IST-2001-34744

RealReflect

D10.4: Perceptual Evaluation of Tone Mapping Operators

Author: Akiko Yoshida <yoshida@mpi-sb.mpg.de>
Max-Planck-Institut für Informatik
Author: Karol Myszkowski <karol@mpi-sb.mpg.de>
Max-Planck-Institut für Informatik
Author: Vlastimil Havran <havran@mpi-sb.mpg.de>
Max-Planck-Institut für Informatik
Author: Alessandro Artusi <artusi@cg.tuwien.ac.at>
Instituite of Computer Graphics and Algorithms

Version: 1.1 (final)
Date: 2004-Sep-14
Due: 2005-Feb-28

Abstract

This document is the first version of the perceptual evaluation of the tone mapping operators within the REALREFLECT project. It deals with the comparison of tone mapping operators for the high dynamic range images from the architectural environments.

Contents

Chapter 1

Introduction

The need of tone mapping high dynamic range (HDR) images has highly increased recently because they are useful not only for static images but also for multimedia applications. Therefore, how to produce visually sufficient HDR images has been one of the important discussions in the computer graphics community for years against the problems which come from the limited number of bits in image formats and the limited display range of the physical devices. A number of techniques have been introduced to conquer those problems including construction and storage of HDR images. To represent HDR images on today's physical devices, a number of successful tone mapping operators have also been presented. They are useful for HDR photography and also for lighting simulations such as realistic rendering and global illumination techniques.

In order to produce visually sufficient images, a knowledge about human visual systems (HVS) cannot be ignored. Many tone mapping operators have been produced based upon HVS theory. Obviously, HVS theory is also useful to judge a quality of images. This document focuses upon the human perception when people compare tone mapped images with their corresponding real-world views.

In this document, overviews of HDR images and our HDR image acquisition are presented in chapter 2. In chapter 3, seven existing tone mapping operators are briefly described. Those operators were used in our perceptual experiment. The procedure and result of our experiment is shown in chapter 4. The set of data was analyzed by using multivariate statistical methods.

Chapter 2

High Dynamic Range Images

2.1 Introduction

Luminance in the real world processed by human observers can have much greater dynamic range than the range of a photographic film and electronic apparatus. Images whose dynamic range is huge (i.e., images which have both very bright and very dark parts) are called high dynamic range (HDR) images. We can create an HDR image from multiple photographs taken with different exposures. The exposure is changed by the shutter speed of a camera. Underexposed pictures can show details in highlighted areas and overexposed pictures can show details in dark areas.

This chapter presents overviews of HDR image encoding techniques, HDR image formats, and our HDR image acquisition.

2.2 HDR Images Encoding Techniques

In 1995, Mann and Picard introduced the first algorithm to construct an HDR image from multiple photographs with different exposures [MP95]. They presented a way to recover a camera response curve function and apply it to input images. Other techniques after the Mann-Picard method also follow this way while each of them has unique philosophy to recover the response curve. The other well-known algorithms are the Debevec-Malik [DM97], the Mitsunaga-Nayar [MN99], the Robertson et al. [RB99], and the Ward [War03] methods. For our HDR image construction, we used the Robertson et al. algorithm to recover the camera response curve.

2.3 HDR Image Formats

There are a number of formats for HDR images: the Pixar Log Encoding, the Radiance RGBE, the LogLuv, the OpenEXR, scRGB, etc. We used the Radiance RGBE format to save our HDR images. The Radiance RGBE format uses one byte for each of read, green, and blue components and one byte for an exponent. The exponent part is a scaling factor on RGB components, which is equal to two raised to the power of the exponent minus 128. A survey over HDR image formats was presented by Ward [WarXX].

2.4 Acquiring HDR Images

Figure 2.1: Scene 1 for our perceptual experiment.

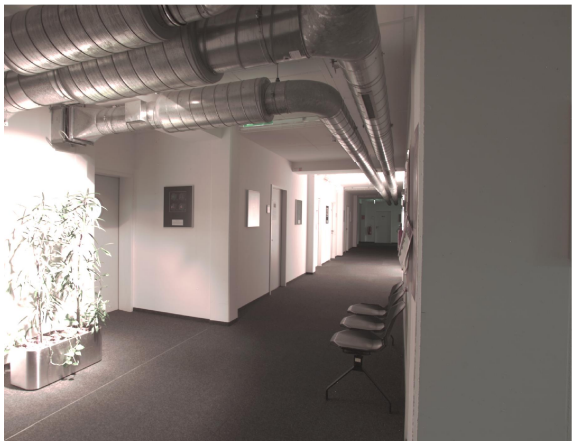

Figure 2.2: Scene 2.

	Max.	**Min.**	**Dynamic range**	$\bar{L}_w$
Scene 1	4878.3040	0.0079	617090:1	4.4096
Scene 2	160.0750	0.0062	25967:1	5.0066

Table 2.1: Maximum and minimum pixel luminances, dynamic ranges, and logarithmic average luminances $\bar{L}_w$.

For our perceptual experiment, static images with different shutter speeds were taken by a camera, Kodak Professional DCS560, with lenses, CANON Lense EF 24mm and 14mm at the Max-Planck-Institut für Informatik in Saarbrücken, Germany. Because those images were saved in a raw format of Kodak, they were converted to 36-bit TIFF format by using a program `raw2image` which is involved in a Kodak DCS Photo Desk package. Then, the response curve of the camera was recovered by using the Robertson et al. method. The advantage of this method is that no assumption is necessary for the response curve and response curves can be reconstructed in any

arbitrary shape while the Debevec-Malik method assumes that the response curve is smooth and the Mitsunaga-Nayar method returns a polynomial response curve. Since our camera was set in a stable condition, it was not necessary to use the Ward method. His algorithm can be applied when hand-held pictures are taken. After the response curve was recovered, HDR images were created and then saved in the Radiance RGBE format. For constructing one HDR image, 15 images were taken for our project with the range of the shutter speeds from 1/2000 to 8.0 seconds. The HDR images for our experiments and statistical analysis are shown in Figure 2.1 and Figure 2.2 and Table 2.1.

The logarithmic average luminance factor indicates whether a scene is light, normal, or dark. It is calculated as

$$\text{Log Dynamic Range} = \log_2 L_{max} - \log_2 L_{min} \tag{2.1}$$

$$\bar{L}_w = \exp\left(\frac{1}{N}\sum_{x,y}\log(\delta + L_w(x,y))\right) \tag{2.2}$$

where L_{max} and L_{min} are the maximum and minimum luminance values of a scene respectively, $L_w(x,y)$ is a luminance value of pixel (x,y), N is the number of pixels, δ is a small value to avoid singularity. The idea of the logarithmic average luminance $\bar{L}_w$ was proposed by Reinhard et al. in [RSSF02].

Chapter 3

Tone Mapping Operators

3.1 Introduction

The concept of a tone reproduction was first introduced to the computer graphics community by Tumblin and Rushmeier in 1993 [TR93]. The goal of tone reproductions is to compress the dynamic range of an image, which has much bigger luminance range than that of the physical devices, to the displayable dynamic range. A number of tone reproduction techniques have been presented, and most of the tone mapping operators can be categorized into two types: global or local operators. Global operators apply the same transformation onto every pixel of an image. Local operators choose different scales onto different areas of an image. Global operators are more easily implemented than local ones because they need to find only one transformation, but local operators can usually provide better results. One of the well-known problems with local operators is halo effect (inverse gradients). This is manifested as a dark aura around a very bright light source.

This chapter provides an overview of the seven existing tone mapping operators which were used in our perceptual experiment.

3.2 Global Operators

The simplest tone reproduction is a linear mapping. It scales the radiances onto the range between 0 and 255 by trivial method. If the logarithm of the radiances is taken and linearly scaled to $[0,255]$, it is called a logarithmic linear mapping.

The histogram adjustment tone mapping operator was presented by Ward Larson et al. in 1997 [WLRP97]. They extended the earlier works of Ward [War94] and Ferwerda et al. [FPSG96]. This method uses the knowledge that human eyes are sensitive to relative rather than absolute changes to luminances. This approach avoids halos altogether. Although this method still has several problems, it can provide the best combination of practical simplicity and uniformly good perceptually advocated results.

In 2000, Pattanaik et al. presented a new time-dependent tone mapping operator which is based upon psychophysical experiments and creates color image sequence from any input scene [PTYG00]. Their model dealt with the changes of threshold visibility, color appearance, visual acuity, and sensitivity over time. It is also useful to predict the visibility and appearance of scene features.

