HRD formátu.
4. Seznamte se s metodami reprezentace map prostředí skupinou bodových světél a implementujte jednu z nich.
5. Implementujte aplikaci pro osvětlení virtuálních objektů pomocí změřených map.
6. Proveďte psychovizuální experiment se skupinou dobrovolníků, porovnávající kvalitu vykreslení 3D objektu osvětleného pomocí pořízených map různých typů prostředí s proměnným počtem světél.
7. Na základě analýzy výsledků určete typy a vizuální rysy změřených map, které výrazněji ovlivňují lidský vizuální vjem.

Rozsah práce:

Rozsah příloh:

Forma zpracování diplomové práce: **tištěná/elektronická**

Seznam odborné literatury:

1. E. REINHARD, G. WARD, S. PATTANAIK, P. DEBEVEC. High Dynamic Range Imaging (Acquisition, Display and Image-Based Lighting), , Morgan Kaufmann Publishers, 2006, ISBN 10: 0-12-585263-0
2. G. WARD, E. REINHARD, P. DEBEVEC. High Dynamic Range Imaging & Image-based Lighting SIGGRAPH 2008 Half-day Class,
3. Perceiving illumination inconsistencies in scenes, Yuri Ostrovsky, Patrick Cavanagh, Perception, 2005
4. A comparison of material and illumination discrimination performance for real rough, real smooth and computer generated smooth spheres, Susan F. te Pas, Sylvia C. Pont, 2005, ISBN:1-59593-139-2
5. Real-world illumination and the perception of surface reflectance properties, Fleming, R. W., Dror, R. O., & Adelson, E. H., 2003, Journal of Vision, 3(5):3, 347-368, <http://journalofvision.org/3/5/3/>,doi:10.1167/3.5.3.

Vedoucí diplomové práce:

Ing. Petr Chalupa, Ph.D.

Ústav řízení procesů

Datum zadání diplomové práce:

19. února 2010

Termín odevzdání diplomové práce:

8. června 2010

Konzultant:

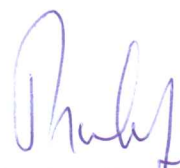
Jiří Filip

Ve Zlíně dne 19. února 2010



prof. Ing. Vladimír Vašek, CSc.

děkan



prof. Ing. Vladimír Vašek, CSc.

ředitel ústavu

Abstract

The visual quality of digital material appearance representation strongly depends, among others, on representation of realistic illumination conditions. While accurate material appearance can be represented by means of measured illumination and view direction dependent textures (i.e., bidirectional texture function), the illumination conditions can be captured in omnidirectional image (i.e., environment map) that can be, for purpose of fast interactive visualization, represented by an appropriate finite set of point-lights.

In this thesis we used two ways of high-dynamic-range environment map acquisition obtained either by means of fish-eye photo-lenses, or by taking pictures of mirrored sphere. The main goal of this thesis was to analyse human visual perception of three different materials illuminated by ten different types of realistic illumination conditions. We performed two psychophysical experiments with 29 naive volunteers to determine appropriate number of lights necessary to achieve the same visual quality across different materials and illumination types. As a result of data analysis from the experiment we suggested a computationally simple method that can predict the appropriate number of lights for any environment map while still preserving the required visual quality. This method successfully decrease the processing times during rendering and still maintains the realistic visual appearance of any digital representation of real-world materials illumination.

Keywords: environment illumination, texture, BTF, HDR, psychophysical experiment, median-cut, tone-mapping.

Acknowledgements

I would like to thank to Jiří Filip for his extraordinary cheerful help with everything concerned with this thesis. Just he knows what I mean.

Further I would like to thank to

Peter Chalupa and Miroslav Červenka, without their willingness it would not be possible to cooperate with the external supervisor and this thesis will never see the light of day.

Michal Haindl for the borrowing fisheye lenses and taking opponency of this thesis.

Francesco Banterle for borrowing his version of median cut code.

Paja, and Robo for their willing assistance with the photographing.

All people who voluntary participated in the experiments.

My last thanks belongs to my mum who cheerfully supported me during all my studies.

Motto

human being belongs to the 6% of species which have two eyes.
one of the biggest advantage of having two
is that one can still use one, when one is lost.

anonym

Prohlašuji, že

- beru na vědomí, že odevzdáním bakalářské práce souhlasím se zveřejněním své práce podle zákona č. 111/1998 Sb. o vysokých školách a o změně a doplnění dalších zákonů (zákon o vysokých školách), ve znění pozdějších právních předpisů, bez ohledu na výsledek obhajoby;
- beru na vědomí, že bakalářská práce bude uložena v elektronické podobě v univerzitním informačním systému dostupná k prezenčnímu nahlédnutí, že jeden výtisk bakalářské práce bude uložen v příruční knihovně Fakulty aplikované informatiky Univerzity Tomáše Bati ve Zlíně a jeden výtisk bude uložen u vedoucího práce;
- byl/a jsem seznámen/a s tím, že na moji bakalářskou práci se plně vztahuje zákon č. 121/2000 Sb. o právu autorském, o právech souvisejících s právem autorským a o změně některých zákonů (autorský zákon) ve znění pozdějších právních předpisů, zejm. § 35 odst. 3;
- beru na vědomí, že podle § 60 odst. 1 autorského zákona má UTB ve Zlíně právo na uzavření licenční smlouvy o užití školního díla v rozsahu § 12 odst. 4 autorského zákona;
- beru na vědomí, že podle § 60 odst. 2 a 3 autorského zákona mohu užít své dílo – bakalářskou práci nebo poskytnout licenci k jejímu využití jen s předchozím písemným souhlasem Univerzity Tomáše Bati ve Zlíně, která je oprávněna v takovém případě ode mne požadovat přiměřený příspěvek na úhradu nákladů, které byly Univerzitou Tomáše Bati ve Zlíně na vytvoření díla vynaloženy (až do jejich skutečné výše);
- beru na vědomí, že pokud bylo k vypracování bakalářské práce využito softwaru poskytnutého Univerzitou Tomáše Bati ve Zlíně nebo jinými subjekty pouze ke studijním a výzkumným účelům (tedy pouze k nekomerčnímu využití), nelze výsledky bakalářské práce využít ke komerčním účelům;
- beru na vědomí, že pokud je výstupem bakalářské práce jakýkoliv softwarový produkt, považují se za součást práce rovněž i zdrojové kódy, popř. soubory, ze kterých se projekt skládá. Neodevzdání této součásti může být důvodem k neobhájení práce.

Prohlašuji,

- že jsem na bakalářské práci pracoval samostatně a použitou literaturu jsem citoval. V případě publikace výsledků budu uveden jako spoluautor.
- že odevzdaná verze bakalářské práce a verze elektronická nahraná do IS/STAG jsou totožné.

Ve Zlíně

.....
podpis diplomanta

Contents

1	Introduction and Motivation	9
I	THEORETICAL PART	13
2	Environment Map Acquisition	14
2.1	Consumer Camera	14
2.2	Fisheye Lenses	15
2.3	Photographing a Mirrored Sphere	16
3	High Dynamic Range Imaging	18
3.1	HDR Imaging Principe	18
3.2	Tone Mapping	19
4	Lights and Material Representations	22
4.1	Median-Cut	22
4.2	Bidirectional Texture Functions	24
II	PRACTICAL PART	27
5	Capturing and Analysis of HDR Environment Maps	28
5.1	Fisheye lenses	28
5.2	Mirrored Sphere	33
5.3	Environment Map Analysis	36
6	Analysis of Human Visual Perception of Illumination Environments	43
6.1	Psychophysical Experiment	43
6.1.1	Experimental stimuli	43

6.1.2	Participants	44
6.1.3	Experimental procedure	45
6.1.4	Analysis of experimental results	45
6.1.5	Discussion	46
6.2	Validation Experiment	53
6.3	Statistical Prediction	53
7	Conclusions and Future Work	60
	References	64
	List of Notations and Acronyms	65
	List of Figures	66
	List of Tables	71

Chapter 1

Introduction and Motivation

In many industrial sectors, demand is currently increasing for accurate virtual representations of real-world materials. Important application areas include safety simulations and computer-aided design. In the first area, the main concern is choosing the right material to fulfil given safety limits of reflectance, while in the second the aim is to avoid costly and time consuming design cycles of material selection, solid model production and visual evaluation [11].

The six dimensional bidirectional texture function (BTF) introduced by Dana in 1999 [2] is the first real and up to date the only practically useful representation of illumination and view dependent surface textures. The visual quality of material visualization using BTFs is dependent on light used to illuminate the object. Figure 1.1 compares a sphere with 3 different BTF materials (*aluminium*, *corduroy* and *leather*) illuminated by a traditional point-light illumination (A, B, C) to its appearance as illuminated by image-based lighting (IBL) environment map (EM) (D, E, F). The difference of visual quality is obvious. Without IBL (A, B, C), the illumination is harsh and simplistic, and the sphere appears noticeably computer generated, i.e. unrealistic. With IBL (D, E, F), the sphere's level of realism and visual quality are increased, i.e., the shadows, reflections, and shading all exhibit complexities and subtleties that are more realistic and internally consistent. In each, a view of the captured environment appears in the background behind the objects (D, E, F). It is great advantage of IBL, because the objects appear to actually belong within the scenes that are lighting them [4].

There is a number of publications on importance of realistic illumination of materials in computer vision and graphics fields. Unfortunately, most of them study perceptual effects of illumination of non-textured objects, where the smooth material

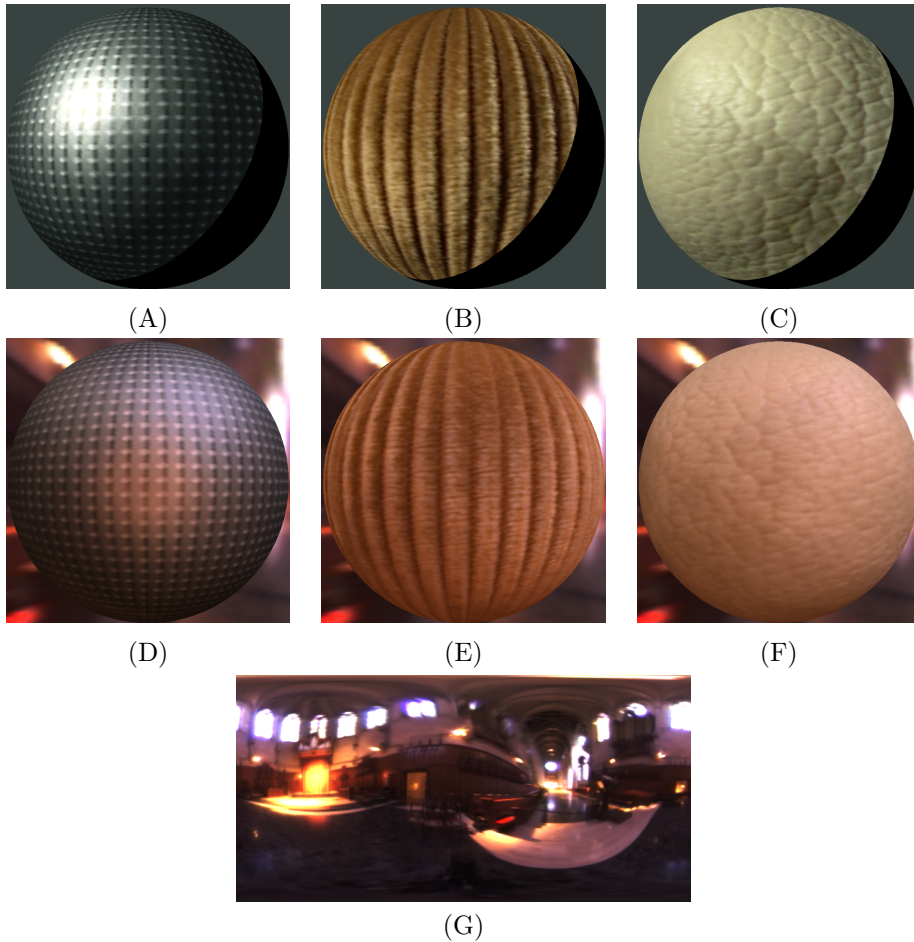


Figure 1.1: Materials illuminated by a point light from left: (A) Aluminium, (B) Corduroy, (C) Leather. The same materials illuminated by the environment illumination (G): (D) Aluminium, (E) Corduroy, (F) Leather. [11].

reflectance is usually represented by a BRDF model.

Fleming et al. [12] shown that peoples' ability to identify the material increases with enhancing of realism of illumination environment. Flemming et al. [13] have demonstrated the importance of real-world illumination for the accurate perception of shape and material properties. Ostrovsky et al. [18] have studied tolerance for illumination inconsistencies. Dror et al. [7] exploited statistical regularities obtained from spatial structure of real-world illumination, which can predict relationships between surface reflectance and certain illumination features. In [25] was shown that changes in surface BRDF and illumination are often confounded, but that adding complex illumination or 3D texture improves the matching. In [23] studied,

among others, impact of illumination map warping on human visual perception of illuminated object.

We are not aware of any research analysing human visual perception of different types of natural environment in combination with realistic material texture measurements represented by BTFs.

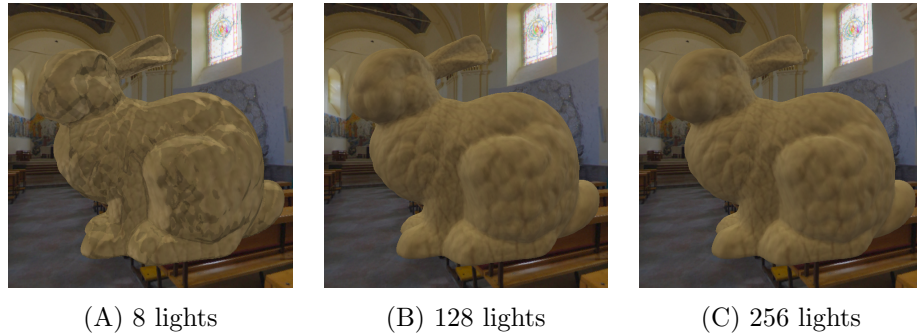


Figure 1.2: Leather bunny illuminated by environment map represented by (A) 8 lights of EM, (B) 128 lights, and (C) 256 lights.

To achieve reasonable visualization speeds the EMs are the most commonly approximated by a finite set of point lights. Number of lights used for material illumination is the most important factor for visual quality when using IBL. Figure 1.2 (A, B, C) compares leather bunny illuminated by IBL from *church* EM for 8, 128 and 256 lights. One can see high visual difference of bunny illuminated by 8 and 128 lights. On the other hand, one can hardly notice any difference between bunny illuminated by 128 and 256 lights. So for the 128 lights visual quality of bunny seems as good as for any higher number of lights.

This is very important because a BTF rendering time can be significantly reduced when one knows appropriate number of lights. This number can be the most reliably estimated by methods of visual psychophysics. Therefore, in this thesis we designed and performed a controlled psychophysical experiment on group of human observers. The relationship of human visual perception on tested BTF materials (*aluminium*, *corduroy* and *leather*) for different number of lights and different EM are tested, analysed, and discussed in this thesis. Outcome of this research is a statistical method that can predict the appropriate number of lights for specific EM in accordance to human visual perception.

The theoretical part of the thesis starts with description of illumination environment map acquisition in Chapter 1. Chapter 2 deals with description of principles

of high-dynamic range imaging and tone-mapping, while Chapter 4 explains representation of environment map by a set of lights using a median-cut algorithm as well as used material appearance representation.

Practical part of the thesis starts with detailed explanation how HDR EMs used for experiment in this thesis were captured. The psychophysical estimation of optimal number of lights and its analysis, validation, and derivation statistical prediction are subject of Chapter 6. Chapter 7 concludes the thesis.

Part I

THEORETICAL PART

Chapter 2

Environment Map Acquisition

This chapter provide general information about the process of environmental map (EM) acquisition. When one wants to capture the incident illumination at a point in space, one must consider taking an image with two following properties.

First an object must be illuminated all directions, because light coming from any direction can affect its appearance. Such an omnidirectional illumination can be represented by an environment surrounding the object. This environment can be captured by a process known as panoramic (omnidirectional) photography. There are several methods of acquiring 360° environment maps. We use the most spread techniques, which include using mirrored spheres, tiled photographs and fish-eye lenses. Each of these methods has certain advantages and disadvantages.

Second, the full dynamic range of the light within the scene must be captured, from the brightest concentrated light sources to the dimmer larger areas of indirect illumination from other surfaces in the scene. This dynamic range should not be affected by used image format with a limited data quantization. HDR photography techniques presented in next chapter satisfy this requirement [4].

2.1 Consumer Camera

Omnidirectional images can be created by taking multitude of photographs capturing different directions and stitching them together. For this purpose can be used a common compact camera together with image stitching software. The most common panoramas are acquired all the way around horizon but only with a limited vertical field of view. To capture lighting is very important to acquire images looking in all directions particularly upward, because much of the light usually comes from

this direction. The quality of stitching depends on stability of camera that can be improved by rotational tripod together with a specific method of taking multitude images [4].

Consumer camera Nikon D-60 was used to capture EM of U13 library in Zlin in the Fig. 2.1. Together 20 images were taken in order to get this EM which does not cover all environment but just a part. Another images will be needed to cover complete EM, which is the main disadvantage of this method. Advantage is that no other special appliances are needed as fisheye objective or mirrored sphere introduced in next chapters. Another advantage is high image quality.

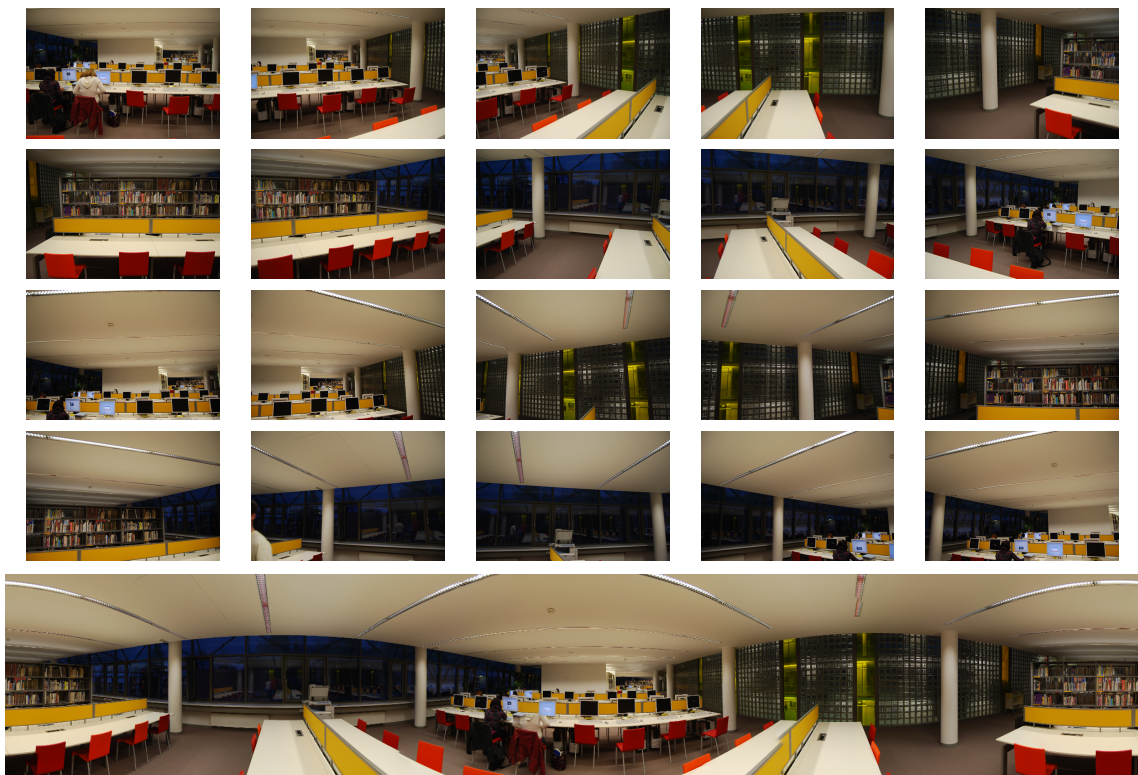


Figure 2.1: EM of U13 library in Zlin consisted of 20 images taken by Nikon D-60 and stitched in Hugin.

