FAKULTÄT
FÜR !NFORMATIK
Faculty of Informatics

Spectral data representation techniques for realistic image synthesis

DIPLOMARBEIT

zur Erlangung des akademischen Grades

Diplom-Ingenieur

im Rahmen des Studiums

Computergraphik/Digitale Bildverarbeitung

eingereicht von

Thomas Kment
Matrikelnummer 0109724

an der
Fakultät für Informatik der Technischen Universität Wien

Betreuung
Betreuer/in: Prof. Dr.Dr.h.c. Werner Purgathofer
Mitwirkung: Dipl.-Ing. Mag.rer.soc.oec. Dr.techn. Andrea Weidlich

Wien, 07.02.2011

(Unterschrift Verfasser)

(Unterschrift Betreuer/in)

Eidesstattliche Erklärung

Hiermit erkläre ich an Eides statt, dass ich die vorliegende Arbeit selbständig und ohne fremde Hilfe verfasst, andere als die angegebenen Quellen und Hilfsmittel nicht benutzt und die aus anderen Quellen entnommenen Stellen als solche gekennzeichnet habe.

Korneuburg, am 07.02.2011

Thomas Kment

Abstract

Digital image synthesis is used in many industrial applications. One goal of this discipline is the physical correct simulation of natural phenomena such as the color of the blue sky or refraction in translucent materials. To accomplish this goal the usage of spectral data for the representation of color information is indispensable. For this purpose different methods were developed over the years. These methods face the challenge of accurately representing complex data (e.g. fluorescent light sources). My work compares these methods using the color difference formula CIEDE2000 as well as a large data set for testing (NCS atlas), and answers the question whether the application of low pass filters during sampling rate conversions does, or does not influence the accuracy of the representation methods.

Kurzfassung

Digitale Bildsynthese wird in vielen industriellen Anwendungen eingesetzt. Ein Teilbereich dieser Disziplin ist die physikalisch korrekte Darstellung natürlicher Phänomene wie beispielsweise die Farbe des blauen Himmels, oder korrekte Lichtbrechung. Die diesen Phänomenen zugrundeliegenden physikalischen Effekte erfordern, dass bei der Verarbeitung von Farbinformationen deren spektrale Eigenschaften berücksichtigt werden. Hierfür wurden über die Jahre einige Methoden entwickelt. Diese stehen vor der gemeinsamen Herausforderung, komplexe Daten (z.B. fluoreszierende Lichtquellen) möglichst genau darzustellen. Meine Arbeit vergleicht diese Verfahren mithilfe der Farbdistanzformel CIEDE2000 und eines großen Testdatensatzes (NCS Farbatlas) und geht der Frage nach, welchen Einfluss Tiefpassfilter bei der Veränderung der Abtastrate der Daten haben.

Contents

List of Figures

CHAPTER 1

Introduction

The goal of realistic image synthesis is the physical correct generation of high quality images of three dimensional scenes on the computer. Computer generated images are widely used in industrial applications and the entertainment industry. For example, they allow product designers to know how a combination of different materials will be perceived under different lighting conditions. A correct computer created prediction makes building a prototype obsolete [Johnson and Fairchild, 1999].

Correct predictions about the behavior of certain physical effects, such as scattering, are dependent on information contained in the spectral appearance of a given color or light source. Hence, the color representation technique used in an application influences the quality of the result. Traditional representation techniques of color information, like RGB values [Smith, 2001], do not contain any spectral information, hence they are not suitable for generating realistic images which involve effects dependent on this information [Peercy, 1993]. In order to simulate these effects correctly other color representation methods which preserve spectral information are needed [Johnson and Fairchild, 1999].

The term "spectral" in this context includes realistic image synthesis and spectral representations of color information [Devlin et al., 2002]. Spectral implies that color information retains the information about the spectral wavelength distribution of the color.

Spectral color information is obtained by the usage of spectrophotometers and data usually consists of 40 sample points or more, depending on the sampling rate of the device. These sample points are the scientific representation of a color and its light intensity distributed along the visible spectrum [Nassau, 2001]. The information of this distribution, usually given as an indexed list of sample values along the spectrum between $380nm$ and $720nm$ [Hunt, 1991], allows to correctly simulate spectral effects. Usually, an application does not utilize the whole number of available sample points of a given spectrum of a color or a light source. In order to reduce storage requirements and calculation costs, fewer sample points or compression methods are used.

This reduction introduces certain issues for spectral representation methods which have to be dealt with. Spectral information, especially high frequent details, may be lost or become

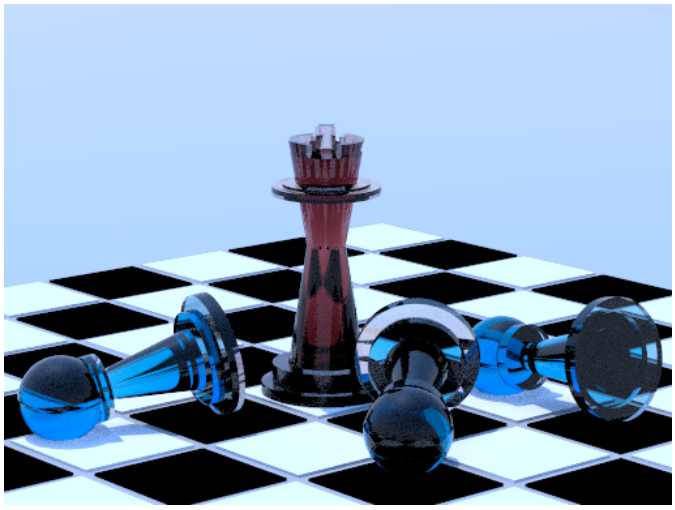

Figure 1.1: This image shows an example of realistic image synthesis. The scene is composed of a simple chessboard scene and involves light scattering (sky color) and refraction (translucent queen and pawns).

falsified [Devlin et al., 2002]. Combined with the requirement of efficiency, preserving spectral information leads to a trade-off between speed, compactness and accuracy.