Drago et al. presented a perception-motivated tone mapping algorithm for interactive display of HDR images in 2003 [DMAC03]. They used logarithmic functions with different bases for darkest and brightest area of an image to compress luminance values. To compress other luminance values for the middle range of luminances, a bias function was used. Their method can be applied to HDR video sequences.

3.3 Local Operators

Reinhard et al. produced the photographic tone reproduction, which is inspired by "dodging and burning" used in photography in 2002 [RSSF02]. This method works on various types of HDR images and its computational speed is very good. Additionally, the recent version of this method operates automatically freeing the user from setting parameters that are not particularly intuitive.

Ashikhmin presented a new tone mapping method which works on multipass approach [Ash02]. His method calculates local adaptation luminance, applies a tone mapping function using the threshold vs. intensity (TVI) functions, and then calculates the final pixel values to preserve details throughout an image. The TVI functions are linearly approximated in his algorithm.

Durand and Dorsey presented an edge-preserving filtering called the bilateral filtering method [DD02]. This method considers two different spatial frequency layers, base and detail, then reduces contrast and preserves details of an image.

Chapter 4

Perceptual Evaluation

4.1 Introduction

Image comparison measurement methods are classified into three categories: human-perception based, objective, and HVS-based methods. Human perception measurement is a subjective measurement method with human participants. Statistical methods are usually used to analyze the result from human subjects. Objective measurement is based upon theoretical models. HVS-based measurement is also defined mathematically as objective methods, but it is based upon HVS theory. A number of techniques and surveys have been presented in each of those three categories.

This chapter presents our perceptual evaluation with human observers and its results. The set of data was analyzed by using the Statistics Toolbox of MATLAB [Mat].

4.2 Experimental Design

For our project, a human perception-based measurement was selected with seven tone mapping operators. The tone mapping operators are the linear tone mapping, fast bilateral filtering by Durand and Dorsey [DD02], the Pattanaik et al. method [PTYG00], the Ashikhmin method [Ash02], the Ward method [WLRP97], the photographic tone mapping by Reinhard et al. [RSSF02], and the adaptive logarithmic mapping by Drago et al. [DMAC03]. For all of the reproductions, the gamma value was set $\gamma = 2.2$ and the maximum display luminance was set to 100 cd/m^2. For Pattanaik method, luminance values are scaled by a factor of 650. The local contrast threshold of Ashikhmin method was set to 0.1. The other parameters used the default values as presented in each of the papers introducing the tone mapping technique. Those images are shown in Appendix B.

Two scenes for our experiment are shown in Figure 2.1 and Figure 2.2. As clearly seen in the figures, the dynamic ranges of both scenes are large enough. Scene 1 has highly bright spot lights around the trees and quite dark areas behind the glass. Scene 2 also has highly bright and dark areas, and in addition, it has gray area (on pipes) which is pointed in a scaling problem [GKF+99]. The scaling problem concerns how the range of luminances in an image is mapped onto a range of perceived grays. The absolute luminances in both scenes were measured by MINOLTA light meter LS-100. In Scene 1 (Figure 2.1), the brightest area is 13,630 cd/m^2 and the darkest area is 0.021 cd/m^2. In Scene 2 (Figure 2.2), the brightest area is 506.2 cd/m^2 and the darkest area is 0.019 cd/m^2.

All of 14 human subjects were graduate students and researchers of the Computer Graphics group in the Max-Planck-Institut für Informatik. Two of them are female and the rest are male. The range of their ages is 24 – 34. All of them had normal or corrected eyesights to normal vision. Additionally, all subjects were naïve enough for the goal of our experiment and tone mapping operators.

Participants were asked to view seven images one after another for each of the two scenes (see Figure 2.1 and Figure 2.2) of the Max-Planck-Institut für Informatik on an sRGB-calibrated monitor (DELL UltraSharp 1800FP) whose resolution is 1280×1024 at 60.0 Hz. For each of the 14 images, they were asked to compare them with their corresponding real-world view and give ratings for image appearance and realism. Image appearance attributes are overall brightness, contrast, details in the dark region, and details in the bright region. The realism rating is only naturalness. For the image appearance attributes, subjects rated how much brightness, contrast, or details each of the images has compared to its corresponding real-world view. For the naturalness, they were asked to rate how real the image is. All of the ratings were done by moving scroll bars. They were allowed to move back and forth among images for a scene. A screenshot of our perceptual experiment is shown in Figure 4.1. Whole procedure for one participant took approximately 20 to 30 minutes.

Figure 4.1: A screenshot of our experiment.

4.3 Results

4.3.1 Introduction

Our experiment is seven (tone mapping operators) $\times$ two (scenes) within-subjects design (see [Tab89] for the details of the design types). In this experiment, there are two independent variables (IVs) and five dependent variables (DVs). IVs are tone mapping operators and scenes. DVs are all of the attributes: overall brightness, contrast, details in dark regions, details in bright regions, and naturalness. Our primary interest on this perception test is whether the images produced by different tone mapping operators are perceived differently when they are compared with their corresponding real-world views. To analyze the set of data obtained from the perception test, the Statistics Toolbox of MATLAB is used [Mat]. As preliminary data processing, all scores were

normalized over each of the attributes on each of the subjects in order to let the standard deviation value equal to 1.0. The normalization was done by

$$x_i \rightarrow \frac{x_i - \mu_x}{\sigma_x} \tag{4.1}$$

where x_i is a score and μ_x and σ_x are respectively the mean and the standard deviation over an attribute of a subject. Those normalized values are shown in Appendix C.

4.3.2 Main Effects of Scenes and Tone Reproductions

4.3.2.1 Scenes

Figure 4.2 – Figure 4.6 show the result for the main effect of the two scenes with F and p values. Both F and p values are the significance scales. As F increases, p decreases. Bigger F value (i.e. smaller p value) indicates that the difference between IVs are more significant. Those results show that the difference between those two scenes is not significant. Only the detail reproductions in dark regions show the significant difference, but for most of the attributes, the effect of Scenes' difference is small enough to be ignored. Multivariate Analysis of Variance (MANOVA) in MATLAB provides the Mahalanobis distance between IVs (see [Tab89] for details of MANOVA). The concept of Mahalanobis distance was introduced by Mahalanobis in 1936 [Mah36]. It is a measure based on correlations between variables and given as

$$\mathbf{S} = \frac{1}{n-1}\sum_{i=1}^{n}(\mathbf{X}_i - \bar{\mathbf{X}})^t(\mathbf{X}_i - \bar{\mathbf{X}}) \tag{4.2}$$

where $\mathbf{X}$ is a data matrix, $\mathbf{X}_i$ is the i-th row of $\mathbf{X}$, $\bar{\mathbf{X}}$ is a row vector of means, and n is the number of rows. A Mahalanobis distance is equivalent to the Euclidean distance if the covariance matrix is the identity matrix. Mahalanobis distance is very useful to determine the similarity of a set of values. It is sensitive to inter-variable changes in a data set. It has been used to classify observations into different groups. For our experiment, the Mahalanobis distance between Scenes is 1.5465, which is very small. This also shows that the difference between Scenes is quite small and negligible. This is why in our following statistical data analysis we considered both Scenes together. This also followed our goal to investigate the tone mapping performance for architectural scenes.

4.3.2.2 Tone Mapping Operators

Figure 4.7 – Figure 4.11 show the result of our experiment for the main effect of the tone mapping operators. As shown in the figure, all of the attributes are highly significant.

4.3.2.3 Brightness

As it is shown in Figure 4.7 the images produced by the linear tone mapping, Pattanaik method, Ward method, and Drago method have substantially higher overall brightness. All these operators are perceived the most differently when compared with their corresponding real-world views. Note that it is shown very clearly that all global methods have stronger overall brightness than the local ones.

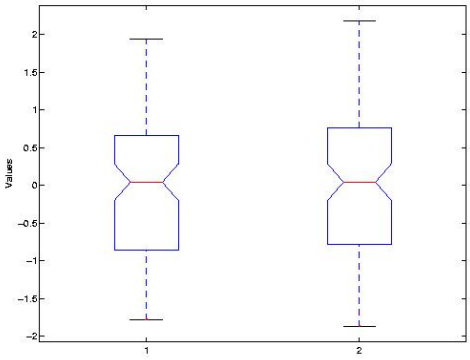

Figure 4.2: Main effect of Scenes (Scene 1 and Scene 2) for overall brightness. $F = 0.04$, $p = 0.8411$. A box shows the lower quartile, median, and upper quartile values. The whiskers are lines extending from each end of the box to show the extent of the rest of the data. Outliers are data with values beyond the ends of the whiskers.