2.2 Fisheye Lenses

The power of fish-eye lenses lies within capability capturing 180 °or more of full view of an environment in one shoot 2.2. As a result it is possible to take full view of the environment in as few as two images [4].

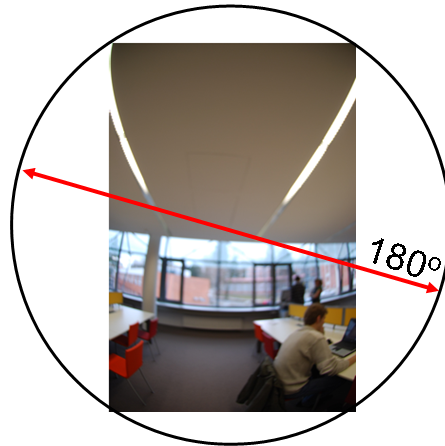


Figure 2.2: Figure shows 180° capturing capability of used fisheye lenses (NIKKOR 10.5mm).

However, the consumer fish-eye lenses, do not capture whole hemisphere of 180° wide field of view as shown in Fig. 2.2. To achieve good quality of stitching it is important to use appropriate method of taking multitude images. The method introduced by Peter Gawthrop together with open source photo stitching software Hugin¹ showed very good results (see Fig. 2.3)[14]. By this technique it is possible to get full EM with 8 pictures using fisheye lens what is big advantage of this method (see Fig. 2.3).

The fisheye lens images are not as sharp as regular photographs, but quality of EM obtained in this way is usually higher than obtained by photographing mirrored spheres and, are sufficient for most light probe applications.

AF DX Fisheye-NIKKOR 10.5mm f/2.8G ED lens were used in this thesis.

2.3 Photographing a Mirrored Sphere

EM can be captured with only two images when photographing a Mirrored sphere (see Fig. 2.4), which is the biggest advantage of this technique. Although one needs two tripods, capturing EM with this technique can be convenient and fast. Mirrored spheres are inexpensive, 15 cm garden chrome ball that was used in this thesis cost just 400 CZK (available from www.homein.cz). Disadvantage of this technique is lower image resolution due to surface imperfection, but in many cases it satisfies

¹<http://hugin.sourceforge.net/>

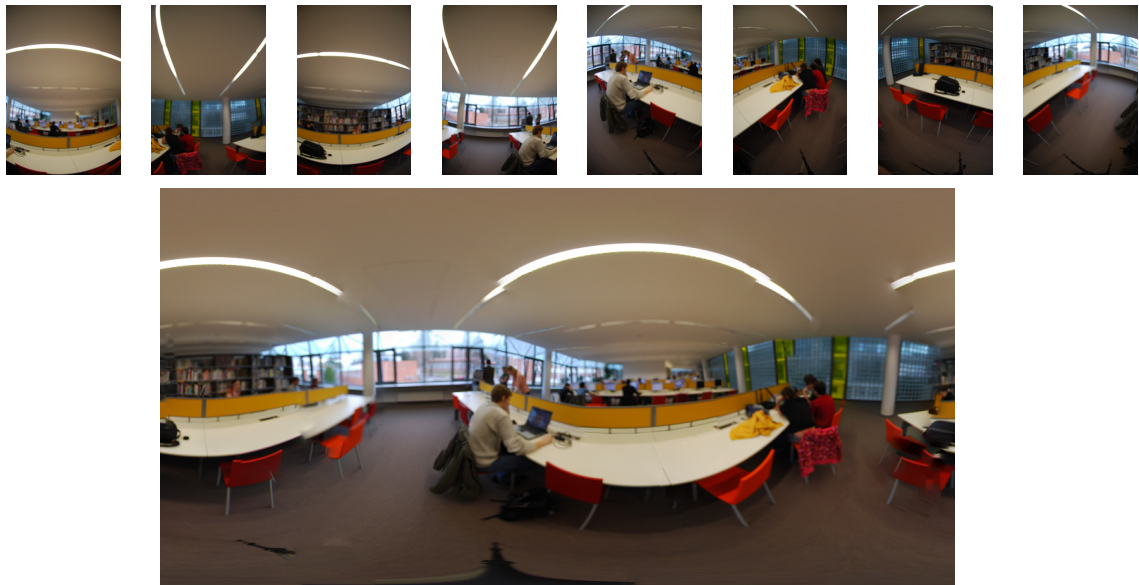


Figure 2.3: EM of U13 library in Zlin consisted of 8 images taken by Nikon D-60 with fisheye lens NIKKOR 10.5mm and stitched in Hugin.

requirements. One can enhance resolution quality by means of using industrial quality mirror sphere and telephoto lenses.



Figure 2.4: EM of U13 library in Zlin consisted of 2 images taken 90 degrees apart by Canon EOS 400D with 300 mm telephoto lenses Sigma, and stitched in HDR Shop and Hugin.

Chapter 3

High Dynamic Range Imaging

In image processing, computer graphics, and photography, high dynamic range imaging is a set of techniques that allow a greater dynamic range of luminances between the lightest and darkest areas of an image than standard provided by digital imaging techniques or photographic methods. This wider dynamic range allows HDR images to more accurately represent the wide range of intensity levels found in real scenes, ranging from direct sunlight to faint starlight [4].

This chapter focus on capturing the high-quality HDR images from real scenes using conventional camera equipment.

3.1 HDR Imaging Principe

Every digital image sensor has inherent limitations which limit the camera capabilities to cover full dynamic range of scene. One has to decide if catches the very dark area of scene and loose information from very bright part, or vice versa. This problem can be solved by taking multiple exposures of the same scene. Usually it is enough to take three low-dynamic range photographs (LDR), i.e. under exposed, nominal and over exposed (see Fig. 3.1), and combine them together to get full HDR image. This method is very fast and can be done with the cheapest consumer camera. One just has to manually set different exposures. Some cameras have auto-bracketing mode that can automatically take several shoots with different camera settings. This function is a key feature for photographing HDR images.

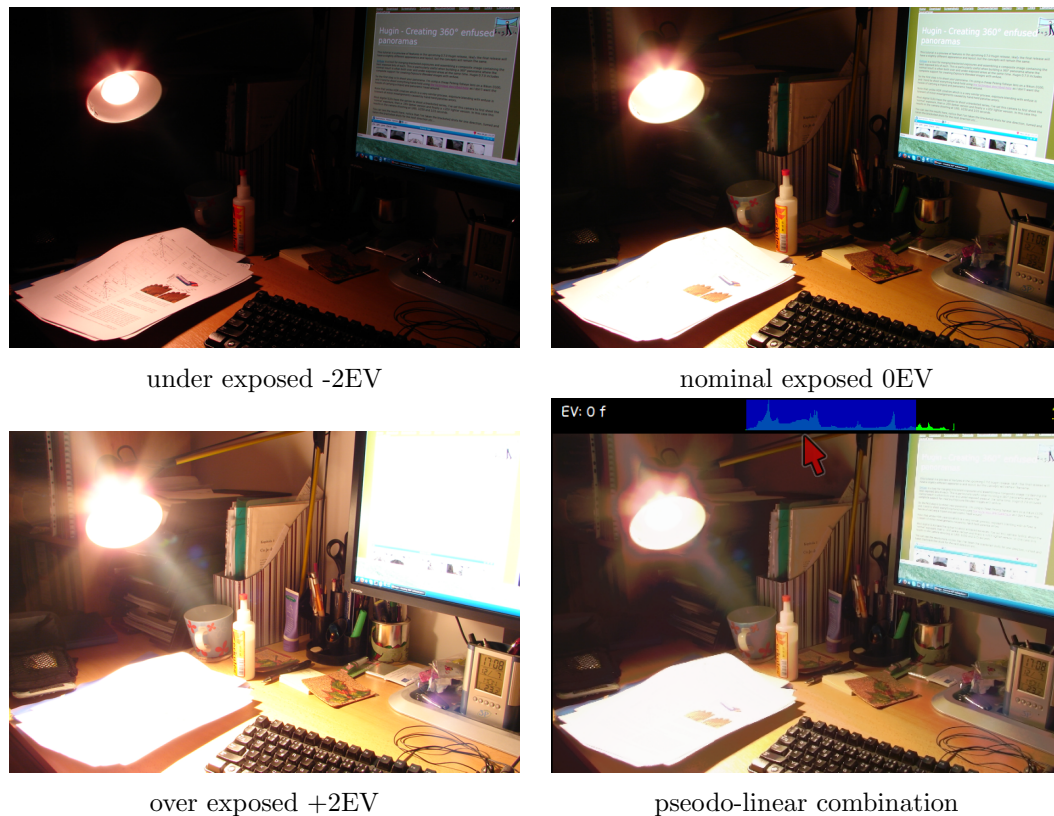


Figure 3.1: The set of three low dynamic range images with different exposure is taken and stitched together using PFSTools (manual can be found at <http://staff.utia.cas.cz/filip/index.html>). The HDR image can be created using pseudo linear combination of these three images. One can interactively change the "exposure" of HDR with the slider to obtain LDR image [10].

3.2 Tone Mapping

The dynamic range of illumination in a real-world scene is high - on the order of 10000 to 1 from highlights to shadows, and higher if light sources are directly visible. Especially if the dark scene is illuminated by an interior light (see Fig. 3.1). Using techniques discussed in the previous chapter is possible to capture full dynamic range of the scene. Unfortunately, most display devices come with moderate absolute output level and a useful dynamic range of less than 256 to 1. The discrepancy between the wide range of illumination that can be captured and the small ranges that can be reproduced by existing display values makes the accurate display of the images of the captured scene difficult. This is the HDR tone-mapping problem. Artists and photographers have been addressing this problem for a long time. The core problem for the artist (canvas), photographer (positive print) and us (display

device) is that the natural light intensity and contrast level can be completely beyond the output level of display medium.

The human visual system deals with a similar problem. Eye's 130 million photo-receptors absorb light, and 1.2 million axons of ganglion cells transmit information from the retina to the brain. Their signal-to-noise ratio of individual channels is about 32 to 1. A significant amount of visual processing arises from the patterns of communication between neurons in the retina [16]. This role in HDR displaying belongs to tone mapping operators which reproduce visual appearance for human visual system [4]. In other words the range of HDR image must be reduced by tone-mapping operators to fit limited display range with realism detail (contrast) visible with human eye (see Fig. 3.3). Generally tone-mapping operators can be split into two main groups, i.e. global and local operators.

Global Operators

Global operator uses non-linear functions based on the luminance and other global variables of the image. A simple example of global tone mapping filter is (see Eq. 3.1).

$$L = \frac{Y}{Y + 1} \quad (3.1)$$

This function will map scene radiance values Y in the domain to a displayable output range (see Eq. 3.2).

$$[0, \infty] \rightarrow [0, 1) \quad (3.2)$$

Global tone mapping is a fast and a very flexible way of obtaining perceptually plausible LDR images [21] (see Fig. 3.2).



(A) simple global filter (see Eq. 3.1)



(B) original image

Figure 3.2: (A) The simple global filter is used to tone map photography of the forest near Prague (see Fig. (B)). As one can see the image is after global tone mapping hardly recognized, contrast is very low.

Local Operators

Another mapping is based on local operators where the parameters of the non-linear function change in each pixel, according to features extracted from the surrounding parameters. In other words, the effect of the algorithm changes in each pixel according to the local features of the image. Those algorithms are more complicated than the global ones, their output can look un-realistic, but they can provide the best performance, since the human vision is mainly sensitive to local contrast [21] (see Fig. 3.3).

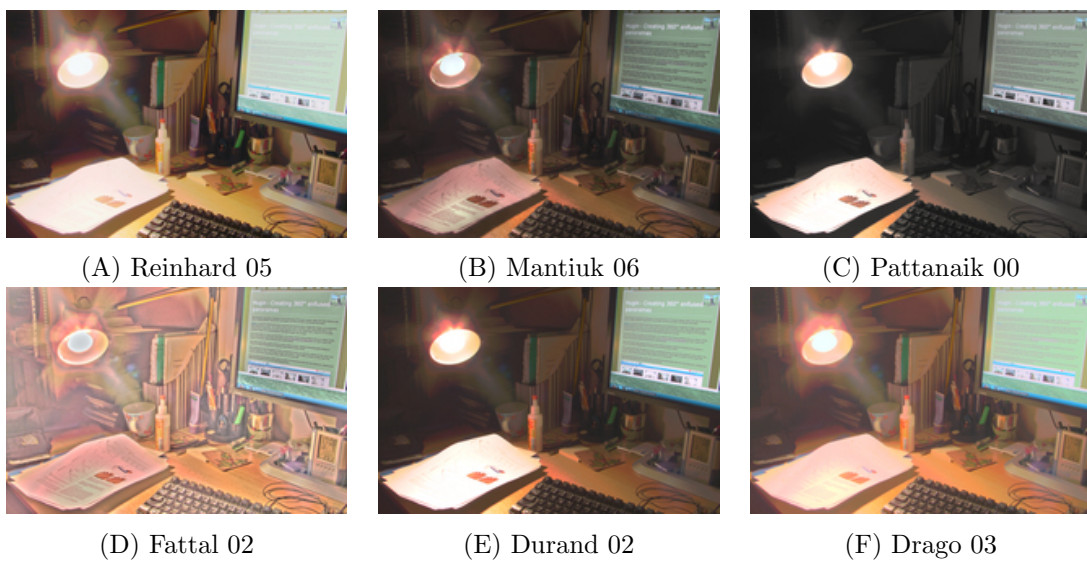


Figure 3.3: Different tone-mapping operators applied on the HDR image (see Fig. 3.1) to scale them into dynamic range of display as shown in [10]: (A) [4] (B) [24] (C) [19] (D) [9] (E) [8] (F) [6]

Chapter 4

Lights and Material Representations

4.1 Median-Cut

The previous chapter described properties and advantages of HDR imagery. This chapter describes how images can be used as a sources of illumination for computer-generated objects and scenes.

HDR images contain information about the shape, colour, and intensity of direct light sources, as well as the colour and distribution of indirect light from surfaces in the rest of the scene. Using suitable rendering algorithms, one can use HDR images to accurately simulate how objects would look if they were illuminated by light from real world. This process is called image-based lighting (IBL) [4].

There is number of approaches for representation of illumination environment maps. They are based either on a set of point lights [3], their clustering [15] or on filtering on different frequency levels [17], [22] with respect to their visibility to achieve faster rendering. The principal idea is to represent illumination environment by a finite set of point lights. Median-cut is the simplest of the algorithms that can do this job very efficiently. The first step, is to choose regions of image from which is rendered object illuminated. The algorithm divides input environment map image into even set of regions by iterative cutting of data pixels at the median point. It is easy to implement, and can realistically represent a complex lighting environment. The pseudo-code is as follows:

1. Add the entire environment map to the region list as a single region.

2. For each region in the list, subdivide it along the longest dimension such that its luminance energy is divided equally.
3. If the number of iterations is less than n , return to step 2.
4. Place a light source at the weighted center or centroid of each region, and set the light source colour to the sum of pixel values within the region (see Fig. 4.1).

Principal disadvantage of this approach is that we can obtain the number of lights that is power of 2 (i.e., 8, 16, 32, ...).

Computing the total light energy is most naturally performed on a luminance Y of the lighting environment rather than the RGB pixel colours. Such a luminance image can be formed as a weighted average (4.1) of the colour channels of the environment map.

$$Y = 0.2125 \cdot R + 0.7154 \cdot G + 0.0721 \cdot B \quad (4.1)$$

While the partitioning decisions are made on the luminance image, the light source colours are computed using the corresponding estimated colour regions in the original RGB image.

The environment map is for rendering wrapped over sphere surrounding the illuminated object. Such a latitude-longitude mapping over-represents regions near the poles. To compensate, the pixels of the environment map image should first be scaled by $\cos \theta$ where θ is the pixels angle of inclination. Determining the longest dimension of a region should also take the over representation into account; this can be accomplished by weighting a regions width by $\cos \theta$ for an inclination θ at center of the region [4].

Every environment map passed through the median-cut algorithm is gamma correction compensated ($\gamma = 2.2$, see Fig. 4.2), prior to lights computation to obtain original linear response of camera sensor.

An example of aluminium sphere illuminated by the different number of lights extracted from the bookshop environment map by median-cut algorithm is shown in Fig. 4.3. The quality of shadows, reflections, and shading is clearly dependent on number of lights. The optimal number of these lights is one of the key issues of this thesis and will be explored in the sets of experiments discussed in the next chapter.

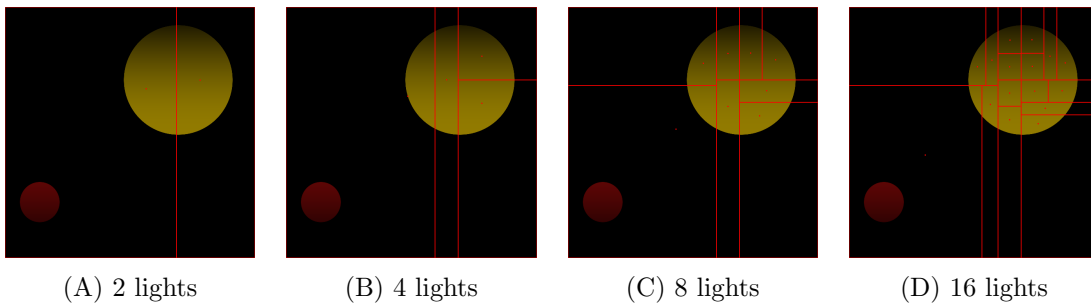


Figure 4.1: Figures (A, B, C, D) show median-cut's performance for the first 4 loops. Every step the regions are cut along the longest dimension such that its light energy is divided evenly. This process is done till the desired number of lights is found. The number of lights that can be found by median-cut is represented by power of two.



Figure 4.2: The same image with 3 different levels of gamma (from the left $\gamma = 1, \gamma = 2, \gamma = 3$) [27].

4.2 Bidirectional Texture Functions

Bidirectional Texture Functions (BTF) are commonly thought to provide the most realistic perceptual experience of materials from rendered images. The first real illumination and view dependent surface texture representation was the Bidirectional Texture Function (BTF), introduced by Dana [11]. A BTF is a six-dimensional function representing the appearance of a material sample surface for variable illumination $\omega_i(\theta_i, \varphi_i)$ and view $\omega_v(\theta_v, \varphi_v)$ directions, resulting in a six-dimensional mono-spectral function $BTF(x, y, \theta_i, \varphi_i, \theta_v, \varphi_v)$, where θ are elevation and φ azimuthal angles. Compared to the four-dimensional BRDF, a BTF depends on two additional parameters, a planar position (x, y) over a material surface (see Fig. 4.4). The BTF preserves effects of masking, shadowing, inter-reflections and sub-surface scattering.