All methods in literature working with spectral information have in common that they were tested against the Gretag Macbeth [McCamy et al., 1976] color checker to test their accuracy, compactness and calculation costs [Sun, 2000]. These three tested properties influence each other, have an impact on the quality of the resulting image of a given project and also influence the economic costs. For example, for a project consisting of thousands of images like a computer generated movie scene, they affect the need for storage and calculation capacity to complete the project in time. The need for storage and calculation capacity results in different operational and rental costs for used computer hardware.

1.1 Motivation

The purpose of my work is to find answers to the following questions:

Question 1: How accurate do methods for the representation of spectral information work when a large data set is used for testing?

My work's contribution to this question is a ranking of the representation methods sorted by the average color difference to reference values derived from dense sampled initial data. This reference data consists of the spectral data from the whole NCS atlas [Hunt, 1991] with its 1,930 entries, which is much larger than the normally used Gretag Macbeth color checker data set. I reviewed the representation methods under the perspectives of accuracy and compactness.

Question 2: Does the application of low-pass filters during sampling rate conversion influence the accuracy of the representation methods?

Since the usage of digital filters for preparing the spectral data has no impact on performance at application runtime, it could be a valuable tool to improve accuracy without reducing compactness or increasing calculation costs. My work gives a specific answer for each tested filter whether it has a positive impact on the accuracy of a method, and which filter is best suited for this task.

This work is structured as follows. The chapters “Color theory”, “Wavelength dependent effects” and “Digital signal processing” are a baseline review of the knowledge needed to work with the topics of this work, such as conversion formulas and physical backgrounds.

In the chapter “Rendering of wavelength dependent effects” I review literature on the topic of spectral data representation for realistic image synthesis.

The results of this work are reviewed and discussed in the chapters “Results” and “Conclusion”.

CHAPTER 2

Color theory

The human eye is able to distinguish between 10 million colors [Hunt, 1991]. Any perceived color is a visual sensation dependent on three components: the involved lighting, the illuminated objects and the observer. The same color of an object is different under different lighting conditions. For example, under artificial light the color of an object is different compared to the color of the same object under daylight [Hunt, 1991]. This difference is subjectively smaller for a human observer, because the human visual system is able to adapt to different lighting conditions. Nassau [Nassau, 2001], thus defines the term color in his marvelous book about color in this way:

> *"The term color is properly used to describe at least three different aspects of reality. First, it describes a property of an object [...] Second, it describes a characteristic of light rays [...] And third, it describes a class of sensations, as in the brain's interpretation of specific manner in which the eyes perceives light [...]"* [Nassau, 2001, p.3]

Nassau also states that these aspects are not properly distinguished in every application, yet it is important to keep their differences in mind.

The observer and his visual system both influence color perception. Therefore, color theory does not only involve material sciences, but also physiology and psychology as well [Nassau, 2001].

The question of how colors will be perceived is important in product design, architecture, dyeing, paint technology and illuminating engineering. To answer this, it is necessary to be able to measure and compare colors [Hunt, 1991]. This chapter gives an overview on the topics of color theory. It starts with the basic theory, continues with human perception and concludes with color systems.

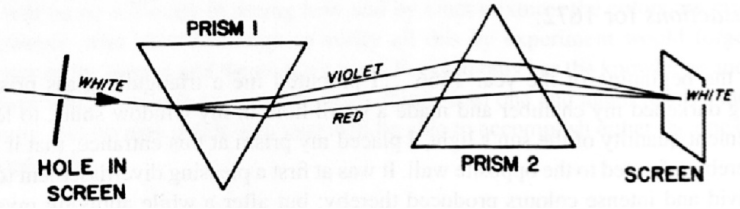

Figure 2.1: This figure shows the design of Newton's experiment which he used to prove that sunlight is a composition of all spectral colors. He split a small beam of sunlight with a prism into its spectral colors. A second prism is used to recombine the rainbow colored display of colors back into a beam of sun light. Newton's experiment also shows the concept of additive color mixing, or in other words the mixing of colored light beams [Nassau, 2001].

2.1 Color and light

Isaac Newton (1643-1727) discovered the spectral nature of light during his well known prism experiments in the year 1666 [Hunt, 1991]. Newton's prism experiments involved the redirection of a beam of sunlight through a small hole into a dark room. There, the beam hit a prism and was split into a rainbow colored display of spectral colors which Newton called a spectrum [Hunt, 1991]. In his second experiment (shown in figure 2.1), Newton added a prism which recombined these spectral colors back into sunlight. Newton concluded from this experiment that white light is a composition of all colors of the spectrum [Hunt, 1991].