4.3.2.4 Detail Reproductions in Bright Regions

The second most differently perceived attribute is the details in bright regions as shown in Figure 4.10. The bilateral filtering, Ashikhmin method, Reinhard method, and Drago method provide significantly more details in bright regions than the others. All of the local operators are perceived with more details than global ones according to the graph. It is obvious because local operators use different scales for small regions of an image while global operators use only one scale for whole part of an image and tend to saturate bright parts.

4.3.2.5 Contrast and Naturalness

Contrast (Figure 4.8) and naturalness (Figure 4.11) show almost same values of significance. The linear mapping, Pattanaik method, and Ward method have higher contrast and Ward, Reinhard, and Drago methods have more naturalness than the others. Global operators have stronger contrast than local ones do as shown in the graph. It corresponds to the expectations because local operators deal with small regions and local contrasts of an image while global operators deal with all pixels together and consider global contrast.

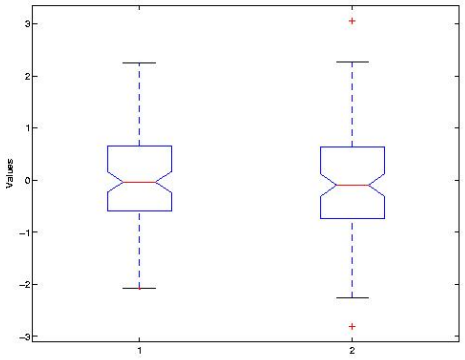

Figure 4.3: Main effect of Scenes (Scene 1 and Scene 2) for contrast. $F = 0.09$, $p = 0.7595$.

4.3.2.6 Detail Reproductions in Dark Regions

Detail reproductions in dark regions (Figure 4.9) show the least significance among the attributes, but it is still significant because its p value is much smaller than the significance level which is usually 0.05 or 0.01. The Drago method is perceived to have the most details in dark regions. The second best in dark regions reproduction is Ashikhmin method. The linear, Pattanaik, Ward, and Reinhard have almost same scores, and the bilateral filtering has slightly more details than those four.

Because the main effect of Scenes for detail reproductions in dark regions are significant, the main effect of the tone mapping operators for detail reproductions in dark regions in each scene was also tested and shown in Figure 4.12 and Figure 4.13. In this case, values were normalized over each attribute of each subject for each scene. They are perceived more differently in Scene 1 than in Scene 2. For both of Scenes, the Ashikhmin and Drago methods are perceived as the two most detailed ones in dark regions. Drago is perceived as the most detailed reproduction in dark regions in Scene 1 and Ashikhmin is perceived as the most in Scene 2. An interesting result from this observation is that although the detail reproductions in bright regions are perceived very differently, the details in dark regions are not perceived as significantly different as in bright regions when tone mapped images are compared with their corresponding real-world views.

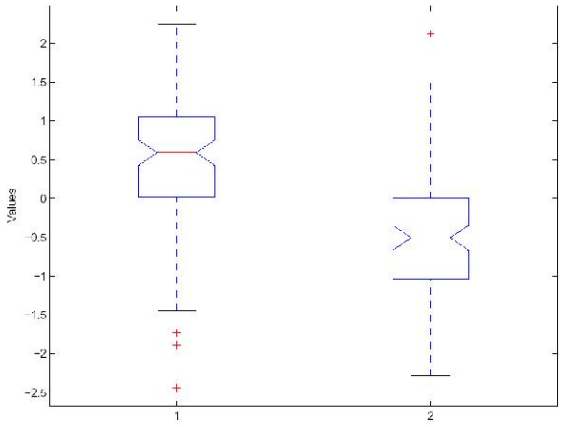

Figure 4.4: Main effect of Scenes (Scene 1 and Scene 2) for detail reproductions in dark regions. $F = 58.98$, $p = 7.67053E - 13$.

4.3.3 Correlations

Correlations of the attributes were calculated by Pearson r correlation coefficients which is the average cross product of standardized variable scores [Pea1896]. Table 4.1 shows all of the values of correlations.

	O.C.	**D.D.**	**D.B.**	**N.**
O.B.	0.3321	-0.1023	-0.55880	-0.0844
O.C.		-0.1114	-0.28220	0.2440
D.D.			0.26340	0.0729
D.B.				0.3028

Table 4.1: Correlations with Pearson r values of all pairs of overall brightness (O.B.), overall contrast (O.C.), details in dark regions (D.D.), details in bright regions (D.B.), and naturalness (N.).

This result shows that the naturalness and each of the overall brightness and details in dark regions have no correlations. On the other hand, the naturalness has slight correlations with the contrast and detail reproduction in bright regions, but they are still small. It can be concluded from this

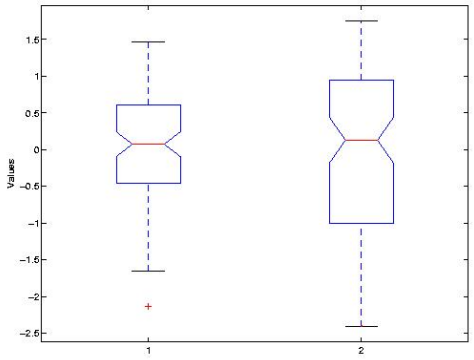

Figure 4.5: Main effect of Scenes (Scene 1 and Scene 2) for detail reproductions in bright regions. $F = 3.2204E-05$, $p = 0.9955$.

result that none of the image appearance attributes has strong influence to determine naturalness by itself and that naturalness is influenced by a combination of the image appearance attributes.

The biggest absolute value of correlation coefficient was obtained between overall brightness and detail reproductions in bright regions with Pearson $r = -0.55880$. It is obvious because as overall brightness decreases, bright parts are less saturated and better visible.

4.3.4 Dimension Estimate and Mahalanobis Distances

MANOVA in MATLAB provides an estimate of the dimension of the space containing the group means and the significant values for each of the dimensions. MANOVA provides the estimate of the dimension d. If value $d = 0$, it indicates that the means are at the same point. Value $d = 1$ indicates that the means are different but along a line, value $d = 2$ shows that the means are on a plane but not along a line and similarly for higher values of d. The null hypotheses are tested by calculating the significant values (p-values) in each of the dimensions such as the means are in N-dimensional space where N is the number of dimensions. From our set of data, MANOVA returns $d = 3$ which indicates that the means are neither along a line nor on a plane but in a three-dimensional space. The p values for each of the dimensions in our perceptual experiment are $p(d = 0) = 0.0000$, $p(d = 1) = 0.0000$, $p(d = 2) = 0.0017$, $p(d = 3) = 0.9397$, and $p(d = 4) = 0.8833$. They show that there is no possibility at all to have all means either on a

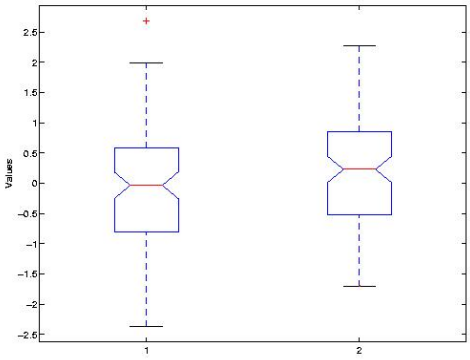

Figure 4.6: Main effect of Scenes (Scene 1 and Scene 2) for naturalness. $F = 2.85$, $p = 0.0928$.

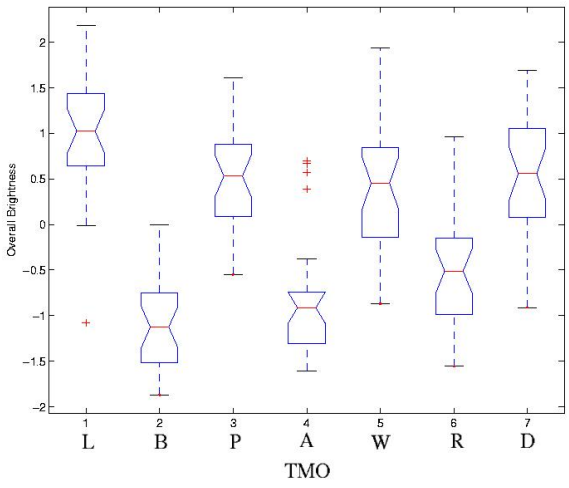

Figure 4.7: Main effects of the tone mapping operators for overall brightness. $F = 46.14$, $p = 0.0$.