Although material visualization using BTFs provides superb visual quality, even a relatively small sample (e.g., 256×256) occupies several gigabytes of data in its raw form and is not suitable for real-time rendering. There has been much research



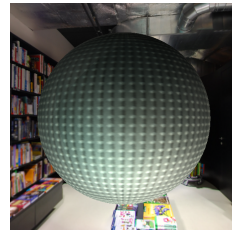
(A) environment map of Archa bookshop in Zlin



(B) 16 lights



(C) 32 lights



(D) 64 lights



(E) 128 lights

Figure 4.3: Figure (A) shows 128 lights found by median-cut algorithm. The aluminium spheres are illuminated by different number of lights from environment map (see Fig. (A, B, C, D)). The quality of shadows, reflections, and shading is clearly dependent on number of lights. All spheres are realistically illuminated as if they were in bookshop, for example the bottom side is more illuminated than top because of indirect light reflected from the floor.

aimed at developing efficient compression techniques that also allow computationally cheap reconstruction and visualization of BTFs (see [11] for a recent summary).

In this thesis we have used measurements from the BTF Database Bonn ¹ as data samples. These data have planar resolution $N \times M = 256 \times 256$ pixels and illumination and viewing directions ($n_i \times n_v = 81 \times 81$) producing uniform sampling of a hemisphere above a material sample (see Figure 4.4). The following BTF samples were used in our experiments (see Figure 1.1): aluminium profile (*alu*), corduroy fabric (*corduroy*) and artificial light leather (*leather*).

Rendering of BTF in ideal illumination environment can be described by the

¹<http://btf.cs.uni-bonn.de/>

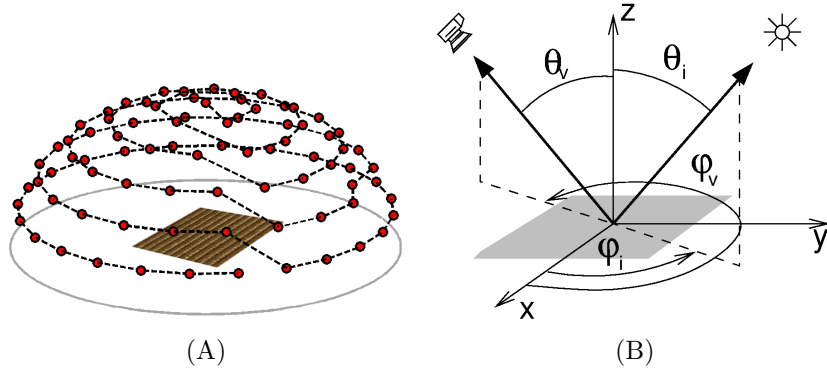


Figure 4.4: (A) Uniform sampling of a hemisphere above a material sample, i.e. 81 sampling directions [11]. (B) Six-dimensional BTF function, representing the appearance of a material sample surface (θ are elevation and φ azimuthal angles, (x, y) planar position over a material surface).

following rendering equation.

$$L_r(x, y, v) = \int_{\Omega_i} BTF(x, y, \omega_i, \omega_v) L_i(x, y, \omega_i) d\omega_i , \quad (4.2)$$

where $L_i(x, y, \omega_i)$ is radiance of individual lights and $L_r(x, y, v)$ is final radiance reflected by the sample, and Ω_i is hemisphere of all possible illumination direction aligned with surface normal .

This function is hard to solve analytically and requires either advanced numerical techniques and simplifications, or stochastic solutions. However, when the environment is represented by a finite set of point lights n_i (as obtained by the median-cut algorithm) the equation can be simplified to equation

$$L_r(x, y, v) = \sum_i^{n_i} BTF(x, y, \omega_i, \omega_v) L_i(x, y, \omega_i) , \quad (4.3)$$

which represent convolution of each BTF pixel with the set of visible lights.

Part II

PRACTICAL PART

Chapter 5

Capturing and Analysis of HDR Environment Maps

In this chapter are step-by-step manuals for acquisition of lighting EM by fisheye lenses and mirrored sphere.

5.1 Fisheye lenses

This section describes the process of panorama composing using camera Nikon D100 and Nikon D60 with AF DX Fisheye-NIKKOR 10.5mm f/2.8G ED lens and cross-platform open source photo stitching software Hugin ¹. By this technique is possible to cover the panoramic sphere with 8 pictures.

The camera is placed on the tripod (Vanguard MK-2). The height of placed camera should be around 130 cm. Two sets of images corresponding to top and bottom camera slant are taken. The top slant is 30° up for azimuthal angles 0°, 90°, 180°, 270°; the bottom slant 30° down for azimuthal angles -45°, 45°, 135°, 225°. The angles can be ticked by coloured insulation tape on the tripod. It is important to take pictures accurately according to method (see Fig. 5.1). If one decides to make HDR panorama one needs to take three different exposures for every position. For this is convenient automatic exposure bracketing embodied in Nikon D100. Exposure can be also changed manually, but it is important to take all pictures from the same position.

Together 24 pictures are taken in order to create HDR panorama with three different exposures. The pictures should be divided into three folders according to

¹<http://hugin.sourceforge.net/>

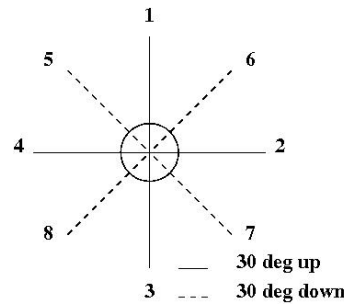


Figure 5.1: Camera orientation. The firm lines show the four directions at 30°, the dashed lines at 30° down. The numbers show the order in which the images are taken [14].

different exposures. Fig. 5.2 shows 8 nominal exposed pictures taken with 0 EV. 8 under exposed pictures (-1.67 EV) are showed in Fig.5.3 and 8 over exposed (1.67 EV) in Fig.5.4.

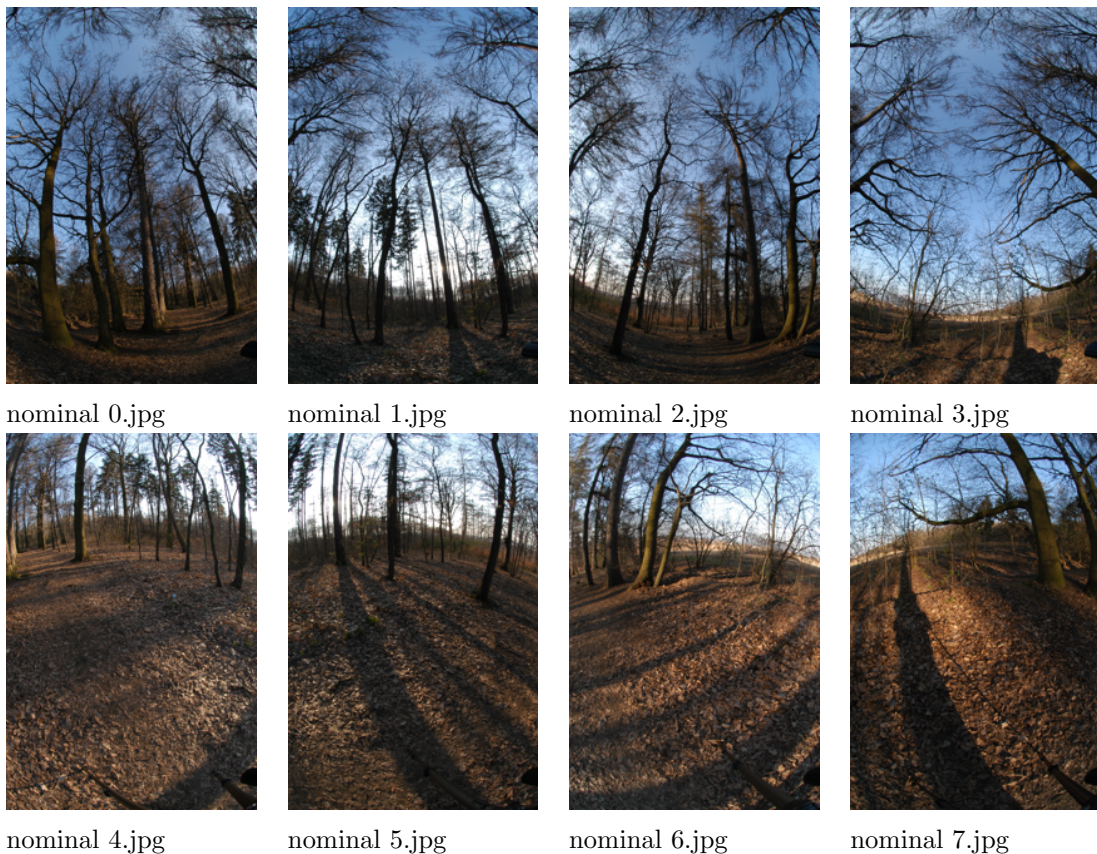


Figure 5.2: Nominal exposure (0 EV) of pictures taken according to Gawthrop method (see Fig.5.1).

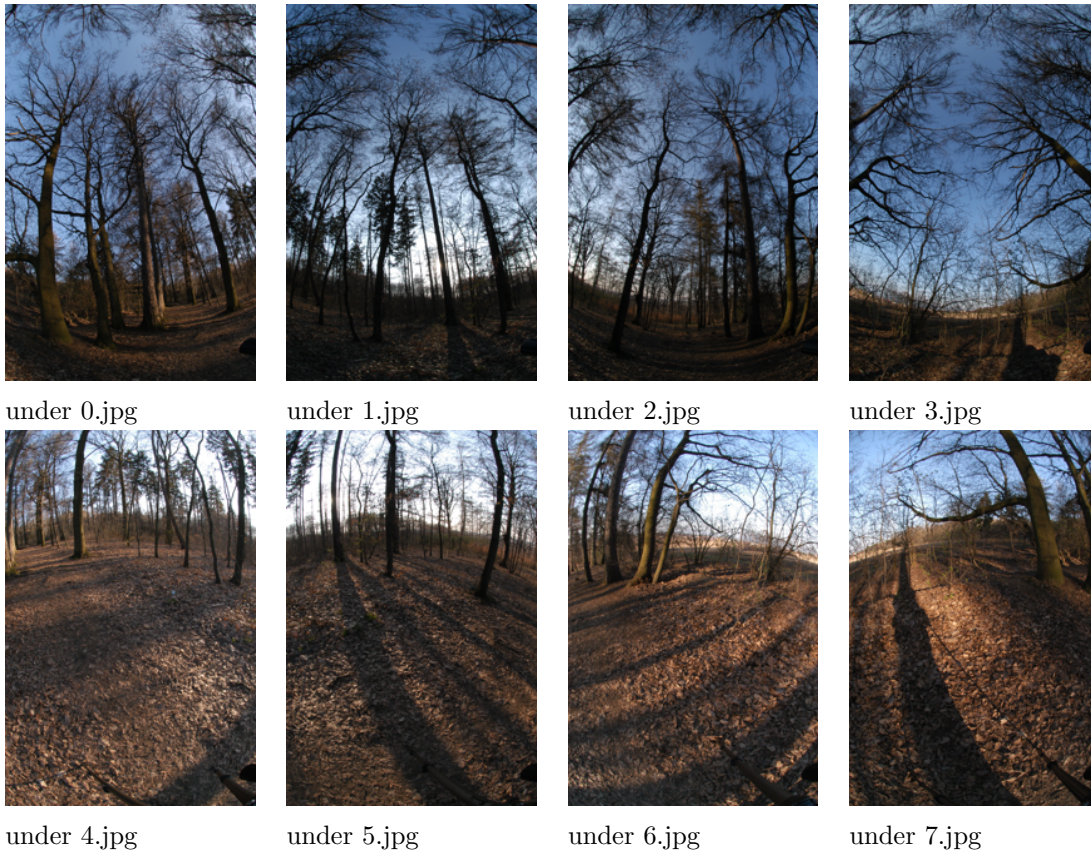


Figure 5.3: Under exposed (0 EV) pictures taken according to Gawthrop method (see Fig.5.1).

Stitching panorama

There are two methods how to connect panoramas. First is manual, second is automatic. Both work very well. For both methods one must tick *Enable rotation search* box which can be found in the *Hugin Preferences* in the tab *Control Points Editor* if one wants auto fine-tune to connect control points correctly. It is because control points between fisheye images almost always have a lot of rotational variation [20].

The automatic panorama stitching:

1. Load 8 pictures with the same exposure in the *Assistant* tab.
2. Choose *Full frame fisheye* Lens type should in the *Assistant* tab. Focal length and focal length multiplier should be loaded automatically. If not enter it manually.

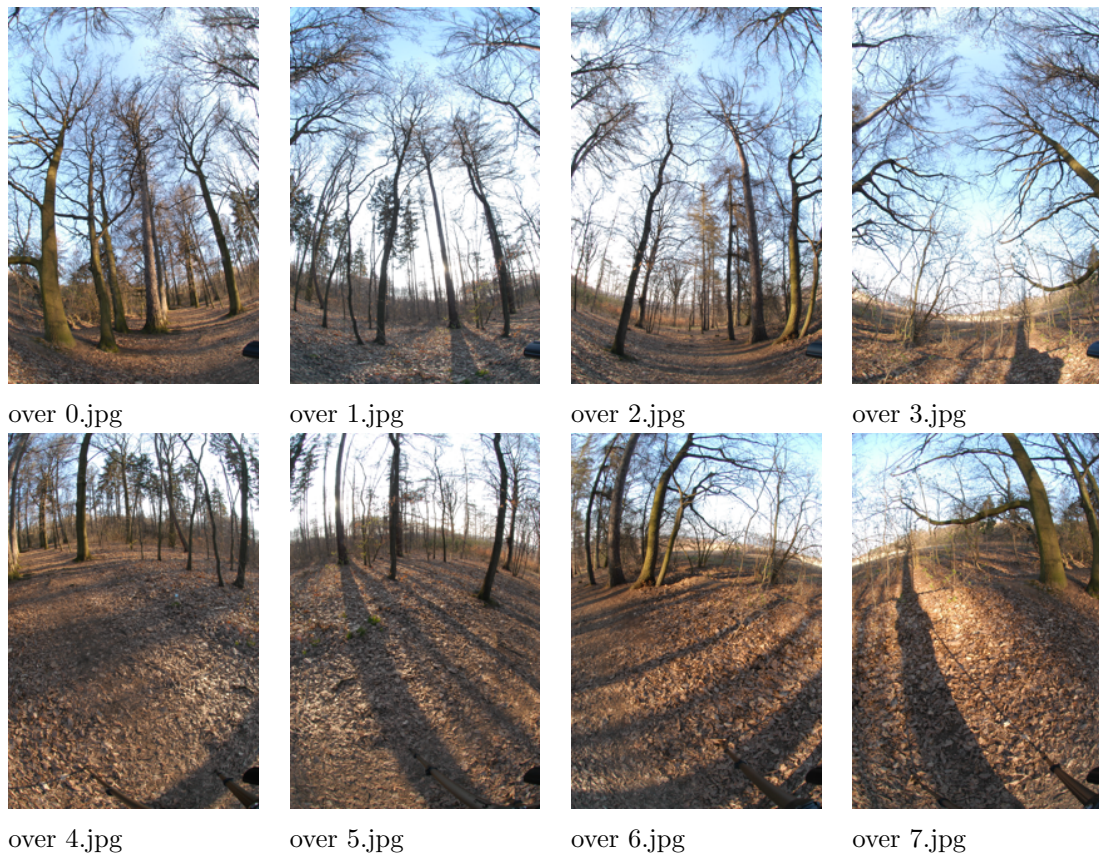


Figure 5.4: Over exposed (0 EV) pictures taken according to Gawthrop method (see Fig.5.1).

3. Button *Align* will automatically connect control points for all pictures. If some pictures stayed unconnected set manual control points in the *Control Points* tab and again align pictures. Check panorama look in *Fast Panorama preview*.

4. In the tab *Stitcher* choose desired projection (in our case Equirectangular), desired panorama canvas size (2048x1024). As a output left default blended panorama. Processing also left default using bi-cubic interpolation. File format can be chosen tiff, jpeg or png based on desired resolution quality of panorama. Stitch panorama using *Stitch Now!*. After that *save project and send to batch*, this will create you *.pto script in which are saved all information about creating this particular panorama as position of control points, type of projection, panorama canvas size etc. It will be used for creating same panoramas for different exposures. If this will be not used one will hardly achieve nice HDR picture due to different control points for every exposure.

So now the panorama and *.pto script are created for the nominal exposure (see Fig. 5.5).

5. Copy the *.pto script to the two other two folders. Create new project and open *.pto script. It will automatically create panorama. Stitch the under and over exposed panoramas. Now one has all three panoramas (see Fig. 5.5).

The manual panorama stitching:

1. Repeat points 1 and 2 from the previous manual.
2. Choose image 0 (for example) as anchor and set the pitch to 30 degrees in the *Images* tab.
3. In the *Control points* tab manually connect pictures as follows:
0 and 4, 0 and 5, 1 and 5, 1 and 6, 2 and 6, 2 and 7, 3 and 7, 3 and 0
4. In the *Optimizer* tab select *Positions(incremental, starting from anchor)* and click *Optimize now!*. Re-optimize using *Positions and barrel distortion*. This should give quite accurate stitching.
5. Repeat points 4 and 5 from the previous manual. More about this method can be found in the [14].

Composing the HDR image

One of the approaches of composing HDR picture is to combine layer masks from different LDR exposures in the Gimp ². In our case, the LDR nominal, under exposed and over exposed panoramas will be composed to one HDR panorama.

1. Load nominal panorama in the Gimp.
2. Bring up under exposed panorama as new layer by *File, Open as Layers*.
3. Add layer masks by *Layer, Mask, Add layer mask*. Choose *Grayscale copy of layer*.

²<http://www.gimp.org/>

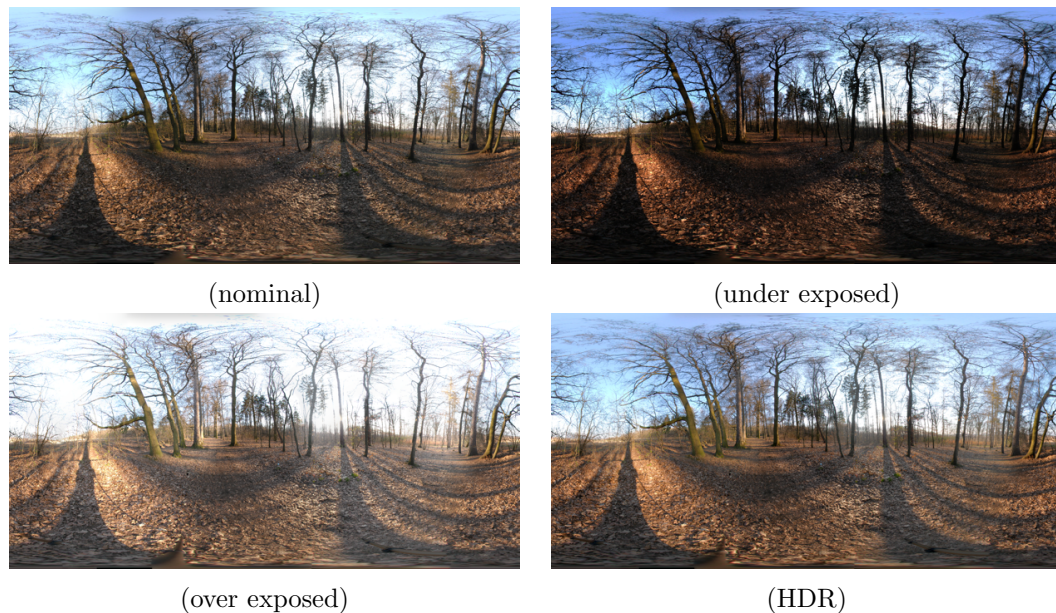


Figure 5.5: Nominal, under exposed, and over exposed EM created from 8 pictures taken in the forest near the Prague according to Gawthrop method (see Fig.5.1). HDR EM was created by stitching them together in the Gimp.