After this experiments, Newton tried to further split the spectral colors with the same technique. This was realized by an additional aperture which separated one spectral color from the spectrum. The result of this experiment was that a spectral color could not be split any further by a prism [Hunt, 1991]. This property of the spectral colors, i.e. that it could not be split any further, lead Newton to label them basic colors. Newton arranged the spectral colors he discovered in a circle (depicted in figure 2.2) which is known as Newton's color wheel and grouped the colors into the (basic) spectral colors red, orange, yellow, green, blue, indigo and violet [Hunt, 1991].

Newton's experiments also introduced the concept of mixing colored light beams, today known as additive color mixing (see figure 2.3), where, the mixture of complementary colored light beams produces the color white. The absence of all light beams results in the color black. The term complementary colors is related to color ordering systems and describes two colors which are oppositionally aligned in these systems (see figure 2.11 for an example and details).

The other known type of color mixing, subtractive color mixing (see figure 2.3) works in a different way. Subtractive color mixing does not involve colored light beams, but instead colored surfaces which absorb light to produce their perceived color. There the mixture of complementary surface colors produces a color near to black. The reason for this is the fact, that for example, a green surface is perceived green because the surface absorbs all colors except green. So if different complementary colors are mixed together, every color is absorbed and only a nearly black surface remains [Nassau, 2001].

A color can be represented as an optical spectrum composed by different vibrations (intensi-

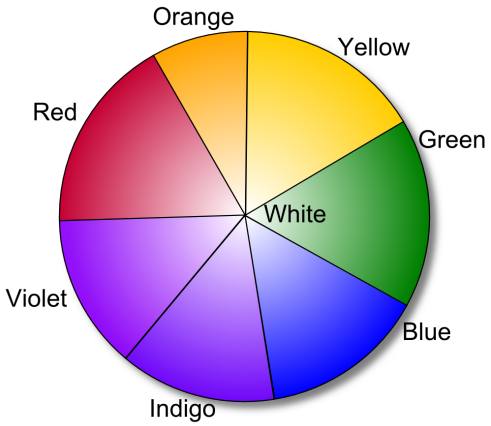

Figure 2.2: Newton's color circle shows the arrangement Newton choose for the spectral colors he discovered. The seven spectral colors are: indigo, violet, red, orange, yellow, green and blue. The number of seven relates to an analogy to seven notes in music [Nassau, 2001].

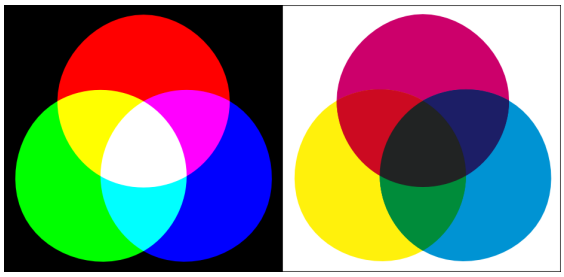

Figure 2.3: Additive color mixing (left image) is the mixing of colored light beams and results in the color white when all colors are present.
Subtractive color mixing (right image) is the mixture of surface colors, and is different to additive color mixing. A surface color results from light absorbance or, in other words, from the color which is not absorbed. When all colors are mixed together the mixture results in the absorbance of all colors and in a color near black [Nassau, 2001].

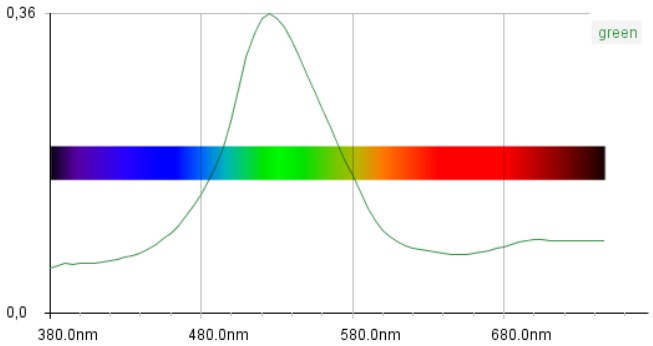

Figure 2.4: This spectral power distribution (SPD) depicts the color green and was taken from the Gretag Macbeth color checker. The illustration was extended by including the spectral distribution of visible colors. The sample spectrum has its dominant wavelength at 500nm which corresponds to the hue of the color green. The absence or presence of other dominant wavelengths influences the pureness of the color. If only one dominant wavelength is present than the resulting color features a high pureness.

ties) and their corresponding frequencies (wavelengths). Although frequencies are usually given in hertz (Hz) this unit is not used in color science, because it results in large unhandy numbers when used in context with colors. Instead nanometers (nm) are used to describe the main bands of a spectrum [Nassau, 2001].

The topic of color also involves a theoretical approach to light. Light's nature is explained by Kurt Nassau [Nassau, 2001] in his book about color:

> *"The best we can say is that light (as well as any other form of radiation) usually acts as it were a transverse electromagnetic wave [...] but that under certain circumstances it also can act as if it were a particle, [...] Just as gravity does not need a medium for its propagation, neither does electromagnetism."* [Nassau, 2001, p.25]

A very important aspect of light is that the result of any mixture of light is perceived as one single color and that some combinations of spectral distributions also produce colors which are not part of Newton's spectral colors [Nassau, 2001].