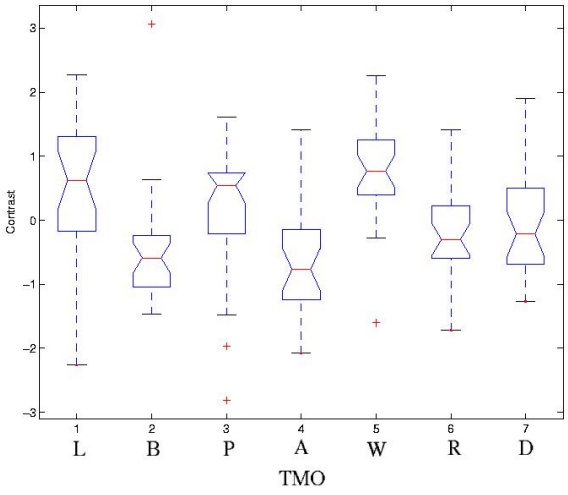

Figure 4.8: Main effects of the tone mapping operators for contrast. $F = 8.74$, $p = 2.1058E-08$.

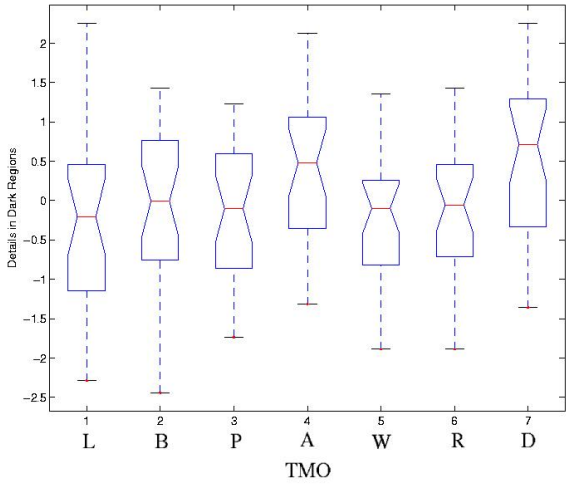

Figure 4.9: Main effects of the tone mapping operators for detail reproduction in dark regions. $F = 3.18$, $p = 0.0054$.

point or along a line. The possibility to have means on a plane is not zero, but it is much smaller than the significance level and it cannot be ignored. For value $d = 3$, the possibility becomes more than the significance level; therefore, the means are located in a three-dimensional space.

Table 4.2 shows the Mahalanobis distances among the tone mapping operators given by MANOVA. According to Table 4.2, the linear tone mapping and bilateral filtering are perceived the most differently when compared with their corresponding real-world views. The second and the third most different combinations come from the combination of the linear tone mapping and Ashikhmin method and of the linear tone mapping and Reinhard method. All of the three biggest differences are provided from the linear tone mapping. On the other hand, the least difference is provided between bilateral filtering and Ashikhmin method. This result is visualized in Figure 4.14. An interesting result shown in Figure 4.14 is that those seven tone mapping reproductions are di-

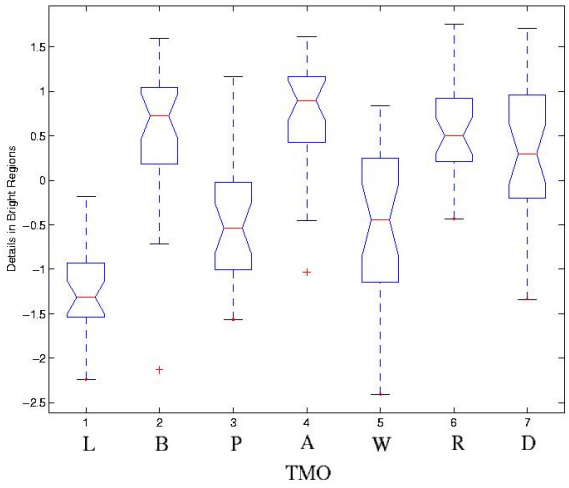

Figure 4.10: Main effects of the tone mapping operators for detail reproduction in bright regions. $F = 30.45$, $p = 0.0$.

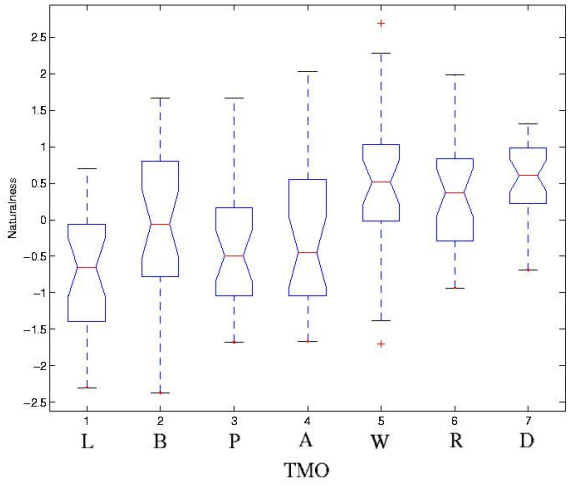

Figure 4.11: Main effects of the tone mapping operators for naturalness. $F = 8.11$, $p = 8.3877E-08$.

vided into global and local methods by Mahalanobis distances. Three local operators (bilateral, Ashikhmin, and Reinhard) are similar to each other and four global operators (linear, Pattanaik, Ward, and Drago) are similar to each other, but both categories of global and local operators are not so similar to each other.

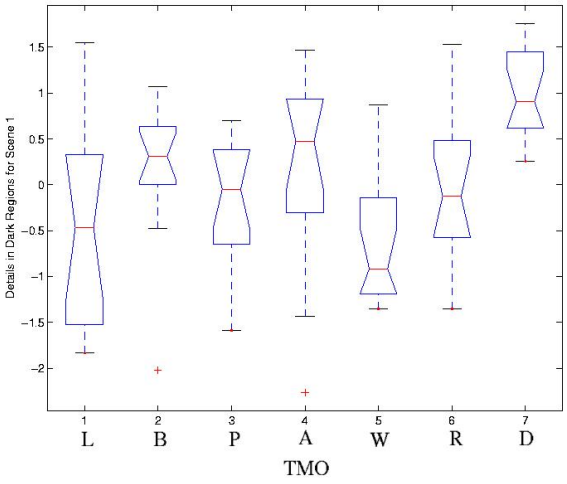

Figure 4.12: Detail reproductions in dark regions for Scene 1. $F = 5.75$, $p = 4.15714E-05$.

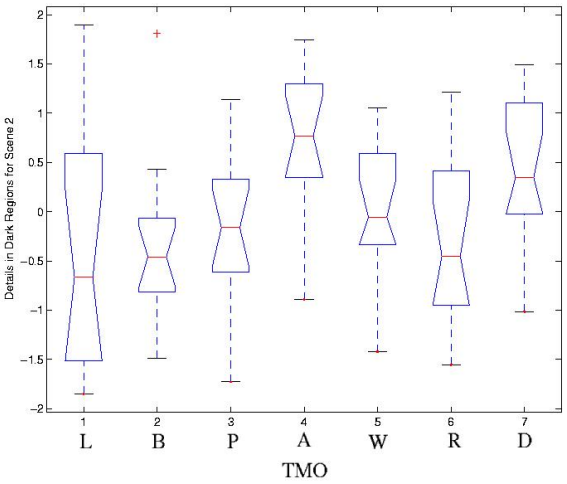

Figure 4.13: Detail reproductions in dark regions for Scene 2. $F = 3$, $p = 0.0101$.

	bilateral	**Pattanaik**	**Ashikhmin**	**Ward**	**Reinhard**	**Drago**
linear	**15.4530**	1.6541	**14.2361**	2.7122	**10.6089**	6.6940
bilateral		7.4749	0.6674	9.2726	1.3353	8.8120
Pattanaik			6.4395	1.1613	3.9887	2.8066
Ashikhmin				8.9709	1.2405	6.2989
Ward					4.5301	2.9536
Reinhard						3.7406

Table 4.2: Mahalanobis distances. The three biggest distances are written in a bold font and the three smallest distances are underlined.