4. Repeat 2 and 3 with over exposed panorama expect that the *Invert Mask* box is ticked in the mask menu.
5. Save the final version as a jpeg.

The final HDR panorama is showed in the Fig. 5.5. One can use different approaches if is not satisfied with this method as interface for generating HDR pictures Qtpfsgui [14] or an open-source set of command line programs PFStools [10].

5.2 Mirrored Sphere

This section describes the process of acquiring light probe images using camera Canon EOS 400D with 300 mm teleobjective Sigma, mirrored ball and Hugin. By this technique it is possible to cover the panoramic sphere with 2 pictures.

The camera is placed on the one tripod and mirrored sphere on the other with height around 130 cm.

1. Two pictures are taken 90 °apart.

2. The mirrored spheres are cutted away from the original background in some image manipulation program (Gimp) and the black background is putted behind the image (see first row of Fig. 5.6).
3. The prepared images are loaded in the HDR Shop ³ (HDR Image Processing and Manipulation software) where are transformed to latitude-longitude mappings using *Image, Panorama, Panoramic Transformations menu* where is chosen *Latitude/Longitude Format of Destination Image*.
4. The final panorama is automatically stitched from the transformed images in Hugin.

One should consider several factors when using this technique described bellow:

Framing and Focus: To have the sphere relatively large in the frame, it is necessary to use a long-focal-length lens [4]. In this thesis are used telelenses Sigma 300 mm. It also minimizes the reflection of the photographer and camera in the sphere.

Calibrating Sphere Reflectivity: The mirrored spheres are generally not optically perfect reflectors. They typically reflect a bit more than half of the light hitting them. Therefore one needs to know the sphere reflectivity which can be done using a setup such as that shown in Fig. 5.6. One needs to compute the average RGB values of each pixel of selected area in direct view and also in mirrored sphere where is reflection of selected area. After that average of mirrored sphere is divided by the average of direct view to obtain the sphere's percentual reflectivity in each of red, green and blue channels. In Fig. 5.6. case result is (0.8181, 0.7689, 0.6361). The light probe image can be corrected by dividing its channels by each of these RGB values [4] (see Fig. 5.6).

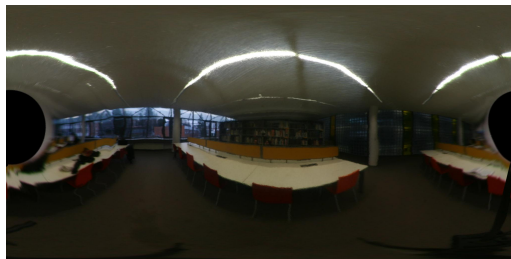
³<http://projects.ict.usc.edu/graphics/HDRShop/>



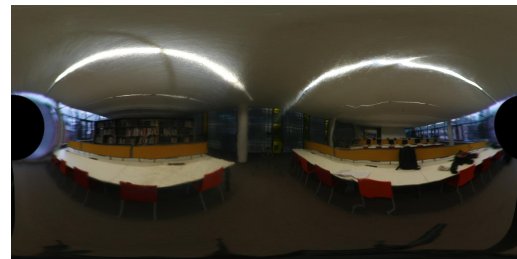
acquired EM



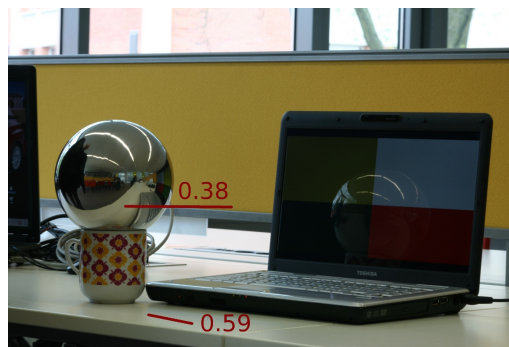
acquired EM taken 90 °apart



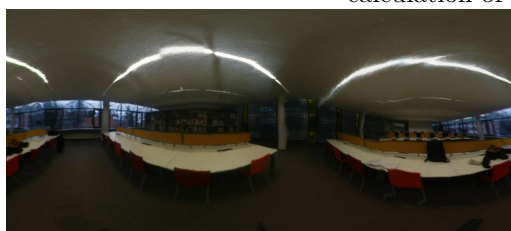
latitute-longitude map in the HDR shop



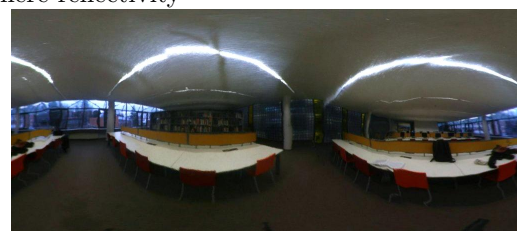
latitute-longitude map in the HDR shop



calculation of sphere reflectivity



before calibration of sphere reflectivity



after calibration of sphere reflectivity

Figure 5.6: Process of acquiring EM with a mirrored sphere. The calibration of sphere reflectivity must be done due to imprecise surface of mirrored ball. In this case, the sphere reflectivity is 64 % $((0.38/0.59)*100)$.

5.3 Environment Map Analysis

Ten different environment maps (EM) were acquired with the Nikon D-60 and Nikon D-100 using AF DX Fisheye-NIKKOR 10.5mm f/2.8G ED lens and method described in the Chapter 5. Environment maps 01, 02, 03, 04, 05, 07, 09, and 10 were captured as LDR images, while maps 06 and 08 were obtained by a tone-mapping (Reinhard 05) [4] of HDR images.

The light is naturally unequally distributed over the EM, what creates unique light for every part of EM. Due to this condition it is appropriate to perform EM analysis in RGB Luminance values (see Eq. 5.1).

$$L(RGB) = 0.3R + 0.59G + 0.11B \quad (5.1)$$

Every EM is analysed as follow:

1. Sum of vertical luminance (sum of all luminances in every pixel in the same row), represented by cartesian graph which shows the vertical direction of luminance intensity (ranging from -90° to 90°).
2. Sum of horizontal luminance (sum of all luminances in every pixel in the same column), represented by polar graph which shows the horizontal direction of luminance intensity (ranging from 0° to 360°).
3. According to these graphs is chosen the angle with the highest horizontal luminance (showed in graph by red dot and angle value). This angle was used as azimuthal viewing angle for illuminated object rendering in the Section 6.2.2.
4. Some EMs have very diverse light conditions. To analyse dependence of human perception on these conditions the rendered object has to be also illuminated from direction of the lowest azimuthal luminance. To make easier to find difference between lowest and highest luminance two blue circles were drawn in polar graph, one touching the highest luminance and other lowest. The lowest luminance is represented by green dot and angle value. This is the case of EM no.03, 04, 06 ($2\times$), and 08).

The EMs analysis results are shown in the Fig. 5.7, 5.8, 5.9, 5.10, 5.11. Fig. 5.12 shows BTF materials (*alu*, *corduroy*, and *leather*) illuminated by 256 lights from these EMs.

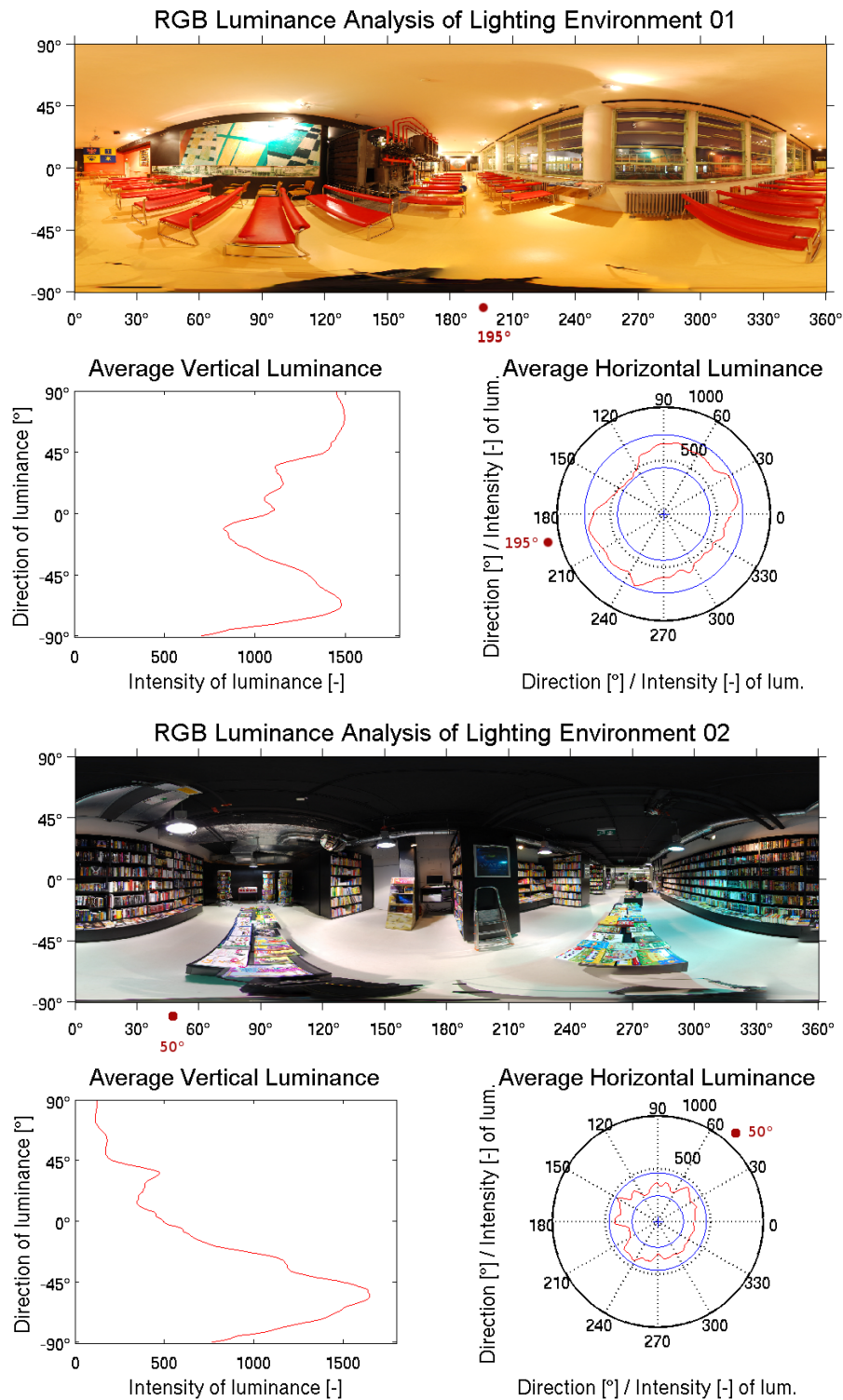


Figure 5.7: Lighting environment 01 is a ground floor of Tomas Bata skyscraper 21. Indoor light is almost uniformly distributed and it has almost highest luminance from all lighting EM (with highest horizontal luminance at 195°). Lighting environment 02 is Archa bookshop. It has specific vertical light distribution, i.e. the brightest part is floor, because of reflected light from the top (with highest horizontal luminance at 50°).

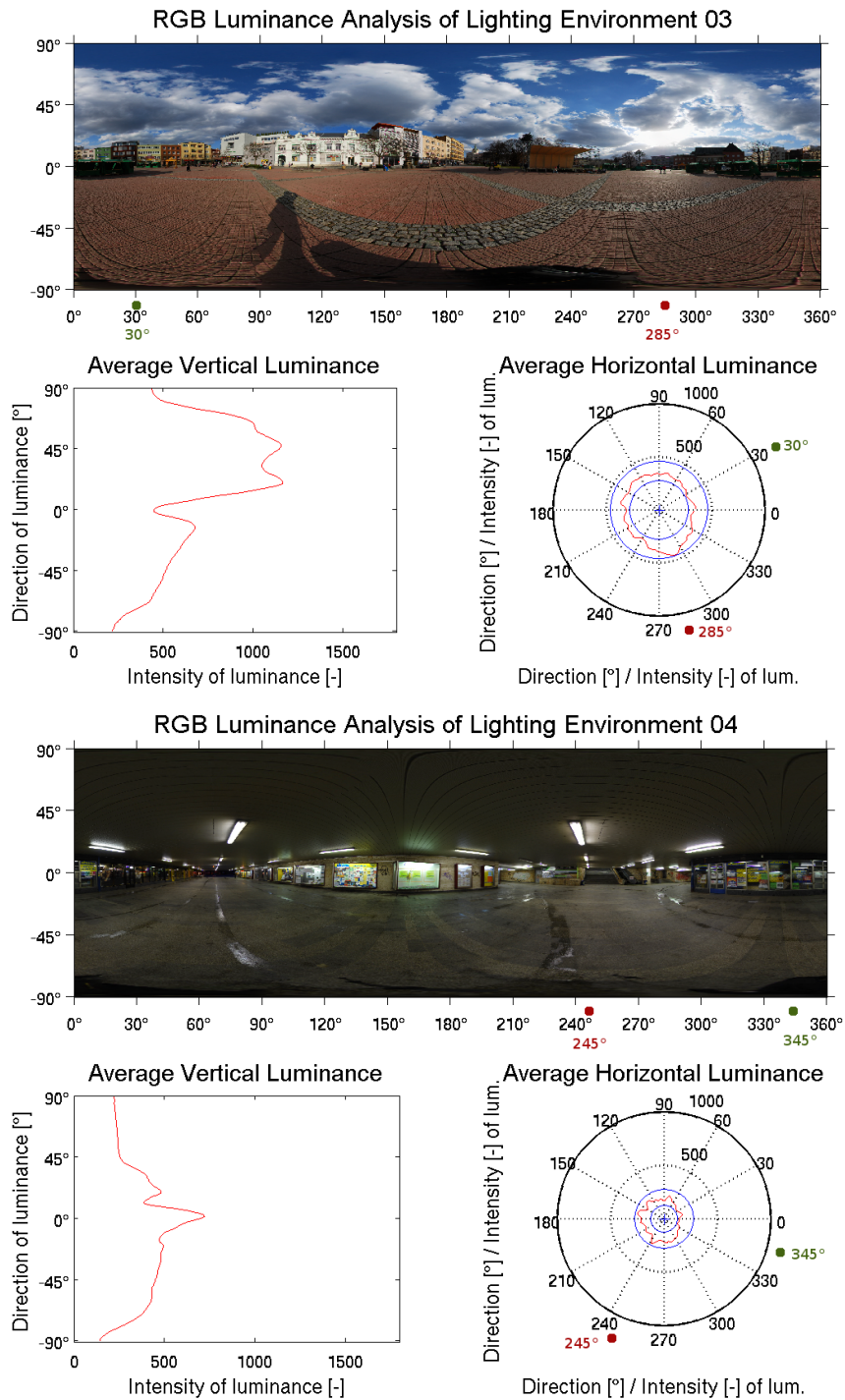


Figure 5.8: Lighting environment 03 is Zlin city centre during bright day. Because of sun hidden behind clouds is naturally highest horizontal luminance at 285°. On the other hand the rest of the EM has much smaller horizontal luminance (see angle 30°). Due to this are chosen two different angles to be used for rendering in next chapter. Lighting environment 04 is Zlin’s subway during the night. The luminance is more than half smaller compared with day light. Again due to non uniform light distribution are chosen two angles for rendering (brightest at 245° and darkest at 345°).

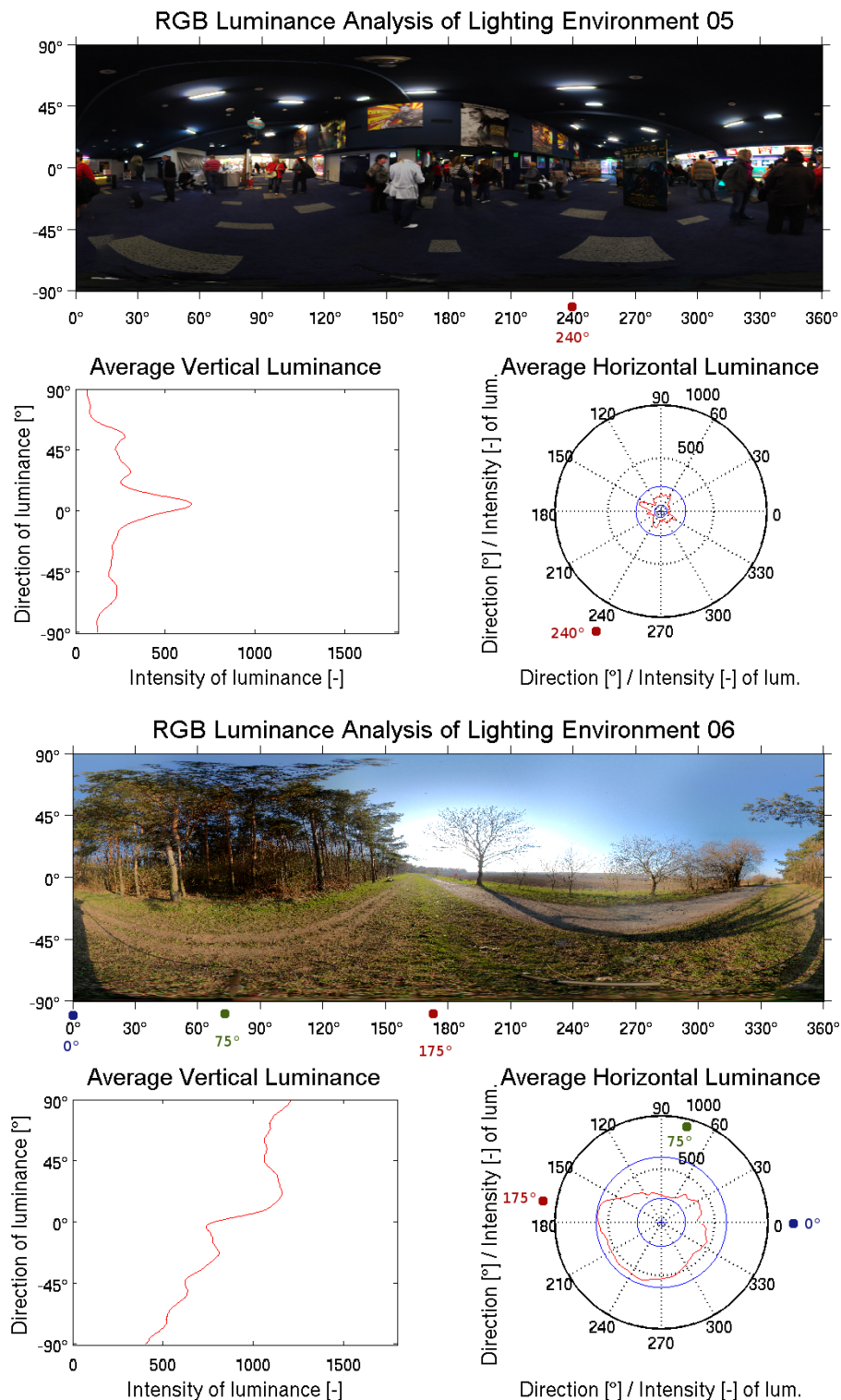


Figure 5.9: Lighting environment 05 is a Golden Apple cinema entrance. The luminance is one of the lowest from all EMs, because light is very dim with the highest horizontal luminance at 240°. Lighting environment 06 is very interesting due to three different horizontal luminance regions. One is when the sun is at 175°, naturally most brightest. The medium luminance is spread around 0° and the lowest is at 75°.