Intensities and wavelength distribution of light are measured by highly sophisticated spectrophotometers. These instruments provide spectral data with high precision (down to 5nm sampling interval and below) [Nassau, 2001].

Besides the scientific representation of light it is also important to keep in mind that light is basically a visual sensation of the (human) eye. In other words, the sensation of a color is not only determined by its wavelength composition or involved light sources, but also by its observer [Hunt, 1991].

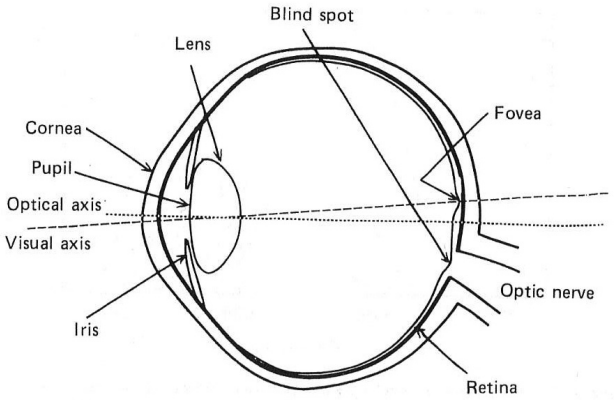

Figure 2.5: The human eye. The *cornea* which is a curved surface provides the optical power for perception. The *lens* dynamically adjusts itself depending on the viewing distance. The *iris* controls the light passing through the *pupil* and controls how much light reaches the *retina*. *Cornea* and *lens* project an inverted image on the light-sensitive surface of the *retina*. Sharpest vision is possible in the *fovea*. At the *blind spot*, where no vision is possible, the nerves leave the eye. *Rods* and *cones* are distributed around the fovea and relate their name to their shape. Cones provide the visual system with the ability to perceive colors and rods provide the visual system with monochromatic information even under dark light conditions [Hunt, 1991].

2.2 Human perception

The human visual system consists of the eye (illustrated in figure 2.5), its nerves and the processing of the output of the nerves in the brain. The sensors in the human eye are distinguished by their shape into rods and cones. Rods provide the visual system with monochromatic information (even under poor lighting conditions). Cones, which are further distinguished by their response behavior, provide the visual system with color information [Hunt, 1991].

It is assumed [Hunt, 1991] that three different types of cones with different spectral sensitivities exist. They have their maximum spectral sensitivities at the colors of yellow-orange, green and blue-violet and are captioned R,G,B. Since their caption is misleading they are called ρ, γ and β. Their sensitivity behavior is depicted in figure 2.6.

Color perception works by processing the response of the cones (see figure 2.7). It may result in the same color perception, although different spectral distributions are involved in the process, which is related to the different spectral sensitivity distributions of the cones. The effect of equal color perception of different colors is well known as metamerism and is an important challenge for the color and print industry [Hunt, 1991].

Metamerism describes the equivalent perceptional appearance of colors which possess a different spectral appearance. The appearance of these colors varies under different lighting conditions and the human eye cannot distinguish between them at least at one lighting condition. Metamerism is described by the metameric index which assigns a high index when a pair of colors are perceptually equivalent under certain lighting conditions, and perceptually different

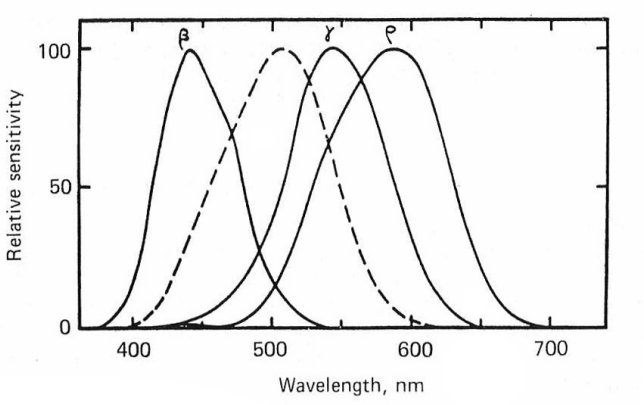

Figure 2.6: Spectral sensitivity distribution of the cones (ρ, γ and β) in the human eye. The dashed line shows the spectral sensitivity of vision based on the stimuli of the rods [Hunt, 1991]. The assumed composition of cone and rod output is shown in figure 2.7.

under other lighting conditions. If a pair of colors appears perceptually equivalent under all lighting conditions it has a metameric index equal to zero [Johnson and Fairchild, 1999].

Human perception can be summarized as a composition of three different sensations: *brightness*, which describes the intensity of the light that reaches the eye; *Hue*, which describes the perceived hue; and *colorfulness*, which describes how pure a perceived color is. In literature, the perceived bandwidth by a human observer is set between the wavelengths of 400nm (violet) and 700nm (red) [Nassau, 2001]. The three different sensations of human perception (brightness, hue and colorfulness) are necessary to specify a color. An example color is given in figure 2.4 to show these elements.