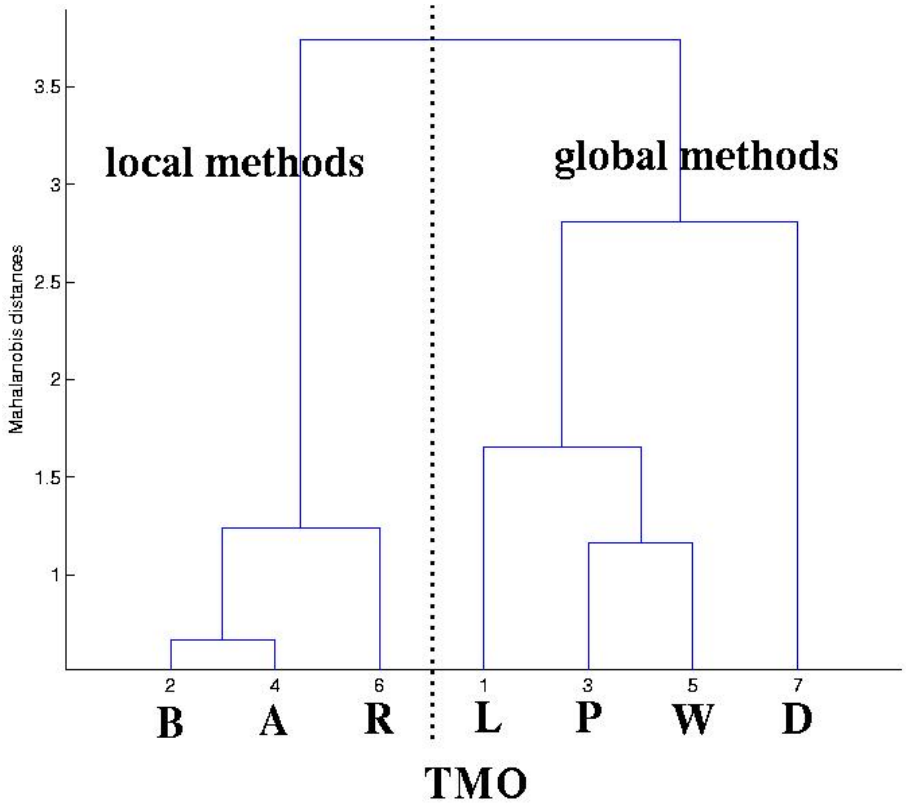

Figure 4.14: Mahalanobis distances. Note that the tone mapping operators are divided into global and local categories by the Mahalanobis distances.

Chapter 5

Conclusions

A perceptual evaluation was done over seven tone mapping operators in two scenes with 14 human subjects. The primary interest of this experiment was to investigate how differently the tone mapping operators are perceived when they are compared with their corresponding real-world views. The linear mapping, the Pattanaik method, the Ward method, and the Drago method were selected from the global tone reproduction group. The fast bilateral filtering, the Ashikhmin method, and the Reinhard method were selected from the local tone mapping group.

In the perceptual experiment, subjects were asked to compare each of the images to its corresponding real-world view and rate its appearance and realism. The image appearance attributes are overall brightness, contrast, details in dark regions, and details in bright regions. The realism rating is naturalness of an image. The subjects rated how much image appearance an image had and how real it was when compared with its corresponding real view. The set of data was analyzed by using a multivariate analysis of variance for the main effect of the tone mapping operators.

The result of the analysis shows that those seven tone mapping operators were perceived very differently in terms of all of the attributes when compared to their corresponding real-world views. Overall brightness shows the most significant differences among the tone reproductions and global operators have more brightness than local ones. The second most differently perceived attribute is detail reproduction in bright regions. In contrast to overall brightness, local operators are perceived with more details in bright regions than global ones. The Ashikhmin method is perceived as the most detailed tone reproduction in bright regions. Contrast and naturalness show almost the same significances. Global operators have more contrast than local ones, but the difference is not as large as for overall brightness. The Ward method shows the strongest contrast reproduction. The Drago method is perceived as the most natural. The least significant attribute among those five is detail reproduction in dark regions, but it is still significant enough. The Drago method is perceived as the most detailed one in dark regions.

Because there are two types of attributes, correlations between naturalness and each of the image appearance attributes were tested. The result shows that none of the image appearance attributes has a strong influence on the perception of naturalness by itself. This may suggest that naturalness is dependent on a combination of the other attributes. All of other possible pairs between attributes were also tested. The biggest correlation happens between overall brightness and details in bright regions.

The multivariate analysis of variance (MANOVA) shows that the means of the set of data are located in a three-dimensional space but neither on a point, along a line, nor on a plane. In terms

of the Mahalanobis distances, the biggest differences are between linear tone mapping and each of the fast bilateral filtering, the Ashikhmin method, and the photographic tone reproduction. The least differences are between fast bilateral filtering and Ashikhmin, between Pattanaik and Ward, and between Ashikhmin and the photographic reproduction. The analysis shows that those tone mapping operators are divided into global and local categories by the Mahalanobis distances.

Appendix A

Abbreviations

HDR High Dynamic Range
HVS Human Visual Systems
TMO Tone Mapping Operator
TVI Threshold Versus Intensity
Tone mapping operators:

- A: the Ashikhmin method [Ash02]
- B: fast bilateral filtering by Durand and Dorsey [DD02]
- D: the Drago method [DMAC03]
- L: linear mapping
- P: the Pattanaik method [PTYG00]
- R: the Reinhard method [RSSF02]
- W: the Ward method [WLRP97]

Appendix B

Images for the Perceptual Experiment

B.1 Tone Mapped Images of Scene 1

Figure B.1 – Figure B.7 show the tone mapped images of Scene 1.

Figure B.1: Linear.

Figure B.2: Bilateral filtering.

Figure B.3: Pattanaik.

Figure B.4: Ashikhmin.

Figure B.5: Ward.

B.2 Tone Mapped Images of Scene 2

Figure B.8 – Figure B.14 show the tone mapped images of Scene 2.

Figure B.6: Reinhard.

Figure B.7: Drago.

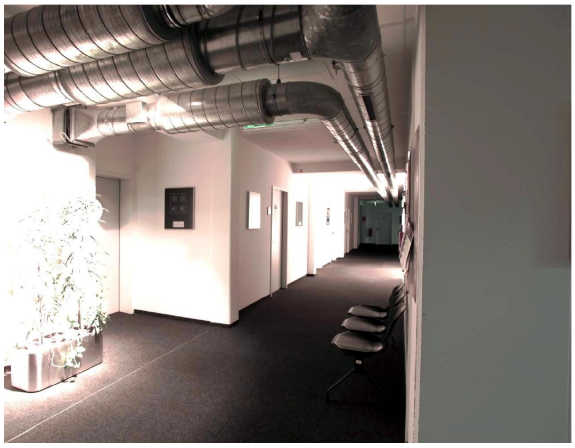

Figure B.8: Linear.

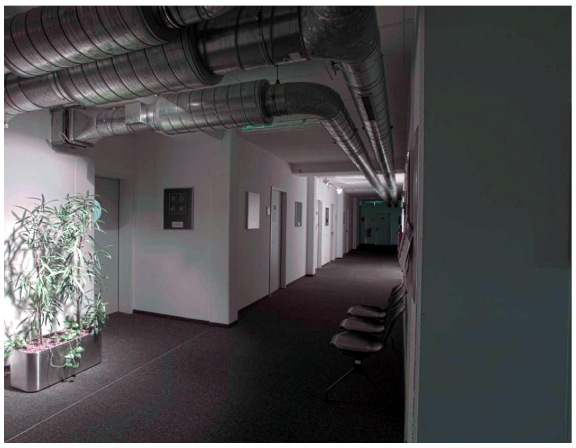

Figure B.9: Bilateral filtering.

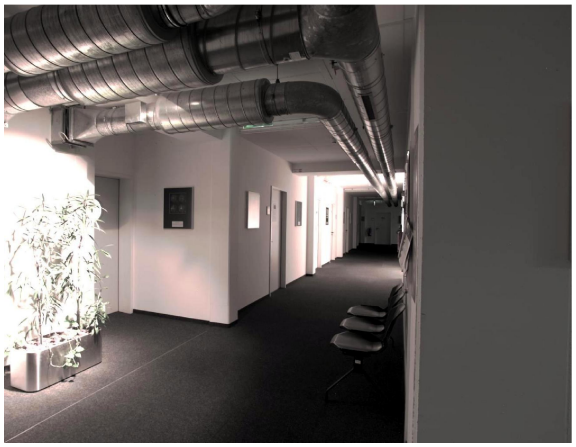

Figure B.10: Pattanaik.