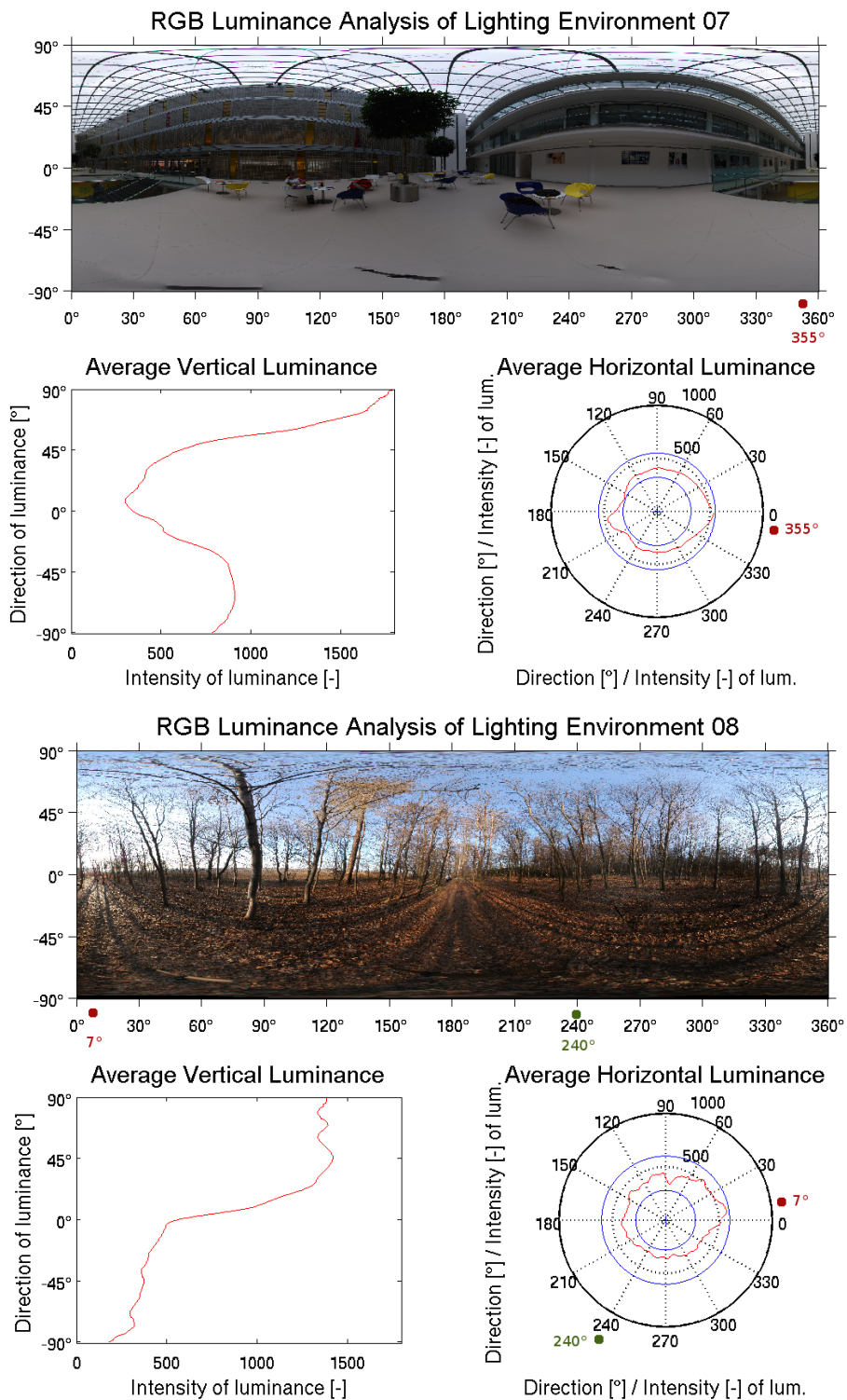


Figure 5.10: Lighting environment 07 is an big open space of Tomas Bata library where natural light comes inside through the glassy roof, so highest vertical luminance is at the top, i.e. 90°. The highest horizontal luminance is at 355°. Lighting environment 08 is a Dáblický háj near Prague, where sun is just coming down at horizontal angle 7°. At angle 240° is the lowest horizontal luminance.

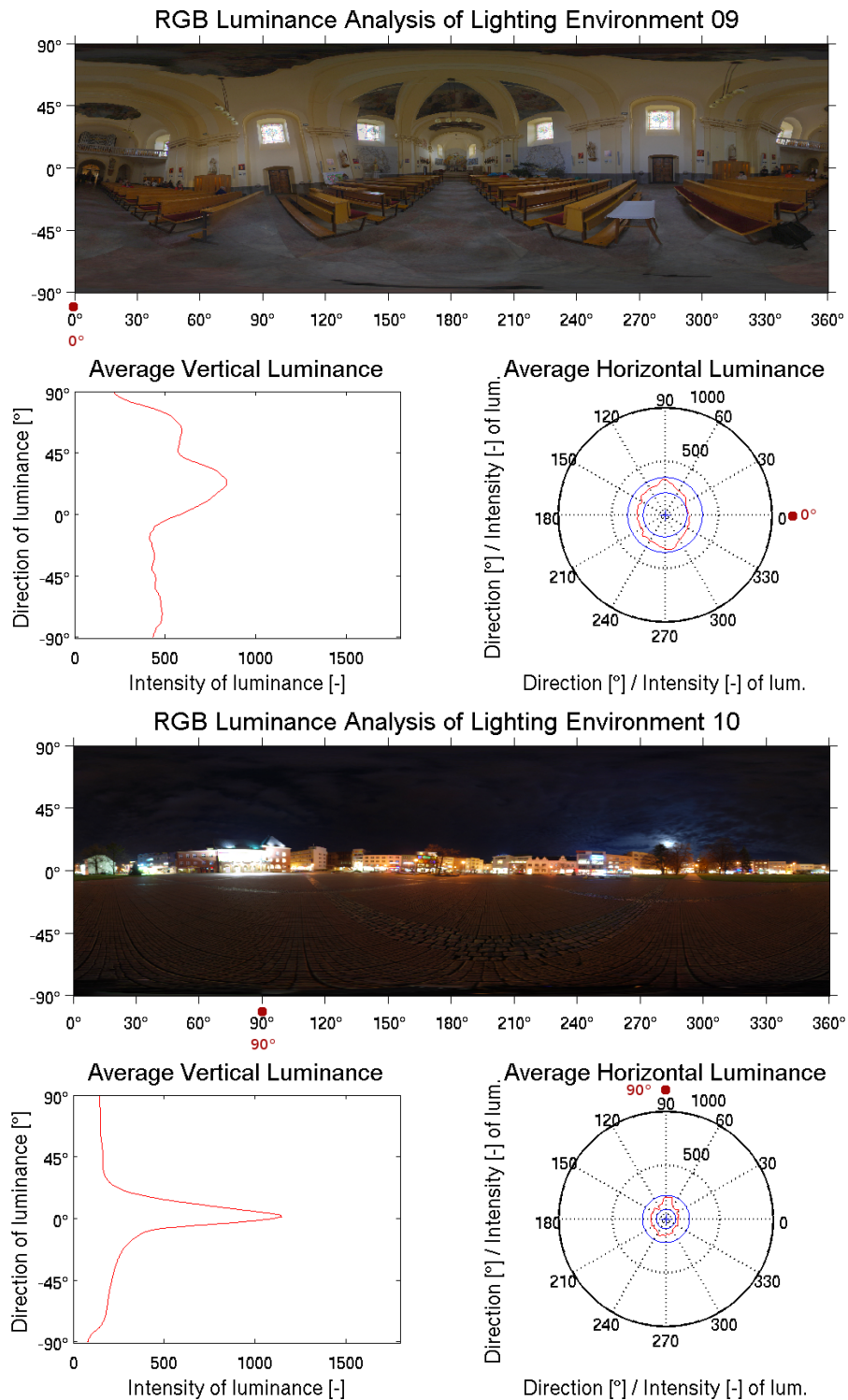


Figure 5.11: Lighting environment 09 is a interior of the church in Zlin with very unusual light conditions that can one find just in church. The lowest horizontal luminance is at 0°. Lighting environment 10 is a Zlin city centre during night, so the light is coming from the moon, stars and surrounding buildings. It is the dimmest lighting EM from all tested. The highest horizontal luminance comes from over-saturated town-hall at 90°.

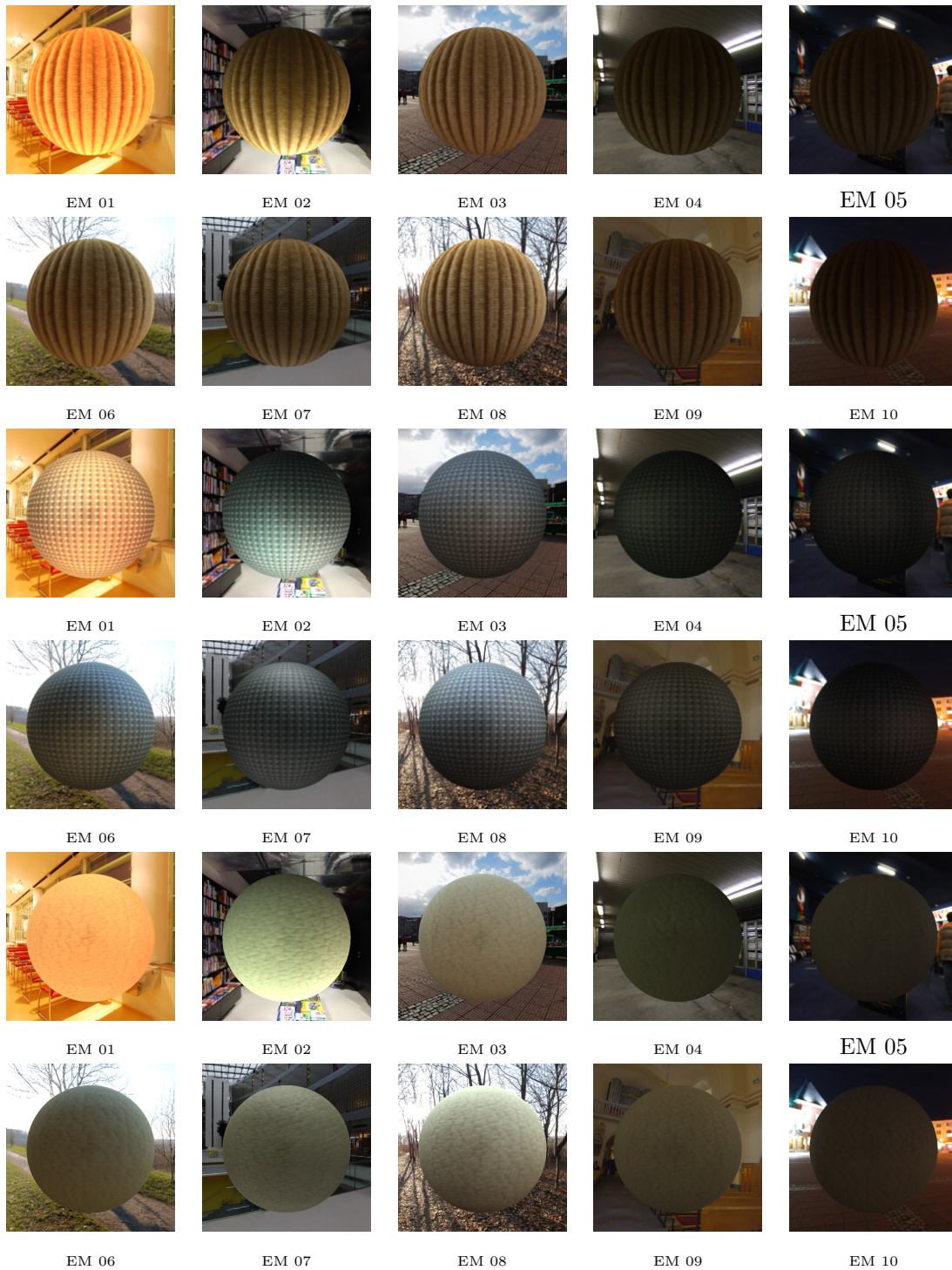


Figure 5.12: BTF materials (*alu*, *corduroy*, and *leather*) illuminated by 256 lights from the real EMs (shown in the Fig. 5.7, 5.8, 5.9, 5.10, 5.11). These rendered images were used as stimuli for the experiment presented in the next chapter.

Chapter 6

Analysis of Human Visual Perception of Illumination Environments

6.1 Psychophysical Experiment

Goal of the experiment was to identify thresholds for the different combinations of tested BTF materials (*alu*, *corduroy*, *leather*) and different illumination environment maps.

6.1.1 Experimental stimuli

Experimental stimuli were used pairs of static images with size of one 640×640 (so the size of pair was 1280×640), each representing a material BTF rendered on a 3D sphere (see Fig. 6.4). We chose to use the sphere shape because its geometry provides a wide range of illumination and viewing combinations without introducing unwanted disturbing effects of higher curvatures. Due to this, the material visual degradation artefacts caused by insufficient environment map representation can be identified much easier and faster on the sphere than on the other more complicated objects. In this work we have not required real-time rendering performance, as we were using static precomputed BTF stimuli, and hence we used original uncompressed BTF data for the rendering to achieve a maximal possible visual quality of our stimuli.

Pairs of images were displayed simultaneously, side-by-side (see Fig. 6.1). Each

pair consisted of one of tested BTF material (*alu*, *corduroy*, *leather*) rendered on the sphere and illuminated by the lighting EM which appropriate part was showed as a background of the stimuli. On the one side of the pair was BTF material illuminated by EM represented by means of: 16, 32, 64, 128, or 256 lights from tested lighting environment, and on the other side was always the same BTF material illuminated by 256 lights by the same lighting EM. Of course their position was randomly changed. So one gets 5 stimuli for every combination of lighting EM and BTF material (16 lights compared with 256 lights, 32 lights compared with 256 lights, 64 lights compared with 256 lights, 128 lights compared with 256 lights and 256 lights compared with 256 lights) what together makes (5 stimuli \times (10 lighting EMs + 5 additional views) \times 3 BTF materials = 225 stimuli).

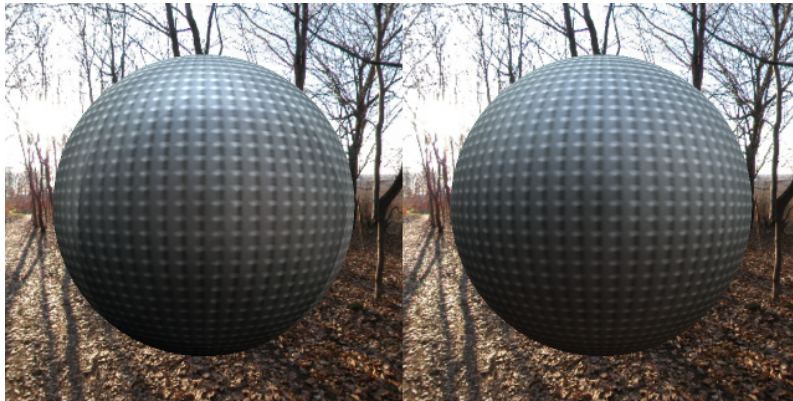


Figure 6.1: Example of stimuli used in experiment. On the left side is showed alu sphere illuminated by 32 lights, which is compared with identical sphere illuminated by 256 lights from the same EM which is also used as a background of the stimuli.

6.1.2 Participants

Twenty participants (8 females, 12 males) voluntary participated in experiment, mostly students under 25 years old with exception of 3 observers who were under 50 years old. All participants had normal or corrected to normal vision. All were naive with respect to the purpose and design of the experiment.

6.1.3 Experimental procedure

Altogether 225 stimuli were shown to the participants in a random order and asked a yes-no question: *Can you spot any differences in the material covering the two objects?* Note that this question is a strict test of the ability to detect a difference and allows participants to make a yes response on the basis of only a minor local difference [11]. Therefore, participants were told to try to answer till 3 seconds, because one will spot eventually any differences if observes the stimulus image too long and compare every pixel. Anyway stimuli were shown as long as they were deciding. However, it turned out around 94 % of the responses were made around 2 seconds. The participants undertook an experiment at once without any break what took usually 8 - 10 minutes. All stimuli were presented on a monitor 14" diagonal wide-screen TruBrite TFT LCD display of notebook Toshiba Satellite L-515 S4925 (60 Hz, 1366 x 768 (16:9), brightness was set to 100 %). The experiment was performed under dim room with one 75W light-bulb during the night. Participants viewed the screen at a distance of 0.5 m, so that stimuli image subtended approximately 28° of visual angle.

6.1.4 Analysis of experimental results

When participants reported a difference between the rendered images their response was assigned a value of 1, and otherwise 0. By averaging the responses of all participants, we obtained psychometric data relating the average response to the different number of lights, EMs and BTF materials [11]. The psychophysical data obtained can be represented by the psychometric function $\psi(x)$ which specifies the relationship between the underlying probability ψ of positive response and the stimulus intensity x

$$\psi(x; \alpha, \beta, \gamma) = \gamma + (1 - \gamma - \lambda)F(x; \alpha, \beta), \quad (6.1)$$

where F is data fitting function with parameters α and β , γ specifies guess rate (i.e., response to zero stimulus), and λ miss rate (i.e., incorrect response for large stimulus). The functions were fitted to the measured data using the `psignifit`¹ tool. As a fitting data was used Weibull cumulative distribution, which is most commonly used in life data analyses due to its flexibility. It is described as

¹www.bootstrap-software.org/psignifit/

$$F(x; \alpha, \beta) = 1 - \exp \left[- \left(\frac{x}{\alpha} \right)^\beta \right] \quad (6.2)$$

for $x \geq 0$, where $\beta > 0$ is the shape parameter and $\alpha > 0$ is the scale parameter of the distribution.

The resulting psychometric functions averaged for all 3 tested BTF materials illuminated by 15 lighting EMs are shown in Fig. 6.2. For example degradation of BTF material illuminated by lighting environment 01 is hardest to identify on corduroy, less on alu and easiest on leather. To exploit psychophysical results, presented in Fig. 6.2, we need to determine the greatest level of degradation of individual BTF samples that is not detectable by the average human observer, so called threshold ε at which a difference between rendered images is detected by only 25 % of participants. Note that due to the applied context of our work we use a more restricting value than 50 %, which is a standard in psychophysical research. These thresholds are computed from the fitted psychophysical functions [26] for different illuminations 6.2 and objects using the equation

$$\varepsilon^{p=0.25} = \alpha \sqrt[\beta]{\ln \left(\frac{1 - \gamma - \lambda}{1 - 0.25 - \lambda} \right)}, \quad (6.3)$$

where α, β are estimated parameters of the Weibull distribution and γ and λ are guess and miss rates estimated during the fitting of (see Eq. 6.2) [11].

All psychophysical data obtained by experiment represented by the psychometric functions are showed in Fig. 6.2. The thresholds obtained for 25 perceptual difference probability level of all EM are showed both in numerical and graphical format in Table 6.1, which also contains the average response time for all EMs.

6.1.5 Discussion

Experiment showed that *corduroy* needs the least number of lights from all 3 tested materials (*alu*, *corduroy*, *leather*) to achieve sufficient visual quality without any visual degradation. It must be illuminated by 64 lights for mostly all types of EMs. For dark EMs (EM 04, 05, 10) even 50 lights is enough. The exceptions are EMs 06, 07 and 08 because of strong luminance caused by direct sun. In this case the corduroy must be illuminated by at least 170 lights.