The human visual system is able to adapt to different lighting conditions. Even if there are great differences in light intensity, as for example, daylight compared to artificial lighting in a room, the human eye is able to adapt to perceived intensity and color. The result of this adaption is that objects are perceived in the same color, although the illumination is different. For example, a green apple is perceived in the same color in daylight at noon and in a room in the evening [Hunt, 1991]. This adaption is called color constancy and is an important topic in color science.

The assumption that human perception relates to the stimuli of three different kinds of cones leads to the theory, that it is possible to obtain their response curves. The *commission internationale de l'eclairage* (CIE) did this in an empiric way, which lead to the RGB color system and the response curves depicted in figure 2.9. They conducted the so called *color matching experiments* (CME). They assumed that the input of the rods in the human eye may be neglected for pure color vision if lighting is sufficient for it. The setup for the color matching experiments basically consists of a device which enables the test subject to mix colored light beams (red, green and blue) (see figure 2.8 for details). The test subject is instructed to mix the colors together to produce a desired color stimuli [Hunt, 1991]. It turned out that in order to produce certain targeted color stimuli in that way, the colors red and green had also to be subtracted from the mix of colored light beams. Because this would require in certain cases something like negative

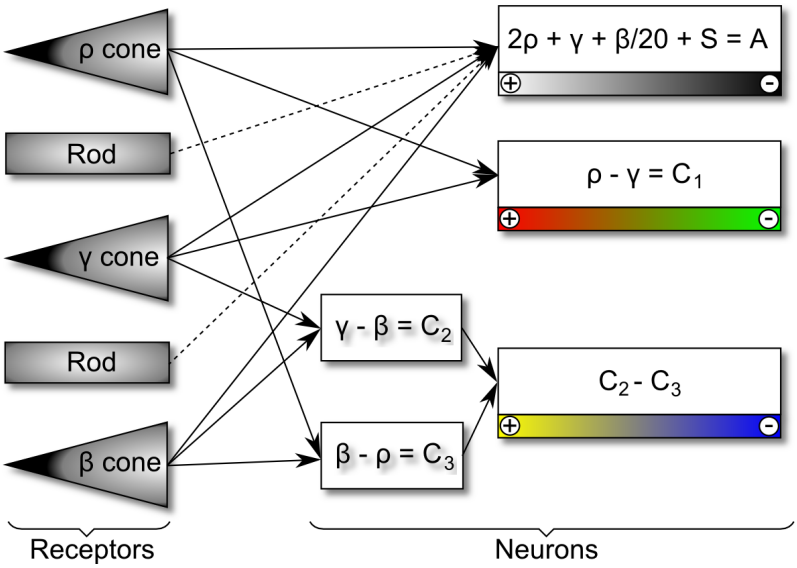

Figure 2.7: This image shows a simplified representation of the widely assumed connection between cone and rod output, which results in human perception. It is assumed that the connection of the output results in an achromatic signal A and three color difference signals C_1, C_2 and C_3. The achromatic signal is composed of the output of all rods and cones. The different occurrence of cones is simply represented by the term $2\rho + \gamma + \beta/20 + S$ where S denotes the scotopic input of the rods. Physiological studies indicate that the output of color difference signals is further processed as two signals C_1 and $C_2 - C_3$ because of the fact $C_1 + C_2 + C_3 = 0$ [Hunt, 1991].

light which does not exist, the light beams which have to be subtracted are instead added to the targeted color [Hunt, 1991].

The results of the CME lead to the CIE RGB color matching functions $\overline{r}(\lambda)$,$\overline{g}(\lambda)$ and $\overline{b}(\lambda)$ (plotted in figure 2.9) and the CIE RGB color space. The values for a given color for R, G and B are calculated by using the equations 2.1, 2.2 and 2.3 [Hunt, 1991]. Equation 2.4 [Hunt, 1991] is the luminance composed of the stimuli of R,G and B. The luminance L result is given in candelas per square meter (cd/m^2) when P_i is given in watts per steradian and per square meter, and k is chosen correctly. P_i denotes the amounts of power sampled throughout the spectrum and $\overline{r}_i$, $\overline{g}_i$ and $\overline{b}_i$ are the corresponding power levels of the color matching functions [Hunt, 1991].

$$R = k(P_1\overline{r}_1 + P_2\overline{r}_2 + P_3\overline{r}_3 + ...) \tag{2.1}$$
$$G = k(P_1\overline{g}_1 + P_2\overline{g}_2 + P_3\overline{g}_3 + ...) \tag{2.2}$$
$$B = k(P_1\overline{b}_1 + P_2\overline{b}_2 + P_3\overline{b}_3 + ...) \tag{2.3}$$
$$L = 1.0000R + 4.5907G + 0.0601B \tag{2.4}$$

The main disadvantage of CIE RGB is the presence of the negative red and green components. This disadvantage was eliminated by mathematically transforming the color space, which resulted in a new color space called CIE XYZ.