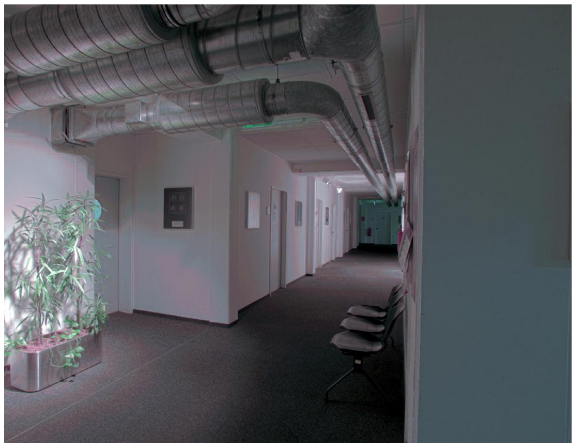

Figure B.11: Ashikhmin.

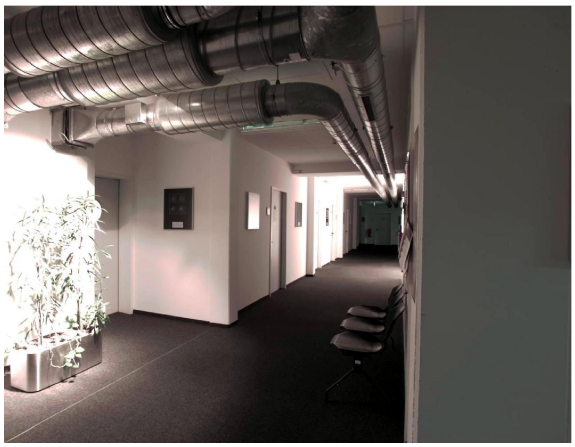

Figure B.12: Ward.

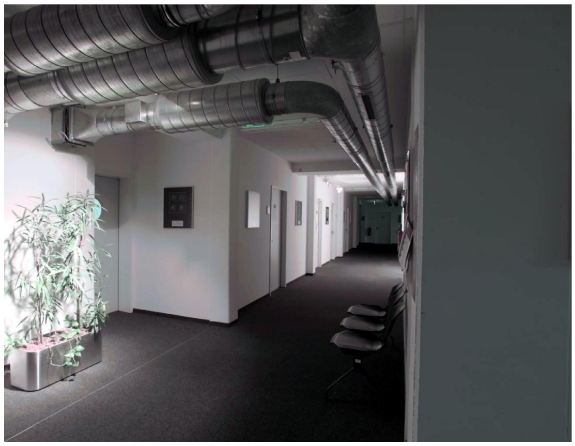

Figure B.13: Reinhard.

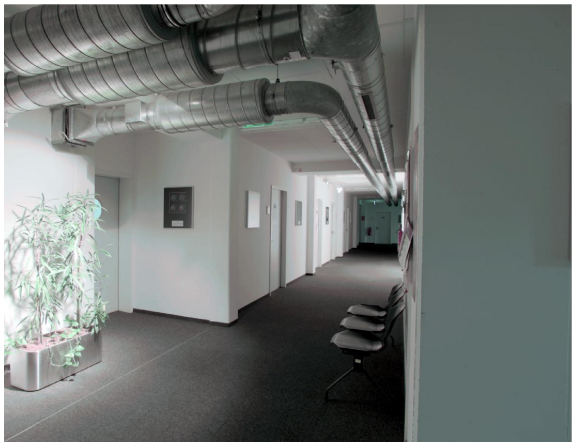

Figure B.14: Drago.

Appendix C

Values of the Perceptual Experiment

Figure C.1 – Figure C.5 show the values of our perceptual experiment. The values were normalized as $x_i \rightarrow \frac{x_i - \mu_x}{\sigma_x}$ where x_i is a score and μ_x and σ_x are respectively the mean and the standard deviation over an attribute of a subject. The normalization was done over each attribute of each subject.
The tone mapping operators and attributes are numbered as follows:

1. the linear mapping
2. fast bilteral filtering by Durand and Dorsey
3. the Pattanaik method
4. the Ashikhmin method
5. the Ward method
6. the Reinhard method
7. the Drago method

- O.B.: overall brightness
- O.C.: overall contrast
- D.D.: detail reproductions in dark regions
- D.B.: detail reproductions in bright regions
- N.: naturalness.

Those values are normalized over each attribute of each subject.

Subject	Scenes	TMO	O.B.	O.C.	D.D.	D.B.	N.
1	1	1	1,4381	-0,0345	0,8633	-1,0362	-1,4284
		2	-0,0052	-0,2758	0,818	-0,3637	0,7631
		3	0,0309	-1,9651	1,2254	-0,0275	-1,5197
		4	-1,3402	-0,0345	-0,8568	-0,1716	0,5805
		5	0,5721	-1,6031	-0,7663	-0,4118	-0,3326
		6	-1,5567	0,5688	0,8633	0,549	0,4892
		7	-0,6185	1,8962	1,678	1,3657	0,7631
	2	1	1,0051	0,6292	0,818	-1,0362	0,4892
		2	-0,4742	0,6292	-0,721	0,4529	1,037
		3	0,5721	-0,1551	-1,1736	-1,5647	-0,5152
		4	-1,4123	0,6292	-0,9473	1,2215	-0,5152
		5	1,4381	0,6895	0,0485	-1,2284	-1,7023
		6	0,5721	-0,0345	-0,8115	0,8372	0,5805
		7	-0,2216	-0,9395	-1,0379	1,4137	1,311

Subject	Scenes	TMO	O.B.	O.C.	D.D.	D.B.	N.
2	1	1	1,2037	-0,7615	0,3215	-0,5916	-0,156
		2	-0,7368	-0,8697	0,5743	0,7215	0,4321
		3	0,0802	0,6997	0,5743	-0,6926	-0,24
		4	-0,9921	0,4832	0,0686	-0,0866	-1,1642
		5	0,2334	0,6455	-0,5382	0,368	0,18
		6	0,2845	0,8079	-0,4371	-0,0361	-0,24
		7	1,1526	0,8079	1,5352	-0,7431	-0,4081
	2	1	1,3058	1,5114	-2,1565	-1,8542	-1,6683
		2	-1,5538	-1,465	-1,2968	1,4285	0,8522
		3	0,2334	0,375	-1,0439	-0,9956	-0,8282
		4	-1,6049	-1,6274	0,8272	1,075	-0,4921
		5	0,3356	0,9161	0,0686	-0,8441	2,2804
		6	-0,89	-0,7615	0,4732	0,9235	0,264
		7	0,9483	-0,7615	1,0295	1,3275	1,1882

Subject	Scenes	TMO	O.B.	O.C.	D.D.	D.B.	N.
3	1	1	-1,0796	-1,0128	-1,1109	-1,4478	-1,3322
		2	-1,3626	-0,7234	1,0796	0,9164	0,5153
		3	0,279	0,3762	-0,0156	-0,8489	-1,5081
		4	-1,1928	-1,1864	0,6415	0,9479	-0,5844
		5	-0,4003	0,7813	-0,6728	0,3175	0,0314
		6	-1,1928	-0,0868	1,0796	1,0425	0,8232
		7	0,279	-0,9549	1,2987	0,5697	0,6472
	2	1	1,2413	-1,36	1,2987	-1,5108	-0,8043
		2	-0,2305	-0,3183	-0,4537	0,7903	-0,4524
		3	1,1281	1,0706	-0,2347	-1,3532	-0,6724
		4	-0,5701	-0,3183	0,2034	0,9794	-0,5404
		5	0,6186	1,6494	-1,7681	-1,0065	1,527
		6	0,9583	1,4179	-1,3299	0,4436	1,7029
		7	1,5244	0,6655	-0,0156	0,1599	0,6472

Figure C.1: Values of the perceptual experiment for the subjects 1 – 3.