The *alu* needs 128 lights to reach the same visual quality for mostly all types of

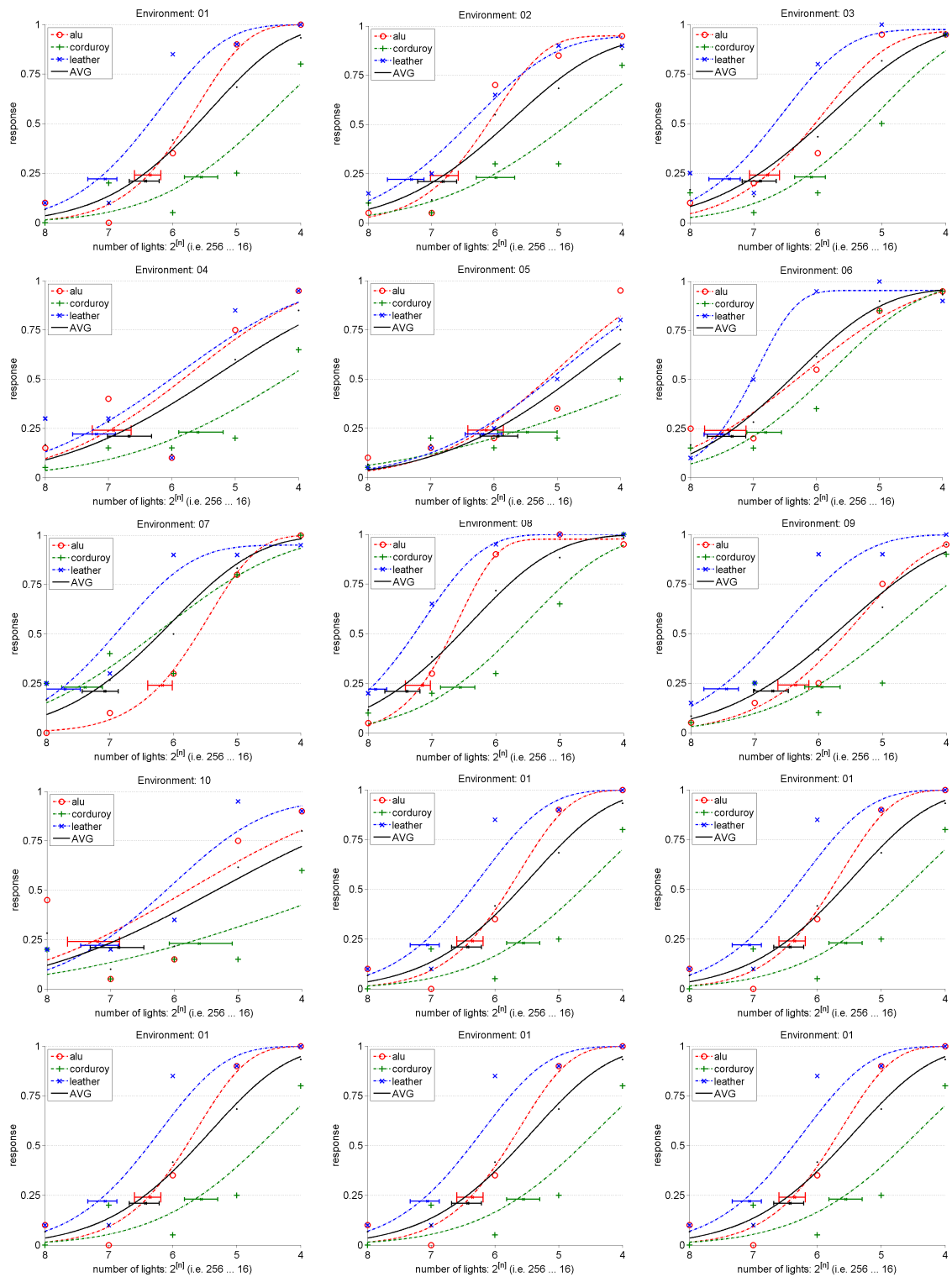


Figure 6.2: Psychometric function for 25 % threshold of psychophysical data from experiment.

EM	alu	cor	lea	AVG
1	82	47	133	86
2	107	64	160	113
3	110	69	167	119
4	121	48	146	102
5	70	44	73	61
6	166	112	182	162
7	72	168	209	135
8	142	94	236	168
9	82	62	174	105
10	147	49	139	117
3a	99	85	118	110
4a	108	38	64	73
6a	108	58	128	97
6b	104	138	115	126
8a	67	79	215	107

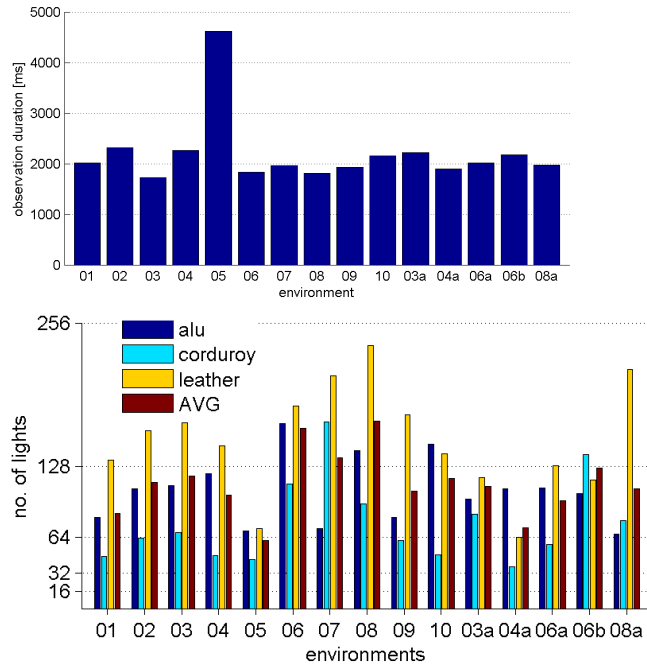


Table 6.1: Thresholds of sufficient number of lights for all tested materials and EMs obtained from the experiment. Figure also shows the AVG of observer’s time response dependent on EM.

EMs. Exception are EMs 06 and 08 with direct sun which require 150 lights. The exception is EM 10 which is town centre during night. For this EM *alu* requires 150 lights as for the EM with a direct sun. This can be caused by a strong concentration of lights (see Fig. 5.11) in the centre of EM which could enhance the conditions for observers to find visual degradation of *alu*. In other dark EMs 04 and 05 (see Fig. 5.8) lights are not centred in any area, but they are mostly uniformly distributed through the EM which could create the worse conditions for observers to find visual degradation, hence more lights are needed.

The *leather* needs the most lights from all tested materials to achieve sufficient visual quality. It keeps desired visual quality for any EM when it is illuminated by 240 lights. The reason for this may be the light colour of the sample, which visually amplifies any degradation.

Figure 6.3 (B) shows the average (AVG) estimated number of sufficient lights for all 3 materials computed over all tested EMs. As one can see, generally for any EMs *corduroy* requires the least number of lights to acquire sufficient visual quality.

Figure 6.3 (A) shows observers time response dependent on number of lights and material. Average time response of observers was about 2 seconds for all EMs with exception 5 seconds for the darkest EM 05 (see Fig. 5.9), where they were longer

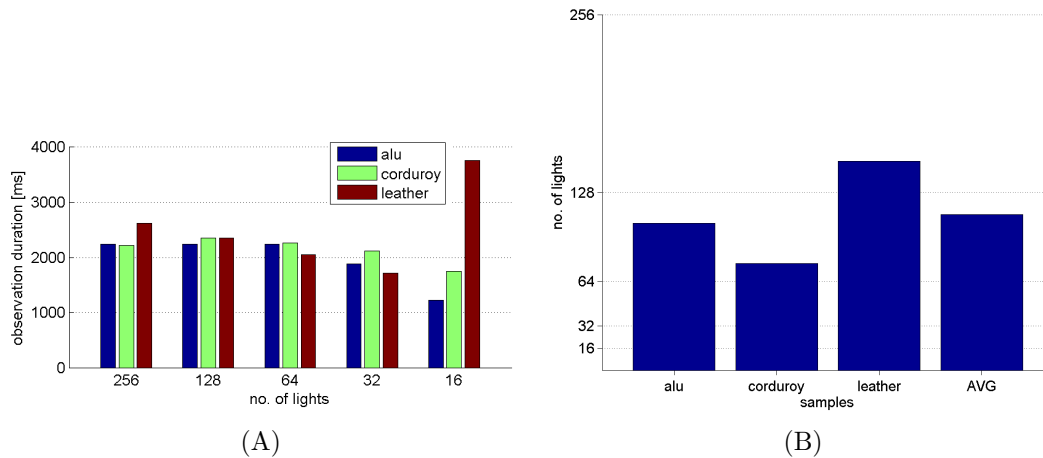


Figure 6.3: (A) Observation duration for all tested materials. (B) Sufficient number of lights for all EMs.

trying to find visual degradations, but after this time has passed away they mostly did not find visual degradations for more than 70 lights (see Table 6.1).

Dependent on material time response was about 2.5 seconds with exception of *leather* illuminated by 16 lights which took almost 5 seconds. This anomaly could be caused by huge visual degradations that could be really easily spotted by observers which could take away the challenge of observing, and which could be even attractive enough to look at for a while (see Fig. 6.4).

Note that human observer fastest reaction on any visual stimuli can range from 150 to 225 ms. It is natural that reaction time is quickest for young adults and gradually slows down with age [5]. This means that in our experiment observers were analysing and processing stimuli image mostly between 1.5 - 2 seconds (Tab. 6.1) prior to their decision.

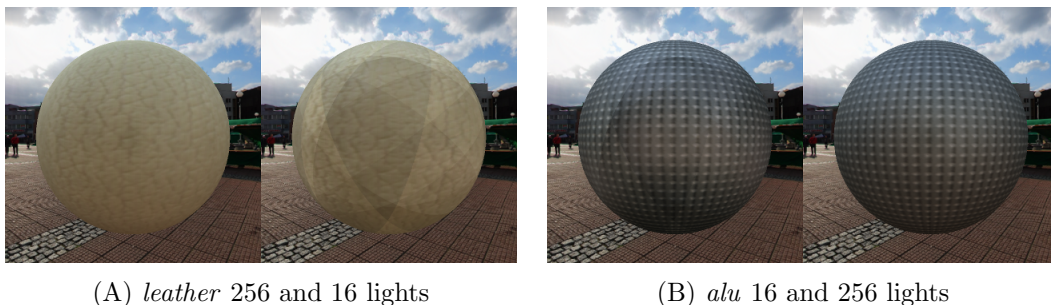


Figure 6.4: Figures compares visual quality of two different BTF materials. Notice that *alu* illuminated by 16 lights has much better visual quality than *leather* illuminated by the exactly same lights.

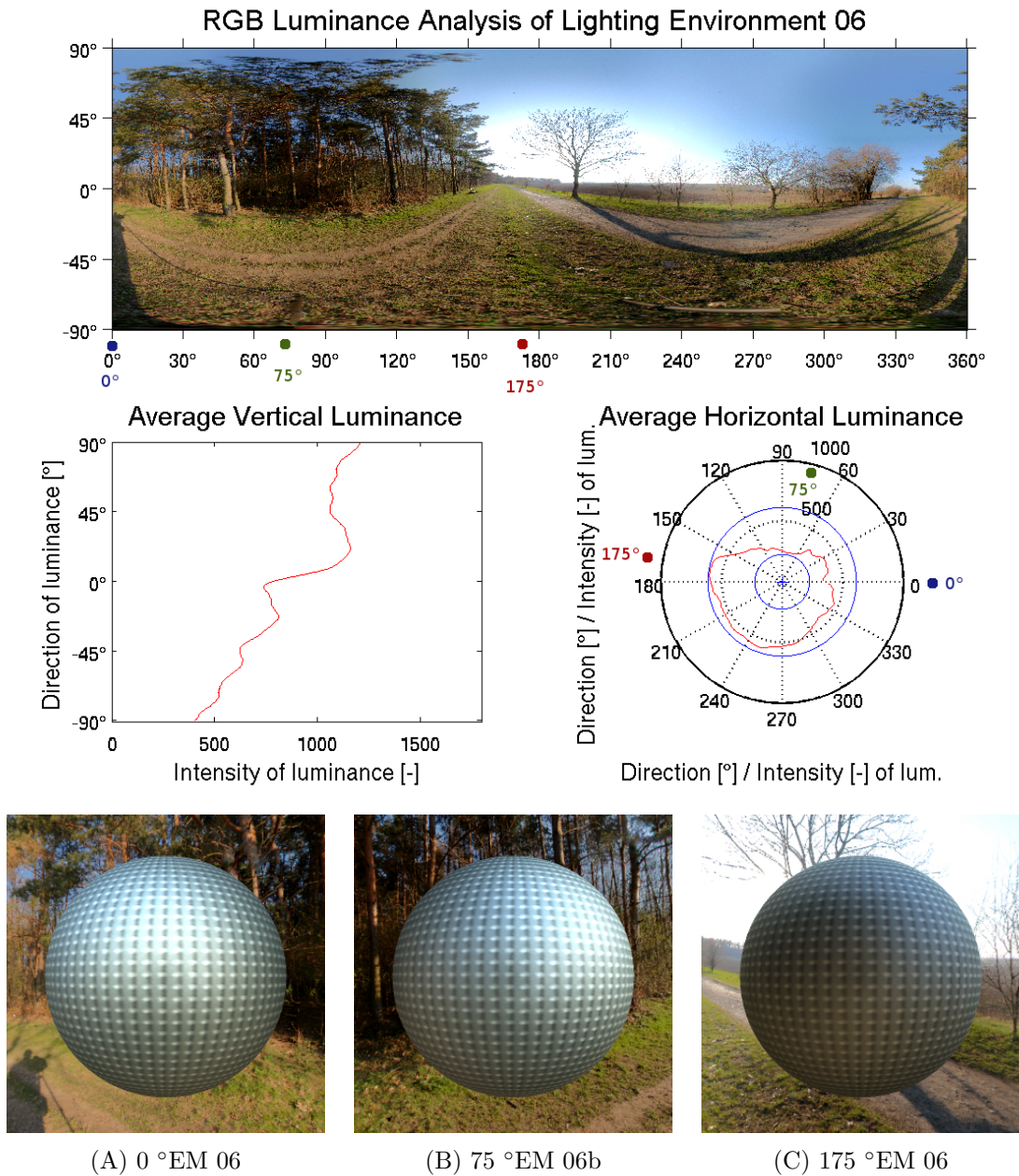


Figure 6.5: Figure compares visual quality of *alu* sphere illuminated by 3 different azimuthal angles from the same EM 06. The highest horizontal luminance is for 175°angle represented in Tab. 6.1 by EM 06. The lowest horizontal luminance is at 0°represented by EM 06a. The medium luminance is for 75°angle represented by EM 06b. The upper figure shows luminance analysis of EM 06.

Another challenge was to study different azimuthal angles of EMs which were used for the sphere illumination. In all previous experiments the spheres were illuminated by azimuthal angles from EMs with the highest horizontal luminance.

But some EMs have too diverse horizontal luminances which could affect the visual quality of tested material. To investigate this condition the spheres were not just illuminated by highest horizontal luminance, but additionally also by the lowest horizontal luminance (it was applied for EM 03a, EM 04a, EM 08a (see Figures 5.8 and 5.10)). For EM 06 was sphere even illuminated by 3 different azimuthal angles (next to highest and lowest also by median luminance (EM 06, EM 06a, and EM 06b) (see Fig. 5.9)).

The psychophysical data obtained from the experiment displayed in Tab. 6.1 showed that azimuthal angle with lowest horizontal luminance had direct effect on observers ability to spot visual degradations of all tested materials. In other words the lower number of lights were sufficient to keep material's high visual quality. These results were expected because it is naturally harder to observe an object illuminated by less light. In the Fig. 6.5 one can see how was the visual appearance of *alu* material affected by different angles of views for EM 06 (all spheres were illuminated by 256 lights).

Together 20 participants (8 females, 12 males) undertook this experiment. So 40 % of observers were women and 60 % were men, so it is almost ratio 50 to 50 which gives opportunity to "compare" visual perception of woman to man according to this experiment. Note that discussion does not want to "state" anything. Figure 6.6 is dedicated to this topic. Generally men were more sensitive to the visual quality of tested materials than women (see Fig. (A, D) 6.6). On the other hand men were observing stimuli longer than women did (see Fig. (B, C) 6.6).

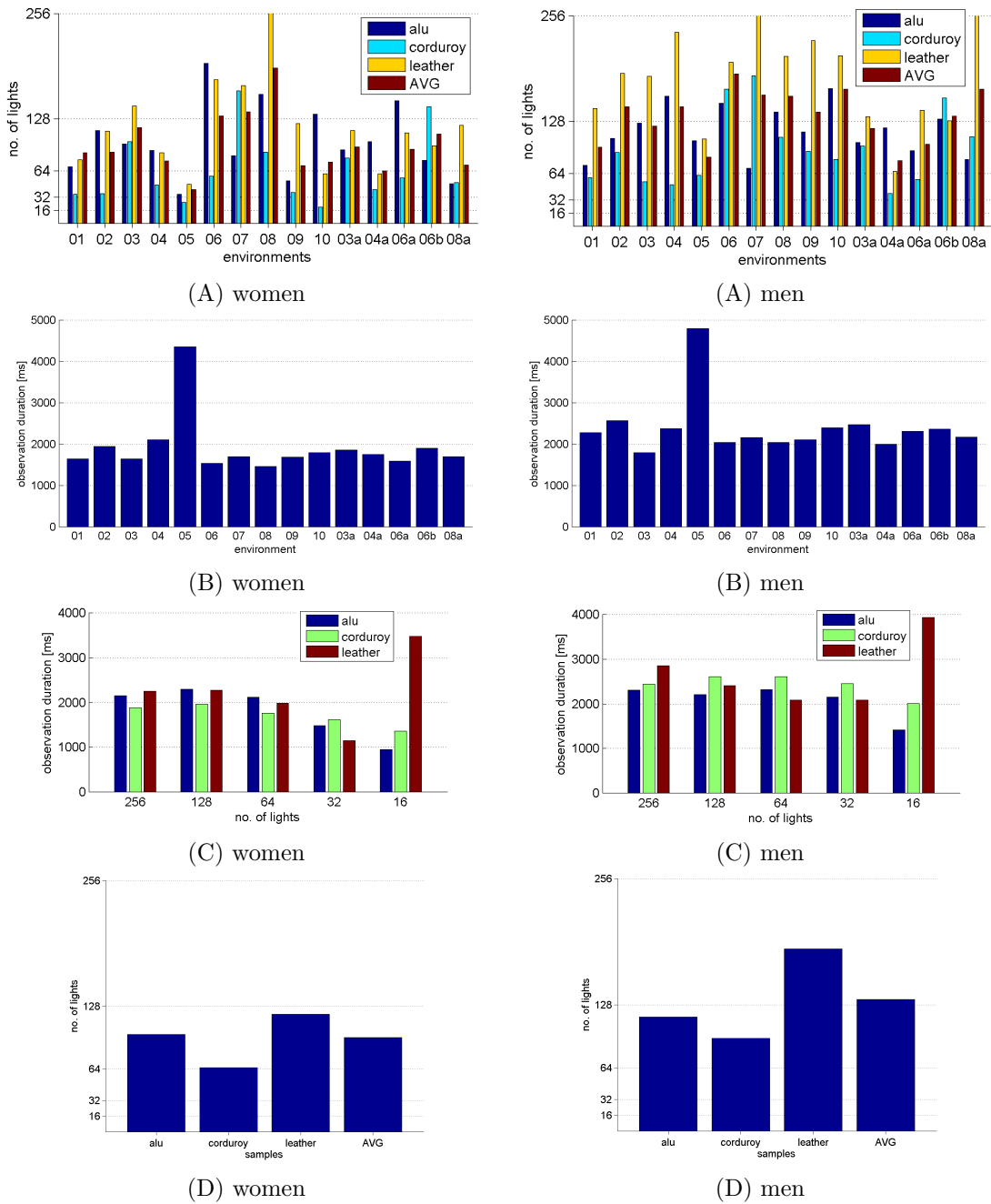


Figure 6.6: Visual perception of women and men is ”compared” according to psychophysical data obtained from the experiment. (A) compares the number of lights for all 3 tested materials related with EMs, (B) compares the observers time related with EMs. (C) compares the observers time related with materials. (D) compares the AVG number of lights from all EMs for all 3 tested materials.