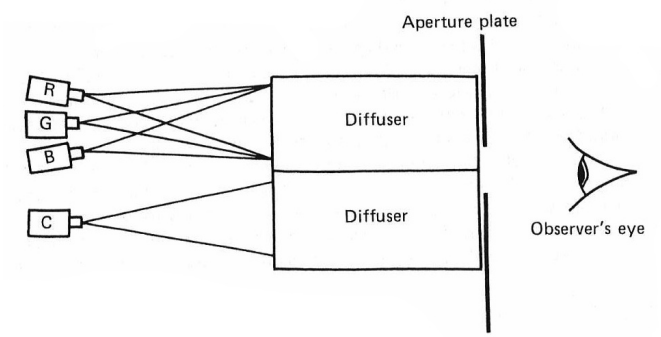

Figure 2.8: Color matching experiment setup which consists of three monochromatic lights R (700nm),G (546.1nm) and B (435.8nm). The light beams are mixed together by a test subject to produce a desired color stimuli C. Negative color components are added to the color stimuli C because there is no negative light [Hunt, 1991].

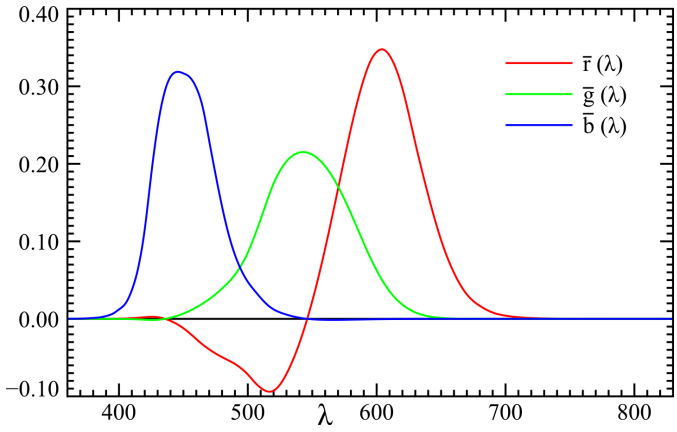

Figure 2.9: Color matching functions for the CIE 1931 standard colorimetric observer which were obtained by color matching experiments. This plot also shows the negative components of red and green.

The transformation from CIE RGB to CIE XYZ is given in the equations 2.5,2.6 and 2.7 [Hunt, 1991]. There, $\overline{x}(\lambda)$,$\overline{y}(\lambda)$, $\overline{z}(\lambda)$, $\overline{r}(\lambda)$, $\overline{g}(\lambda)$ and $\overline{b}(\lambda)$, denote the color matching functions of both color spaces. The color matching functions $\overline{x}(\lambda)$,$\overline{y}(\lambda)$ and $\overline{z}(\lambda)$, which are the most important spectral functions in colorimetry, are plotted in figure 2.10 [Hunt, 1991].

$$\overline{x}(\lambda) = 0.49\,\overline{r}(\lambda) + 0.31\,\overline{g}(\lambda) + 0.20\,\overline{b}(\lambda) \tag{2.5}$$

$$\overline{y}(\lambda) = 0.17697\,\overline{r}(\lambda) + 0.81240\,\overline{g}(\lambda) + 0.01063\,\overline{b}(\lambda) \tag{2.6}$$

$$\overline{z}(\lambda) = 0.00\,\overline{r}(\lambda) + 0.01\,\overline{g}(\lambda) + 0.99\,\overline{b}(\lambda) \tag{2.7}$$

2.3 Color space: CIE XYZ

The values for X,Y and Z of an emissive source for a given wavelength distribution $I(\lambda)$ may be obtained by applying equations 2.8, 2.9 and 2.10 [Lindbloom, 2007]. The values $\overline{x}(\lambda)$,$\overline{y}(\lambda)$ and $\overline{z}(\lambda)$ represent the color matching functions. $P(\lambda)$ describes the spectral power distribution of the given emissive source.

$$X = \int_{\lambda} \bar{x}(\lambda)P(\lambda)d\lambda \tag{2.8}$$

$$Y = \int_{\lambda} \bar{y}(\lambda)P(\lambda)d\lambda \tag{2.9}$$

$$Z = \int_{\lambda} \bar{z}(\lambda)P(\lambda)d\lambda \tag{2.10}$$

For reflective or transmissive samples the equations 2.11, 2.12 and 2.13 are used [Nassau, 2001]. There, $S(\lambda)$ denotes the spectral power distribution of the involved illuminant. $R(\lambda)$ denotes the reflectance factor of the object. The color matching functions are represented by $\bar{x}(\lambda)$, $\bar{y}(\lambda)$ and $\bar{z}(\lambda)$.

$$X = k \int S(\lambda)R(\lambda)\bar{x}(\lambda)d\lambda \tag{2.11}$$

$$Y = k \int S(\lambda)R(\lambda)\bar{y}(\lambda)d\lambda \tag{2.12}$$

$$Z = k \int S(\lambda)R(\lambda)\bar{z}(\lambda)d\lambda \tag{2.13}$$

The constant k is dependent on the illuminant $S(\lambda)$ and is given in equation 2.14 [Nassau, 2001]. There, k is defined in a way that Y in equation 2.12 yields 100 for a perfectly reflecting diffuser [Nassau, 2001]. In practical applications the range of the integration (or summation) is set between $380nm$ and $720nm$ [Hunt, 1991] or even wider between $360nm$ and $830nm$ [Lindbloom, 2007].

$$k = \frac{100}{\int S(\lambda)\bar{y}(\lambda)d\lambda} \tag{2.14}$$

In CIE XYZ the Y component of a color correlates to its brightness [Hunt, 1991]. But X and Z do not correlate to any perceptual property. As Hunt [Hunt, 1991] points out in his book about color:

> *"Important colour attributes are related to the relative magnitudes of the tristimulus values. It is therefore helpful to calculate a type of relative tristimulus values called chromaticity co-ordinates[.]"*[Hunt, 1991]

The chromaticity coordinates are given in equations 2.15, 2.16 and 2.17 [Nassau, 2001]. The

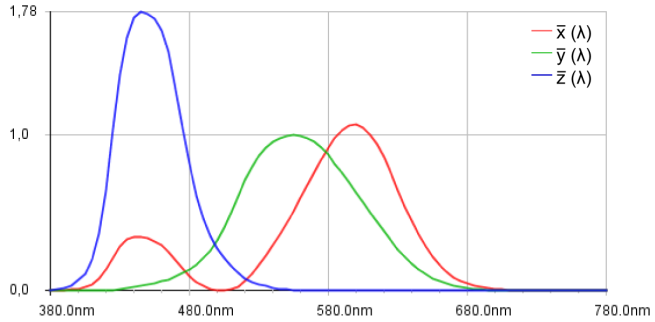

Figure 2.10: Color matching functions for CIE XYZ. These were obtained by transforming the CIE 1931 RGB color matching functions [Hunt, 1991].

normalization by $(X+Y+Z)$ in these equations leads to $x+y+z=1$ and to the implication, that if two chromaticity coordinates are known the third can be deduced [Hunt, 1991]. The chromaticity coordinates x and y were used to construct the CIE 1931 x,y chromaticity diagram. In this diagram any combination of x and y is represented as a colored point in a two dimensional representation [Hunt, 1991]. Regarding the nature of x and y, which both relate to the RGB system, the chromaticity has to be seen as a map of different color stimuli [Hunt, 1991].

$$x = X/(X+Y+Z) \tag{2.15}$$
$$y = Y/(X+Y+Z) \tag{2.16}$$
$$z = Z/(X+Y+Z) \tag{2.17}$$

The CIE chromaticity diagram is a normed way to represent a colored light beam. In order to represent a color of a surface an additional dimension (Y), which represents the brightness of the color, is necessary [Nassau, 2001].

The chromaticity diagram and the CIE XYZ system are suitable to display differences in color stimuli, but not differences in actual color perception. The need to handle differences in color perception leads to the development of the *CIE 1976 uniform chromaticity scale diagram* (CIE 1976 UCS diagram) and to uniform color spaces like CIE L*a*b and CIE L*u*v [Hunt, 1991]. In the CIE 1976 UCS diagram the two axis u' and v' are calculated from X,Y and Z, or x and y. The coordinates u' and v' are given in equations 2.18 and 2.19 [Nassau, 2001].

$$u' = \frac{4X}{(X+15Y+3Z)} = 4x(3-2x+12y) \tag{2.18}$$
$$v' = \frac{9Y}{(X+15Y+3Z)} = 9y(3-2x+12y) \tag{2.19}$$

There, the differences between two colors (or two points in the diagram) correspond better to the same difference in color perception than in the CIE 1931 chromaticity diagram. It is important to mention that there are no two dimensional representation which may reduce this non-uniformity entirely [Hunt, 1991].

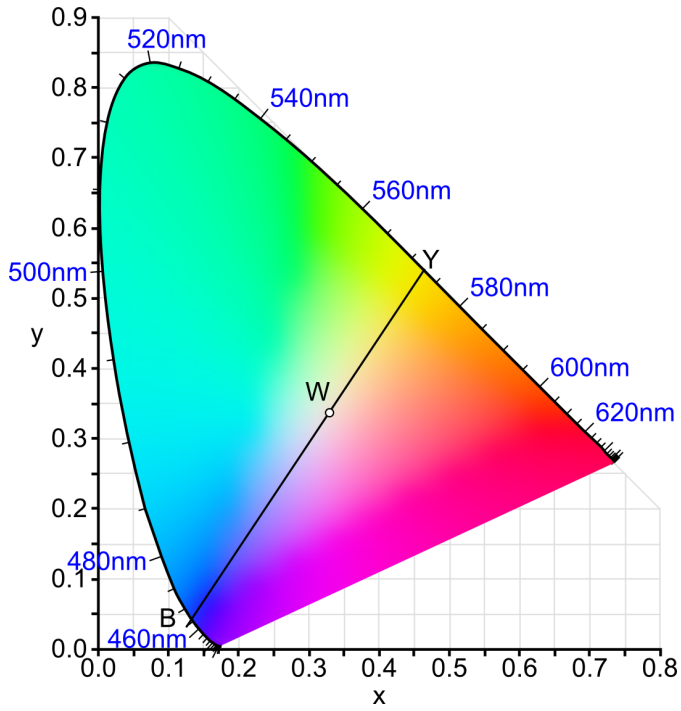

Figure 2.11: CIE1931 chromaticity diagram. The central point W of the diagram denotes the color white. The curved edge of the shape, also known as the "spectral locus", denotes the spectrum. Hues with highest saturation are on the edge of this curved edge. The straight line from red to violet is also known as "non spectral hue" or " purple line" and is obtained by mixing red and violet. B and Y denote two complementary color examples. They are complementary because the straight line between them runs through the color white (point W). By mixing two complementary colors, like B and Y, any color on the straight line between them can be obtained. When adding a third or fourth color the resulting shape defines the colors which may be obtained by mixing these colors [Hunt, 1991].