Subject	Scenes	TMO	O.B.	O.C.	D.D.	D.B.	N.
4	1	1	0,2255	0,3443	0,8578	-0,6595	-0,3392
		2	-0,469	-0,5209	1,4336	1,0442	-0,0424
		3	0,2886	0,5297	-0,2938	0,1649	-0,8586
		4	-0,8478	-0,8299	1,0223	0,6045	0,0318
		5	0,6674	-0,2737	-0,047	0,2198	0,848
		6	-0,0902	-0,2119	1,1868	-0,3847	0,4028
		7	-0,0902	0,2825	1,0223	-0,4397	0,477
	2	1	2,1827	-0,2737	-1,7743	-1,9785	-1,3779
		2	-1,0372	3,0637	-0,7873	1,2091	0,9964
		3	0,9831	-1,0771	-1,1986	-1,4289	-1,3037
		4	-1,5423	-0,7063	0,4465	1,4289	2,0351
		5	1,1094	-0,0265	-0,7873	-0,7694	-1,3779
		6	-0,469	0,2207	-0,4583	0,2198	-0,1166
		7	-0,9109	-0,5209	-0,6228	0,7694	0,6254

Subject	Scenes	TMO	O.B.	O.C.	D.D.	D.B.	N.
5	1	1	1,0135	1,0979	2,2494	-1,6589	-2,2984
		2	-1,042	0,4205	1,0163	0,3927	-0,0707
		3	1,6103	1,6189	-1,7345	-0,592	-0,5658
		4	-0,9093	-0,2568	-0,4065	-0,0176	0,5893
		5	1,9418	0,3684	-1,4499	-1,3306	-0,9783
		6	-1,1083	-0,4131	-0,4065	0,2286	0,6718
		7	-0,18	-0,2047	0,9214	-0,4279	0,6718
	2	1	0,8146	0,0558	0,1626	-0,51	-1,4733
		2	-0,7767	-0,2047	-0,5014	1,0492	0,8369
		3	-0,3126	-2,8098	-0,122	-0,0176	0,0943
		4	-0,7767	-0,3089	-0,3117	0,8851	0,6718
		5	-0,18	0,4205	0,1626	-1,0024	-0,0707
		6	-0,5115	-0,3089	-0,0271	1,2954	0,9194
		7	0,4168	0,5247	0,4472	1,7058	1,0019

Subject	Scenes	TMO	O.B.	O.C.	D.D.	D.B.	N.
6	1	1	1,6033	1,5824	1,2272	-1,4613	-0,4852
		2	-1,6673	-1,1181	-2,4406	-2,132	-2,3682
		3	0,528	1,5366	0,7446	0,9822	0,0528
		4	0,5728	-2,0793	1,6133	-1,0301	-1,4267
		5	0,304	0,0719	0,0207	0,2635	1,0616
		6	0,6176	-0,5689	-0,0276	0,551	0,6581
		7	1,6929	0,85	1,2272	1,4613	-0,6869
	2	1	-0,0096	0,6212	-0,4137	-0,2635	-0,2162
		2	-0,7264	-0,34	-0,4137	0,2635	-0,1489
		3	-0,5472	0,5754	-0,3171	-0,2635	1,0616
		4	-1,3089	-0,5231	-0,2689	0,7905	0,4563
		5	-0,3232	0,2092	-0,2206	0,2635	0,9944
		6	-0,9953	-0,2942	-0,5102	-0,3114	0,2546
		7	0,2592	-0,5231	-0,2206	0,8863	0,7926

Figure C.2: Values of the perceptual experiment for the subjects 4 – 6.

Subject	Scenes	TMO	O.B.	O.C.	D.D.	D.B.	N.
7	1	1	1,7087	0,0148	0,1244	-1,3747	0,6879
		2	-1,6431	-0,1407	-0,2566	0,9909	-1,3838
		3	0,6986	0,6371	-0,2022	-0,3723	-0,208
		4	-0,8626	-1,4371	0,9954	0,7904	-1,6637
		5	0,1476	1,4149	-1,8897	-0,1718	0,4639
		6	-0,4952	0,2222	-1,8897	0,3093	-0,4879
		7	0,0558	-0,1926	0,5055	-0,2119	1,2478
	2	1	0,6067	1,2593	-0,6377	-1,2945	0,5199
		2	-1,2758	-1,0223	0,0156	1,5923	-1,4957
		3	1,1577	0,6371	0,7232	-0,6931	1,1918
		4	-1,3217	-1,5408	1,1587	1,1112	0,5199
		5	0,6067	1	-0,0933	-1,535	0,8559
		6	-0,0361	-1,3334	-0,0389	0,9508	-0,7679
		7	0,6527	0,4815	1,4854	-0,0916	0,5199

Subject	Scenes	TMO	O.B.	O.C.	D.D.	D.B.	N.
8	1	1	1,4268	1,5673	-0,9006	-1,5595	0,7029
		2	-1,3512	-0,5561	0,9864	0,6443	-0,4521
		3	0,5669	0,6067	0,8148	-0,4832	-0,4983
		4	0,6992	-1,6178	0,6433	0,388	-1,5147
		5	0,2362	0,6067	0,2144	0,4393	2,6896
		6	-1,2189	-0,3539	1,0293	1,4643	0,2871
		7	0,4347	-0,4045	2,0157	0,4905	0,1023
	2	1	0,5669	1,0111	-0,5575	-0,9957	0,1947
		2	-1,0866	-1,2639	-1,1151	0,9518	-0,6369
		3	0,1039	0,4045	-0,0858	-1,252	0,2409
		4	-1,5497	-1,1123	0	1,003	-1,2375
		5	1,5591	1,5673	-0,9006	-1,4057	0,5181
		6	-0,4914	-0,3033	-1,4582	-0,432	-0,3597
		7	0,1039	-0,1517	-0,6862	0,7468	-0,0363

Subject	Scenes	TMO	O.B.	O.C.	D.D.	D.B.	N.
9	1	1	1,6757	1,1197	-0,4732	-0,9135	0,1036
		2	-1,5509	-0,3949	0,1679	-0,1994	0,1761
		3	0,8691	-0,3191	0,0076	-0,7836	-1,4914
		4	0,6674	-1,3415	1,5037	0,8393	-0,7664
		5	-0,7442	1,1954	0,4885	0,8393	1,191
		6	-0,7442	0,6275	0,3816	0,7094	1,9885
		7	1,2052	-0,3191	2,2517	0,255	0,1761
	2	1	0,8691	0,5896	0,0076	-0,459	0,0311
		2	-1,0131	-0,7356	-0,9007	-0,7187	-1,4914
		3	-0,1392	0,7032	-1,7021	1,1638	0,901
		4	-0,9459	-1,19	-0,0458	1,2288	-1,0564
		5	-0,3409	1,3469	-0,0992	-2,4065	0,6835
		6	-0,677	-1,7201	-0,9541	0,3849	-0,3314
		7	0,8691	0,4381	-0,6335	0,0603	-0,1139

Figure C.3: Values of the perceptual experiment for the subjects 7 – 9.

Subject	Scenes	TMO	O.B.	O.C.	D.D.	D.B.	N.
10	1	1	0,3485	-0,0806	0,027	-0,9396	-0,7983
		2	-1,1658	-1,0612	0,7832	-0,1323	-0,1409
		3	-0,2404	-1,2083	0,7292	0,865	-1,6748
		4	-0,7452	0,753	1,4854	1,3399	1,5027
		5	-0,8713	0,753	0,8913	-0,2748	0,407
		6	-1,3341	-0,3257	0,8372	1,1499	-0,7983
		7	1,4422	-0,2277	1,1073	-0,1798	0,9548
	2	1	1,274	2,273	-1,2154	-1,652	-1,3461
		2	-1,2079	-1,0122	-1,5935	0,7225	-0,2504
		3	0,8954	0,5569	-0,9993	-0,9396	-0,36
		4	0,3906	-0,4238	-0,9993	1,2924	1,0644
		5	1,1057	1,2433	-0,9453	-1,5095	0,8452
		6	-0,661	-0,3747	-0,135	0,2001	-0,5791
		7	0,7692	-0,8651	0,027	0,0577	1,1739