6.2 Validation Experiment

A validation experiment was done in order to support and confirm the results obtained from the first experiment. Stimuli and experimental procedure was same as in previous experiment with exception that the experiment was not performed under dim room during the night as in previous study, but instead it was performed during the day in the school library. Note that during the day is harder to observe display due to reflections caused by sun. Also notice that the experiment was done during the day for no specific purpose, just that during the day one can find full library of potential voluntary participants for the experiment instead of night when is library closed. On the other hand it allowed us to compare results from our previous experiment where the illumination conditions were strictly controlled.

This experiment undertook 9 voluntary observers (4 females, 5 males) all students under 25 years old. The psychophysical data obtained from experiment were analysed the exactly same way as data in previous experiment. Fig. 6.7 compares data from both experiments. The results are very similar with the previous experiment, which enhance the significance of the first experiment. We can conclude that the validation experiment did not shown any differences with the prior experiment.

6.3 Statistical Prediction

In this chapter we attempt to find statistical method that predicts the number of lights for all tested materials (*alu*, *corduroy*, *leather*) for specific environments. So one will exactly know how many lights should be used for rendering these materials to acquire the high visual quality in specific environment. This method is based on the data obtained from our previous psychophysical experiments.

The relationship between data obtained from the psychophysical experiments and statistical properties of environment maps was studied by means of the correlation function 6.8 which is statistical function expressing linear relationships among observed data values. Correlations are useful because they can indicate a predictive relationship that can be exploited in practice. So in our case, they can help to identify the appropriate statistics related to number of lights obtained from the psychophysical experiment.

There are several correlation coefficients, often denoted to mean or standard deviation, measuring the degree of correlation. The most common of these is the Pearson correlation coefficient 6.8, which is mainly sensitive to a linear relationship

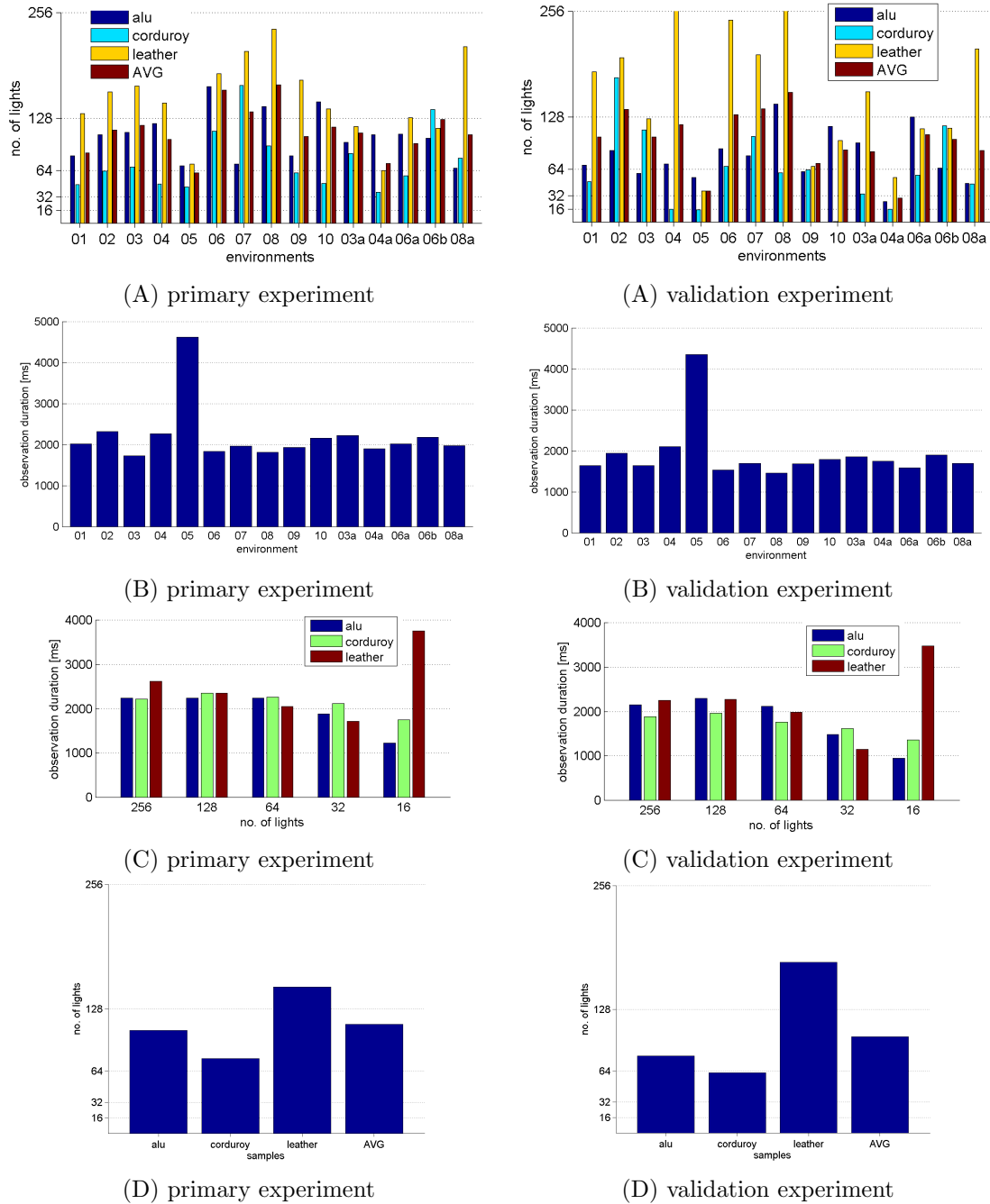


Figure 6.7: Primary experiment is compared with the validation experiment. (A) compares the number of lights for all 3 tested materials related with EMs, (B) compares the observers time related with EMs. (C) compares the observers time related with materials. (D) compares the AVG number of lights from all EMs for all 3 tested materials.

between two variables, one that is used in for the purpose of statistical prediction

$$\rho_{X,Y} = \text{corr}(X, Y) = \frac{\text{cov}(X, Y)}{\sigma_X \sigma_Y} = \frac{E[(X - \mu_X)(Y - \mu_Y)]}{\sigma_X \sigma_Y} \quad (6.4)$$

where E is the expected value operator , cov means covariance, and, corr a widely used alternative notation for Pearson’s correlation. The Pearson correlation is +1 in the case of a perfect positive (increasing) linear relationship, -1 in the case of a perfect decreasing (negative) linear relationship, and some value between -1 and +1 in all other cases, indicating the degree of linear dependence between the variables. As it approaches zero there is less of a relationship. The closer the coefficient is to either -1 or +1, the stronger the correlation between the variables [1]. Figure 6.8 shows the example of two sets of (x,y) points with their correlation coefficient.

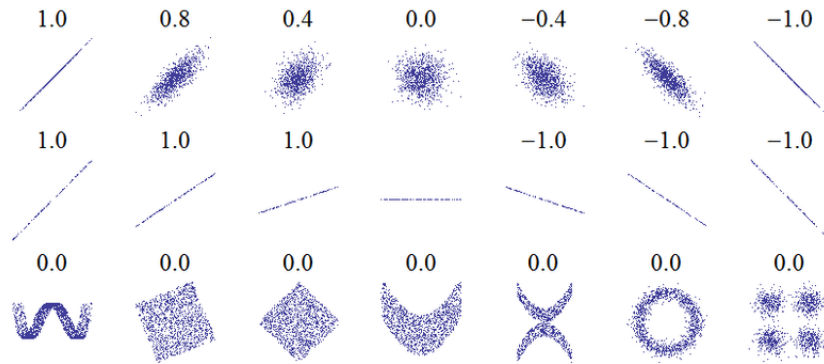


Figure 6.8: The correlation between two sets of points (x,y). Note that correlation reflects the noisiness and direction of a linear relationship (top row), but not the slope of that relationship (middle), nor many aspects of non-linear relationships (bottom) [1].

To analyse the environment maps by means of correlation we needed to get data from EMs. Due to this we computed several simple statistics from the analysed EM. The standard deviations (std) were calculated for every row (stdV) and column (stdH) of the each EM. After that the average (avg) and standard deviation of stdV and stdH were calculated from all previously calculated stdH and stdV in order to get values that characterise all EMs (so we got avgstdV, avgstdH, stdstdV, and stdstdH). For better orientation:

stdstdV - is std from std calculated for every row of EM

stdstdH - is std from std calculated for every column of EM

avgstdV - is avg from std calculated for every row of EM

	stdstdV	stdstdH	avgstdV	avgstdH	avg	std
pior exp.	0.1211	0.3481	0.6458	0.5082	0.4766	0.6680
validation exp.	-0.1457	0.4812	0.7794	0.6463	0.6385	0.7874

Table 6.2: Correlations between environment maps and data obtained from the psychophysical experiments.

avgstdH - is avg from *std* calculated for every column of EM

avg - is avg for all EM pixels

std - is std for all EM pixels

For these data and data based upon the observers answers from the psychophysical experiment were calculated correlations that are shown in the Tab. 6.2. For the computation were used psychophysical thresholds averaged over all materials (see Tab. 6.1).

The highest correlations were obtained for *std* of EMs in both primary and validation experiments. This could mean that there is a relationship between *std* of EM and visual perception of the observer which can be used as a statistical prediction for the number of lights to illuminate the specific material by the specific EM. Also the correlations between *avgstdV* in both experiments are very high, while *avgstdH* are lower. This implies that average horizontal variance in individual rows *avgstdV* in environment map affects visibility of illumination degradation artefacts less than average horizontal variance *avgstdH*. In other words, the less uniform and more variable is the environment map horizontally the more lights it requires to achieve the required visual quality.

Relationship between *std* value and number of lights for the tested materials is shown in Fig. 6.10. For example, how many lights is needed to acquire sufficient visual quality of *alu* material illuminated by some specific EM? One just needs to calculate *std* for the specific EM and according to the calculated value choose a number of lights in the Fig. 6.10 for the desired material. With this method one should get high visual quality of rendered illuminated material, while the number of lights used to represent the illumination is estimated in a way to save excessive computation of visual details which are visually indistinguishable by a human eye.

Five different EMs ((A) indoor EM during day with artificial light, (B) indoor EM during night with artificial light, (C) night EM, (D) EM with bright sun, and (E) EM during early sunset (see Fig. 6.11 and 6.9)) were tested for number of

lights obtained according to this method. For example for EM (A) was calculated std 54.6 for all pixels (using command *std2* in Matlab). According to Fig. 6.10 the appropriate number of lights for EM with std 54.6 is 128 for *alu*, 128 for *corduroy* and 256 *leather* (see Fig.6.11). As one can see the visual quality of all tested materials is high for all tested EMs, i.e., there are not visual material degradations, which confirms the suitability of the described statistical prediction method to choose the appropriate number of lights for the specific EM.

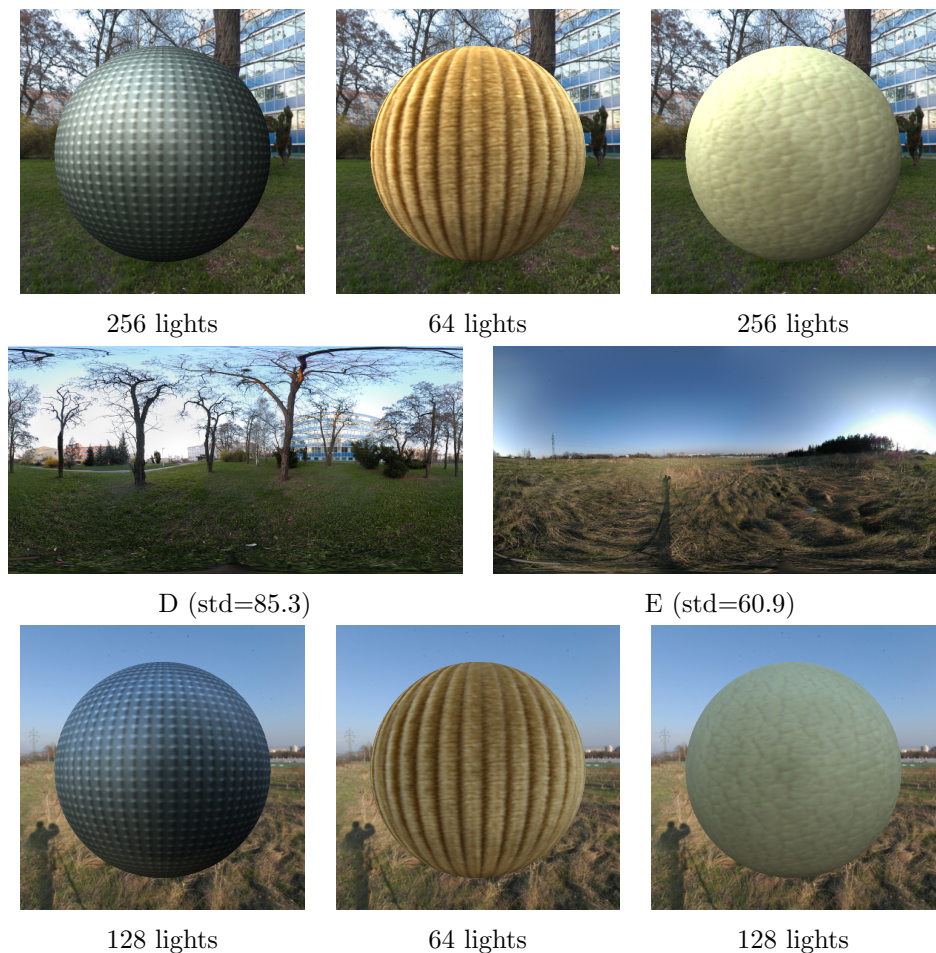


Figure 6.9: Application of the statistical prediction method (see Fig. 6.10) for the different EMs and BTF materials.

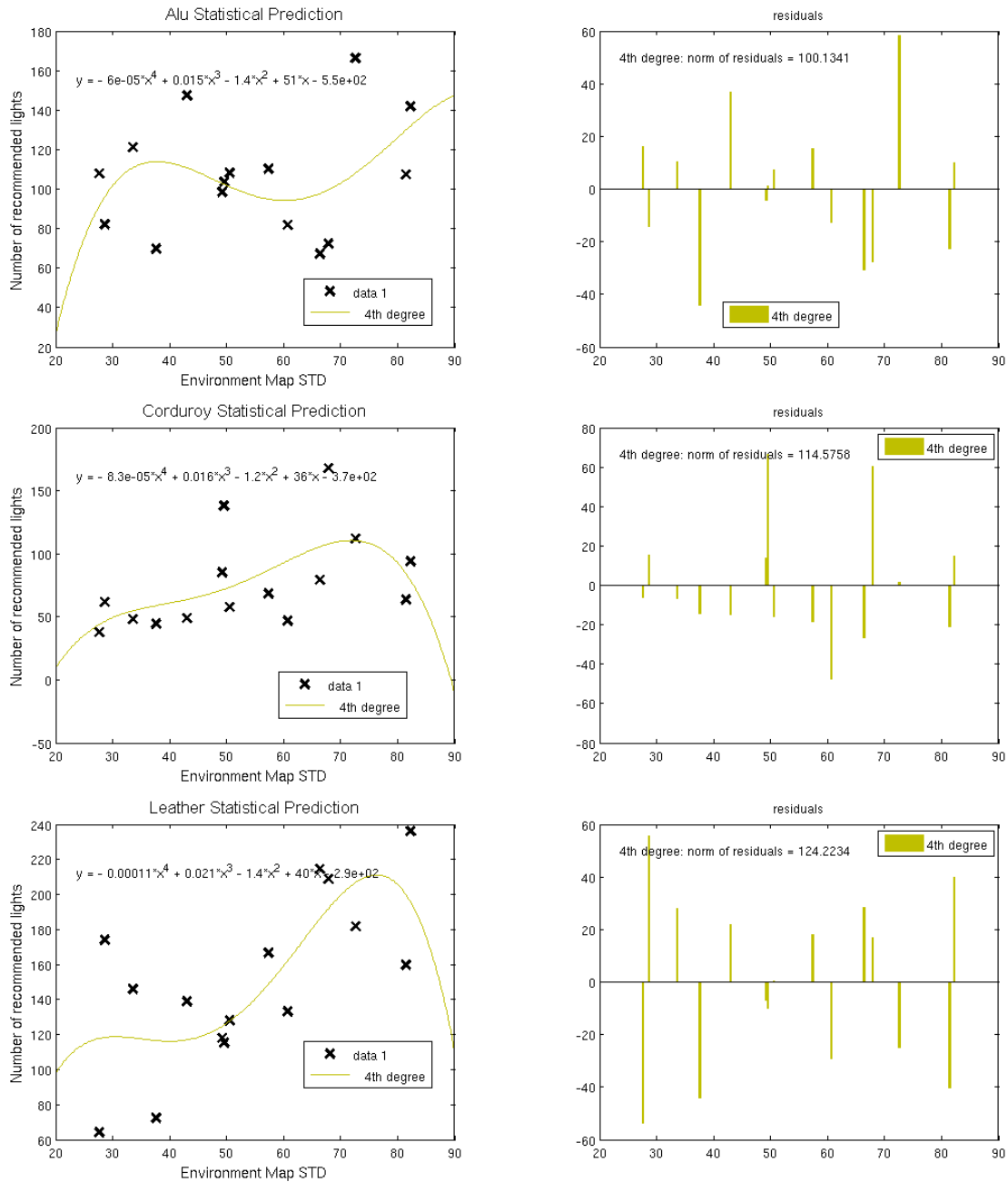


Figure 6.10: Statistical prediction of number of lights for the specific EMs. The data were fitted (using fitting tools in Matlab) by 4th degree polynomial function (PF) which is showed in the graph. Norm of residuals is quite high, due to high diversion of data, but nevertheless tested calculations showed that PF approximates number of lights for certain ranges of std. One just needs to put calculated std from the specific EM to the PF in order to get appropriate number of lights. If one is not satisfied with this result, there is still possibility to read data from the graphs visually without use of PF. This way is more accurate.

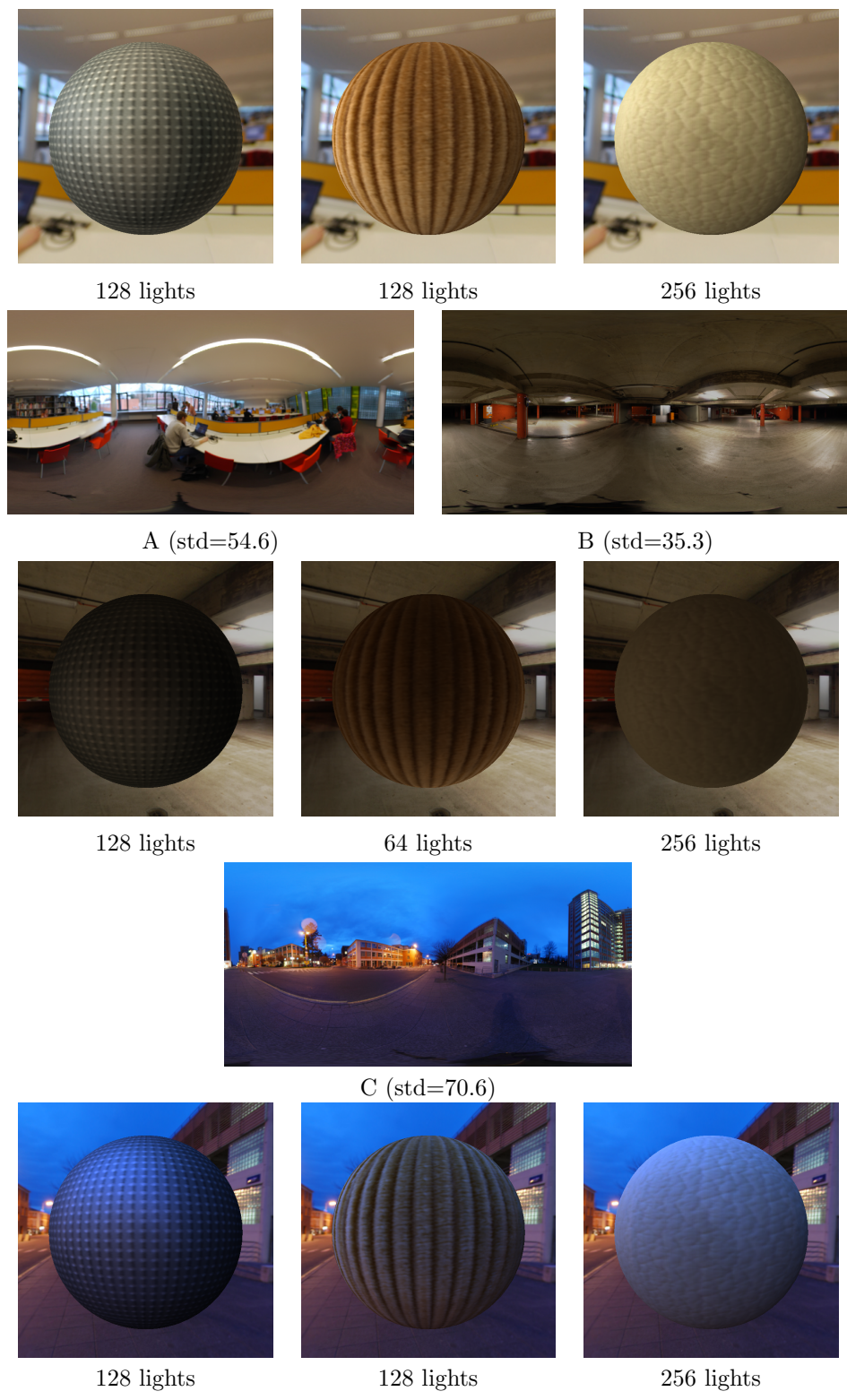


Figure 6.11: Application of the statistical prediction method (see Fig. 6.10) for the different EMs and BTF materials.

Chapter 7

Conclusions and Future Work

In this thesis, we have given an account to the human visual perception of materials illuminated by the real environment maps. This analysis was performed by means of visual psychophysical experiments with voluntary participants.

First, the different methods of acquiring 360° environment maps were studied (mirrored spheres, tiled photographs and fish-eye lenses). Capturing EM by the mirrored sphere together with telelenses showed as very fast and reliable, but the quality of captured EM was not so high as EM captured with fish-eye lenses which brought the best results of all. As a result of this study were created step-by-step manuals of EM acquisition for fisheye and mirrored sphere.

Second, the principle of HDR imaging was studied in order to acquire HDR environment maps. Several tone-mapping operators were tested to tone-map these maps from HDR to LDR.

After this prior preparation, ten 360° environment maps were acquired by fish-eye lens, which were used in experiment. Two of them were HDR which were tone-mapped to LDR. These maps were analysed based on their RGB luminance. The analysis resulted in consistent selection of visual stimuli that were used in the psychophysical experiment.

The captured environment maps were represented by different numbers of lights using median-cut algorithm. Median cut algorithm provided by Francesco Banterle together with rendering program and BTF materials provided by Jiří Filip were used to render 225 stimuli for the experiment. These stimuli were used for the psychophysical experiments with 20 voluntary participants which was aimed to study human visual perception of three tested BTF materials (*alu*, *corduroy*, *leather*) illuminated by the ten previously captured real environment maps. The data averaged

across experiment participants were fitted by psychometric functions, which were used to obtain the desired visual thresholds for all BTF materials according and environment maps. These thresholds represent the estimated number of lights necessary to achieve the required visual quality of illumination environment. They characterise the relationship of human visual perception with tested materials and environment maps. We also studied relationship of experiment results with several statistics computed over environment maps. We have found that standard variance over environment maps shows high correlation with human visual perception. We have show how can be this statistics used for prediction of appropriate number of lights for arbitrary environment. To verify the results from the experiment we performed a validation experiment with 9 voluntaries. The validation results successfully confirmed the results of the primary experiment.

Our results can be practically applied for speed enhancement of digital material visualization methods, by preserving a required visual quality while computing only the data only up to resolution of human-eye. Due to this, we can gain considerable speed increase namely for rendering of accurate and complex material appearance representations (BRDF, BTF, etc.).

In the future our work can be tested on other BTF materials and extended to other algorithms representing illumination environment by a set of point lights.

Bibliography

- [1] Frederick E. Croxton, Dudley J. Cowden, and Sidney Klein. *Applied general statistics*. 1939.
- [2] K. J. Dana, B. van Ginneken, S. K. Nayar, and J. J. Koenderink. Reflectance and texture of real-world surfaces. *ACM Transactions on Graphics*, 18(1):1–34, January 1999.
- [3] P. Debevec. A median cut algorithm for light probe sampling. In *Siggraph 2005 (Posters)*. ACM Press / ACM SIGGRAPH / Addison Wesley Longman, 2005.
- [4] P. Debevec, E. Reinhard, G. Ward, and S. Pattanaik. *HIGH DYNAMIC RANGE IMAGING Acquisition, Display, and Image-Bases Lighting*. The Morgan Kaufmann Series in Computer Graphics. Morgan Kaufman, 2006.
- [5] G. Der and I. J. Deary. Age and sex differences in reaction time in adulthood: Results from the united kingdom health and lifestyle survey. *Psychology and Aging*, (21):62–73, 2006.
- [6] Frederic Drago, William Martens, Karol Myszkowski, and Norishige Chiba. Design of a tone mapping operator for high dynamic range images based upon psychophysical evaluation and preference mapping. 5007:321–331, April 2003.
- [7] R. O. Dror, E. H. Adelson, and A. S. Willsky. Surface reflectance estimation and natural illumination statistics. In *Workshop on Statistical and Computational Theories of Vision*, 2001.
- [8] Frédo Durand and Julie Dorsey. Interactive tone mapping. pages 219–230, 2000.
- [9] Raanan Fattal, Dani Lischinski, and Michael Werman. Gradient domain high dynamic range compression. pages 249–256, 2002.

- [10] J. Filip. *High Dynamic Range Imaging Using PFStools*. <http://staff.utia.cas.cz/filip/index.html>, November 2009.
- [11] J. Filip and M. Haindl. Bidirectional texture function modeling: A state of the art survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 31(11):1921–1940, October 2009.
- [12] R. W. Fleming, R. O. Dror, and E. H. Adelson. How do humans determine reflectance properties under unknown illumination? In *Proceedings of CVPR Workshop on Identifying Objects Across Variations in Lighting: Psychophysics and Computation*, 2001.
- [13] R. W. Fleming, R. O. Dror, and E. H. Adelson. Real-world illumination and the perception of surface reflectance properties. *Journal of Vision (2003)*, (3):347–368, 2003.
- [14] Peter Gawthrop. *Creating Spherical Panoramas with the Canon 5D and 15mm Fisheye Lens*. <http://www.lightspacewater.net/Tutorials/PhotoPano2/paper/>, September 2007.
- [15] Vlastimil Havran, Miloslaw Smyk, Grzegorz Krawczyk, Karol Myszkowski, and Hans-Peter Seidel. Interactive system for dynamic scene lighting using captured video environment maps. In Kavita Bala and Philip Dutre, editors, *Rendering Techniques 2005 (Eurographics Symposium on Rendering)*, pages 31–42,311, Konstanz, Germany, June 2005. Eurographics Association.
- [16] J.E.Dowling. *The Retina: An Approachable Part of the Brain*. Belknap Press, Cambridge, 1987.
- [17] Jan Kautz, Pere-Pau Vázquez, Wolfgang Heidrich, and Hans-Peter Seidel. Unified approach to prefiltered environment maps. In *Proceedings of the Eurographics Workshop on Rendering Techniques 2000*, pages 185–196, 200.
- [18] Y. Ostrovsky, P. Cavanagh, and P. Sinha. Perceiving illumination inconsistencies in scenes. *Perception*, 34:1301–1314, 2005.
- [19] Sumanta N. Pattanaik, Jack Tumblin, Hector Yee, and Donald P. Greenberg. Time-dependent visual adaptation for fast realistic image display. pages 47–54, 2000.

-
- [20] Bruno Postle. *Hugin - Creating 360 enfused panoramas*. <http://hugin.sourceforge.net/tutorials/enfuse-360/en.shtml>, January 2008.
- [21] Guoping Qiu, Jian Guan, Jian Duan, and Min Chen. Tone mapping for hdr image using optimization - a new closed form solution. *ICPR 2006*, vol.1, 2006.
- [22] Ravi Ramamoorthi and Pat Hanrahan. Frequency space environment map rendering. *ACM Transactions on Graphics*, 21(3):517–526, July 2002.
- [23] G. Ramanarayanan, J. Ferwerda, B. Walter, and K. Bala. Visual equivalence: Towards a new standard for image fidelity. 26(3), 2007.
- [24] Erik Reinhard, Michael Stark, Peter Shirley, and James Ferwerda. Photographic tone reproduction for digital images. pages 267–276, 2002.
- [25] S.F. te Pas and S.C. Pont. A comparison of material and illumination discrimination performance for real rough, real smooth and computer generated smooth spheres. In *APGV '05: 2nd Symposium on Applied perception in graphics and visualization*, pages 57–58, 2005.
- [26] FELIX A. WICHMANN and N. JEREMY HIL. The psychometric function: I. fitting, sampling, and goodness of fit. *Perception & Psychophysics*, (63):1293–1313, 2001.
- [27] Wikipedia.org. *Gamma correction*.

List of Notations and Acronyms

AVG	Average
BTF	B idirectional T exture F unctions
Eq.	E quation
EM	E nvironment M ap
Fig.	F igure
HDR	H igh D ynamic R ange I mage
IBL	I mage B ased L ighting
LDR	L ow D ynamic R ange I mage
STD	standard d eviation
x, y, z	row, column and spectral index

List of Figures

1.1	Materials illuminated by a point light from left: (A) Aluminium , (B) Corduroy, (C) Leather. The same materials illuminated by the environment illumination (G): (D) Aluminium, (E) Corduroy, (F) Leather. [11].	10
1.2	Leather bunny illuminated by environment map represented by (A) 8 lights of EM, (B) 128 lights, and (C) 256 lights.	11
2.1	EM of U13 library in Zlin consisted of 20 images taken by Nikon D-60 and stitched in Hugin.	15
2.2	Figure shows 180° capturing capability of used fisheye lenses (NIKKOR 10.5mm).	16
2.3	EM of U13 library in Zlin consisted of 8 images taken by Nikon D-60 with fisheye lens NIKKOR 10.5mm and stitched in Hugin.	17
2.4	EM of U13 library in Zlin consisted of 2 images taken 90 degrees apart by Canon EOS 400D with 300 mm telephoto lenses Sigma, and stitched in HDR Shop and Hugin.	17
3.1	The set of three low dynamic range images with different exposure is taken and stitched together using PFSTools (manual can be found at http://staff.utia.cas.cz/filip/index.html). The HDR image can be created using pseudo linear combination of these three images. One can interactively change the "exposure" of HDR with the slider to obtain LDR image [10].	19
3.2	(A) The simple global filter is used to tone map photography of the forest near Prague (see Fig. (B)). As one can see the image is after global tone mapping hardly recognized, contrast is very low.	20
3.3	Different tone-mapping operators applied on the HDR image (see Fig. 3.1) to scale them into dynamic range of display as shown in [10]: (A) [4] (B) [24] (C) [19] (D) [9] (E) [8] (F) [6]	21

4.1	Figures (A, B, C, D) show median-cut's performance for the first 4 loops. Every step the regions are cut along the longest dimension such that its light energy is divided evenly. This process is done till the desired number of lights is found. The number of lights that can be found by median-cut is represented by power of two.	24
4.2	The same image with 3 different levels of gamma (from the left $\gamma = 1, \gamma = 2, \gamma = 3$) [27].	24
4.3	Figure (A) shows 128 lights found by median-cut algorithm. The aluminium spheres are illuminated by different number of lights from environment map (see Fig. (A, B, C, D)). The quality of shadows, reflections, and shading is clearly dependent on number of lights. All spheres are realistically illuminated as if they were in bookshop, for example the bottom side is more illuminated than top because of indirect light reflected from the floor.	25
4.4	(A) Uniform sampling of a hemisphere above a material sample, i.e. 81 sampling directions [11]. (B) Six-dimensional BTF function, representing the appearance of a material sample surface (θ are elevation and φ azimuthal angles, (x, y) planar position over a material surface).	26
5.1	Camera orientation. The firm lines show the four directions at 30° , the dashed lines at 30° down. The numbers show the order in which the images are taken [14].	29
5.2	Nominal exposure (0 EV) of pictures taken according to Gawthrop method (see Fig.5.1).	29
5.3	Under exposed (0 EV) pictures taken according to Gawthrop method (see Fig.5.1).	30
5.4	Over exposed (0 EV) pictures taken according to Gawthrop method (see Fig.5.1).	31
5.5	Nominal, under exposed, and over exposed EM created from 8 pictures taken in the forest near the Prague according to Gawthrop method (see Fig.5.1). HDR EM was created by stitching them together in the Gimp.	33
5.6	Process of acquiring EM with a mirrored sphere. The calibration of sphere reflectivity must be done due to imprecise surface of mirrored ball. In this case, the sphere reflectivity is 64 % $((0.38/0.59)*100)$	35

5.7 Lighting environment 01 is a ground floor of Tomas Bata skyscraper 21. Indoor light is almost uniformly distributed and it has almost highest luminance from all lighting EM (with highest horizontal luminance at 195°). Lighting environment 02 is Archa bookshop. It has specific vertical light distribution, i.e. the brightest part is floor, because of reflected light from the top (with highest horizontal luminance at 50°). 37

5.8 Lighting environment 03 is Zlin city centre during bright day. Because of sun hidden behind clouds is naturally highest horizontal luminance at 285°. On the other hand the rest of the EM has much smaller horizontal luminance (see angle 30°). Due to this are chosen two different angles to be used for rendering in next chapter. Lighting environment 04 is Zlin’s subway during the night. The luminance is more than half smaller compared with day light. Again due to non uniform light distribution are chosen two angles for rendering (brightest at 245° and darkest at 345°). 38

5.9 Lighting environment 05 is a Golden Apple cinema entrance. The luminance is one of the lowest from all EMs, because light is very dim with the highest horizontal luminance at 240°. Lighting environment 06 is very interesting due to three different horizontal luminance regions. One is when the sun is at 175°, naturally most brightest. The medium luminance is spread around 0° and the lowest is at 75°. 39

5.10 Lighting environment 07 is an big open space of Tomas Bata library where natural light comes inside through the glassy roof, so highest vertical luminance is at the top, i.e. 90°. The highest horizontal luminance is at 355°. Lighting environment 08 is a Ďáblický háj near Prague, where sun is just coming down at horizontal angle 7°. At angle 240° is the lowest horizontal luminance. 40

5.11 Lighting environment 09 is a interior of the church in Zlin with very unusual light conditions that can one find just in church. The lowest horizontal luminance is at 0°. Lighting environment 10 is a Zlin city centre during night, so the light is coming from the moon, stars and surrounding buildings. It is the dimmest lighting EM from all tested. The highest horizontal luminance comes from over-saturated town-hall at 90°. 41

5.12 BTF materials (*alu*, *corduroy*, and *leather*) illuminated by 256 lights from the real EMs (shown in the Fig. 5.7, 5.8, 5.9, 5.10, 5.11). These rendered images were used as stimuli for the experiment presented in the next chapter. 42

6.1	Example of stimuli used in experiment. On the left side is showed alu sphere illuminated by 32 lights, which is compared with identical sphere illuminated by 256 lights from the same EM which is also used as a background of the stimuli.	44
6.2	Psychometric function for 25 % threshold of psychophysical data from experiment.	47
6.3	(A) Observation duration for all tested materials. (B) Sufficient number of lights for all EMs.	49
6.4	Figures compares visual quality of two different BTF materials. Notice that <i>alu</i> illuminated by 16 lights has much better visual quality than <i>leather</i> illuminated by the exactly same lights.	49
6.5	Figure compares visual quality of <i>alu</i> sphere illuminated by 3 different azimuthal angles from the same EM 06. The highest horizontal luminance is for 175°angle represented in Tab. 6.1 by EM 06. The lowest horizontal luminance is at 0°represented by EM 06a. The medium luminance is for 75°angle represented by EM 06b. The upper figure shows luminance analysis of EM 06.	50
6.6	Visual perception of women and men is "compared" according to psychophysical data obtained from the experiment. (A) compares the number of lights for all 3 tested materials related with EMs, (B) compares the observers time related with EMs. (C) compares the observers time related with materials. (D) compares the AVG number of lights from all EMs for all 3 tested materials.	52
6.7	Primary experiment is compared with the validation experiment. (A) compares the number of lights for all 3 tested materials related with EMs, (B) compares the observers time related with EMs. (C) compares the observers time related with materials. (D) compares the AVG number of lights from all EMs for all 3 tested materials.	54
6.8	The correlation between two sets of points (x,y). Note that correlation reflects the noisiness and direction of a linear relationship (top row), but not the slope of that relationship (middle), nor many aspects of non-linear relationships (bottom) [1].	55
6.9	Application of the statistical prediction method (see Fig. 6.10) for the different EMs and BTF materials.	57

- 6.10 Statistical prediction of number of lights for the specific EMs. The data were fitted (using fitting tools in Matlab) by 4th degree polynomial function (PF) which is showed in the graph. Norm of residuals is quite high, due to high diversion of data, but nevertheless tested calculations showed that PF approximates number of lights for certain ranges of std. One just needs to put calculated std from the specific EM to the PF in order to get appropriate number of lights. If one is not satisfied with this result, there is still possibility to read data from the graphs visually without use of PF. This way is more accurate. 58
- 6.11 Application of the statistical prediction method (see Fig. 6.10) for the different EMs and BTF materials. 59

List of Tables

6.1	Thresholds of sufficient number of lights for all tested materials and EMs obtained from the experiment. Figure also shows the AVG of observer's time response dependent on EM.	48
6.2	Correlations between environment maps and data obtained from the psychophysical experiments.	56