Subject	Scenes	TMO	O.B.	O.C.	D.D.	D.B.	N.
11	1	1	0,5541	-0,5902	-0,5105	-1,0272	-1,5481
		2	-1,0848	0,1914	0,699	-0,3424	0,4297
		3	0,2263	-0,2552	0,2592	-1,2554	-1,2184
		4	-1,0848	1,4196	1,1388	-0,4565	-0,9712
		5	-0,1015	0,5264	1,3588	0,7989	0,1825
		6	-0,2107	1,4196	0,2592	0,4565	0,3473
		7	0,0078	-0,2552	1,6886	0,3424	0,3473
	2	1	1,9745	-2,2649	-1,94	-2,0544	-1,5481
		2	-0,757	0,5264	-1,0603	1,3696	1,6658
		3	0,5541	-1,4834	-0,7304	-0,2283	0,3473
		4	-1,3033	0,0798	-0,4006	0,9131	1,0065
		5	1,5375	-0,2552	-0,7304	-0,4565	-0,3944
		6	-0,9756	0,3031	-0,0707	0,9131	1,0889
		7	0,6634	0,638	0,0393	1,0272	0,2649

Subject	Scenes	TMO	O.B.	O.C.	D.D.	D.B.	N.
12	1	1	0,7266	1,4301	0,3059	-1,3767	-0,2107
		2	-1,4884	-0,6206	0,3708	0,7605	-1,7156
		3	0,7266	0,8095	0,3708	-0,4441	-0,2709
		4	-0,3809	-1,2413	-1,3164	0,6051	-1,6554
		5	-0,3399	0,8365	0,3708	-0,4441	1,3544
		6	-0,4219	0,1079	0,4357	0,5662	1,2942
		7	1,67	-1,2682	0,3708	-1,3378	-0,2709
	2	1	1,67	1,4301	-1,3813	-1,3378	-0,2709
		2	-1,4473	-0,6206	0,4357	1,5765	1,3544
		3	-0,3809	0,7825	0,3059	-0,4441	0,4515
		4	-0,8731	-1,2413	2,123	1,6154	-0,2107
		5	0,7266	0,8365	0,3059	-0,4441	0,5117
		6	-0,3809	-0,6206	-1,3813	0,6051	-0,9331
		7	0,1934	-0,6206	-1,3164	0,0999	0,5719

Figure C.4: Values of the perceptual experiment for the subjects 10 – 12.

Subject	Scenes	TMO	O.B.	O.C.	D.D.	D.B.	N.
13	1	1	0,6724	0,6662	0,3963	-0,1804	-0,5019
		2	-0,6211	-0,5083	0,7542	0,642	1,203
		3	0,888	0,3208	0,9928	0,7595	-1,3284
		4	-1,0523	-0,8538	1,0525	1,0532	-1,0185
		5	0,1694	2,2554	0,4559	0,2308	1,6679
		6	-0,5492	-0,7847	0,3366	0,2895	0,8413
		7	1,1755	0,1135	1,2911	1,347	0,583
	2	1	1,3192	-0,3701	-2,2881	-2,2365	-1,2768
		2	-1,7709	-1,2683	-0,5582	0,642	-0,9152
		3	0,0975	0,459	-1,0354	-1,4141	-0,7085
		4	-0,7648	-1,1992	0,5156	0,4658	-0,3985
		5	0,9599	1,2881	-0,8565	-1,0616	0,5314
		6	-1,2679	-0,6465	-0,6178	-0,1804	0,2214
		7	0,7443	0,5281	-0,4389	-0,3567	1,0997

Subject	Scenes	TMO	O.B.	O.C.	D.D.	D.B.	N.
14	1	1	1,0336	0,8252	0,5163	-1,1232	-1,0682
		2	-1,7787	-1,1003	1,1111	0,1028	-1,8247
		3	1,1696	0,9697	0,6078	0,2094	-0,4863
		4	-0,9169	0,2957	1,0653	0,0495	-0,0208
		5	0,6253	0,4401	0,4706	-0,1104	0,0956
		6	-0,5087	-1,4854	1,4313	0,1561	0,6775
		7	0,58	-0,9078	0,9738	-0,2703	1,2594
	2	1	0,6707	1,3548	-1,1764	-1,4964	-1,4173
		2	-1,8695	-1,1966	-0,0327	1,0623	0,0956
		3	0,5346	0,9697	-0,5817	-1,0166	1,6667
		4	-0,7355	-0,5708	-0,7189	1,4888	-0,1954
		5	0,716	1,2585	-1,0392	-1,2832	-0,3117
		6	-0,0551	0,0069	-1,2679	1,7553	1,0266
		7	0,5346	-0,8596	-1,3594	0,476	0,5029

Figure C.5: Values of the perceptual experiment for the subjects 13 – 14.

Bibliography

[WarXX] G. WARD. *High Dynamic Range image Encodings, Anyhere Software Corp*. http://www.anyhere.com/gward/hdrenc/hdr_encodings.html ().

[MP95] S. MANN AND R.W. PICARD. *On being 'undigital' with digital cameras: extending dynamic range by combining differently exposed pictures*. IS ampersand T's 48th Annual Conference, pages 422-428, (1995).

[DM97] P.E. DEBEVEC AND J. MALIK. *Recovering High Dynamic Range Radiance Maps from Photographs*. T. Whitted, editor, Proceedings of ACM SIGGRAPH, pages 369-378, (1997).

[MN99] T. MITSUNAGA AND S.K. NAYAR. *Radiometric self calibration*. Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, (1999).

[RB99] M.A. ROBERTSON, S. BORMAN, AND R.L. STEVENSON. *Dynamic range improvement through multiple exposures*. IEEE International Conference on Image Processing, (1999).

[War03] G. WARD. *Fast, Robust Image Registration for Compositing High Dynamic Range Photographcs from Hand-Held Exposures*. Journal of Graphics Tools, 8(2):17-30, (2003).

[TR93] J. TUMBLIN AND H.E. RUSHMEIER. *Tone Reproduction for Realistic Images*. IEEE Computer Graphics and Applications, 13(6):42-48, November, (1993).

[WLRP97] G.W. LARSON, H. RUSHMEIER, AND C. PIATKO. *A Visibility Matching Tone Reproduction Operator for High Dynamic Range Scenes*. IEEE Transactions on Visualization and Computer Graphics, 3(4):291-306, (1997).

[War94] G. WARD. *A Contrast-Based Scalefactor for Luminance Display*. P.S. Heckbert, editor, Graphics Gems IV, pages 415-421, Academic Press, (1994).

[FPSG96] J.A. FERWERDA, S. PATTANAIK, P. SHIRLEY, AND D.P. GREENBERG. *A Model of Visual Adaptation for Realistic Image Synthesis*. Proceedings of ACM SIGGRAPH, pages 249-258, (1996).

[PTYG00] S.N. PATTANAIK, J. E. TUMBLIN, H. YEE, AND D.P. GREENBERG. *Time-dependent visual adaptation for realistic image display*. Proceedings of ACM SIGGRAPH, pages 47-54, (2000).

[DMAC03] F. DRAGO, K. MYSZKOWSKI, T. ANNEN, AND N. CHIBA. *Adaptive logarithmic mapping for displaying high contrast scenes*. P. Brunet and D. Fellner, editors, Proceedings of Eurographics, (2003).

[RSSF02] E. REINHARD, M. STARK, P. SHIRLEY, AND J. FERWERDA. *Photographic Tone Reproduction for Digital Images*. Proceedings of ACM SIGGRAPH, (2002).

[Ash02] M. ASHIKHMIN. *A Tone Mapping Algorithm for High Contrast Images*. P. Debevec and S. Gibson, editors, 13th Eurographics Workshop on Rendering, pages 145-155, (2002).

[DD02] F. DURAND AND J. DORSEY. *Fast Bilateral Filtering for the Display of High-Dynamic-Range Images*. Proceedings of ACM SIGGRAPH, (2002).

[Mat] MATHWORKS, INC. *MATLAB*. http://www.mathworks.com/, ().

[Tab89] B.G. TABACHNICK. *Using multivariate statistics*. Harper Collins Publishers, Inc., 2nd edition, ISBN 0-06-046571-9, (1989).

[Mah36] P.C. MAHALANOBIS. *On the Generalised Distance in Statistics*. Proceedings of the National institute of Science in India, A2:49 – 55 (1936).

[Pea1896] K. PEARSON. *Regression, heredity, and panmixia*. Philosophical Transactions of the Royal Society of London, A(187):253 – 318 (1896).

[GKF+99] A. GILCHRIST, C. KOSSYFIDIS, F. BONATO, T. AGOSTINI, J. CATALIOTTI, X. LI, B. SPEHAR, V. ANNAN, AND E. ECONOMOU. *An Anchoring Theory of Lightness Perception*. Psychological Review, 106(4):795-834 (1999).