2.4 Color space: CIE L*a*b and CIE L*u*v

The CIE published the CIE L*a*b at the same time as CIE L*u*v in the year 1976. Both systems compensate the disadvantage of non-uniformity of CIE XYZ [Nassau, 2001]. CIE L*a*b and CIE L*u*v are non-linear transformations based on CIE XYZ. Both result in a three dimensional color space.

CIE L*u*v which is used in television [Hunt, 1991] and correlates to the chromaticity coordinates u' and v' (given in equations 2.18 and 2.19), is defined by these values: L for Luminance, and two values u and v. The following equations 2.20-2.23 [Hunt, 1991] describe the conversion from CIE XYZ to CIE L*u*v. There, the values u'_n and v'_n are the values of the reference white [Hunt, 1991],

$$L^* = 116(Y/Y_n)^(1/3) - 16 \;\; for\, Y/Y_n > 0.008856 \tag{2.20}$$
$$L^* = 903.3(Y/Y_n) \;\; for Y/Y_n \; \leq 0.008856 \tag{2.21}$$
$$u^* = 13L^*(u' - u'_n) \tag{2.22}$$
$$v^* = 13L^*(v' - v'_n) \tag{2.23}$$

CIE L*a*b which is used in colorant industries [Hunt, 1991] is defined by the values: L for Luminance, and two values a and b. Equations 2.24-2.29 [Lindbloom, 2007] describe the conversion from CIE XYZ to CIE L*a*b color space. This calculation involves the use of a reference white source for color conversion [Hunt, 1991, Wyszecki and Stiles, 2000]. The reference white source defines which color is assumed as the color white for the conversion.

$$L = 116f_y - 16 \tag{2.24}$$
$$a = 500(f_x - f_y) \tag{2.25}$$
$$b = 200(f_y - f_z) \tag{2.26}$$
$$f_x = \begin{cases} \sqrt[3]{x_r} & x_r > \epsilon \\ \frac{\kappa x_r + 16}{116} & x_r \leq \epsilon \end{cases} \quad x_r = \frac{X}{X_r} \tag{2.27}$$
$$f_y = \begin{cases} \sqrt[3]{y_r} & y_r > \epsilon \\ \frac{\kappa y_r + 16}{116} & y_r \leq \epsilon \end{cases} \quad y_r = \frac{Y}{Y_r} \tag{2.28}$$
$$f_z = \begin{cases} \sqrt[3]{z_r} & z_r > \epsilon \\ \frac{\kappa z_r + 16}{116} & z_r \leq \epsilon \end{cases} \quad z_r = \frac{Z}{Z_r} \tag{2.29}$$

The two constants κ and ϵ are defined by the CIE and are given in equations 2.30 and 2.31 [Lindbloom, 2007].

$$\kappa = 903.3 \tag{2.30}$$
$$\epsilon = 0.008856 \tag{2.31}$$

CIE L*a*b and CIE L*u*v introduce measurements called hue-angles which better correlate to hue differences. They are defined in equation 2.32 and 2.33 [Hunt, 1991].

$$h_{uv} = arctan[(v' - v'_n)/(u' - u'_n)] \tag{2.32}$$
$$h_{ab} = arctan(b^*/a^*) \tag{2.33}$$

Both systems arrange colors similar to the Munsell color system. Albert Henry Munsell (1858-1918) aimed to arrange colors in his system in such a way as to keep the perceived color difference between any two neighboring colors constant. He also published an color atlas, the Munsell book of color [Hunt, 1991].

The main advantage of CIE L*a*b and CIE L*u*v is that their uniformity makes it possible to represent a perceived color difference between two colors through the geometrical distance

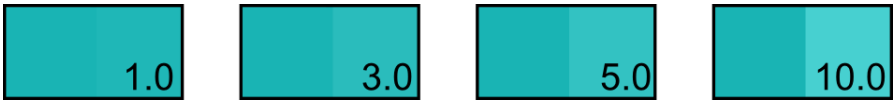

Figure 2.12: This image shows color patches containing examples for different color differences. Based on a blue hue every patch contains the reference color on the left side and a different color on the ride side. The color on the right side was chosen in such a way that the color difference equation CIE76 yields the result labeled in the patch.

between them. This is also known as the color difference equation CIE76 which is given for CIE L*a*b in equation 2.37 [Hunt, 1991]. The results of equation 2.37 are greater than or equal to zero and can be interpreted as the higher the result of the equation the higher the perceived color difference between the two colors $L_1a_1b_1$ and $L_2a_2b_2$. Figure 2.12 illustrates some example color differences.

$$\Delta L = L_1 - L_2 \tag{2.34}$$

$$\Delta a = a_1 - a_2 \tag{2.35}$$

$$\Delta b = b_1 - b_2 \tag{2.36}$$

$$\Delta d_{76} = \sqrt{\Delta L^2 + \Delta a^2 + \Delta b^2} \tag{2.37}$$

2.5 Color difference equation: CIEDE2000

The CIE has released other color difference calculation methods over the years. These new methods, CIE94 and CIEDE2000, are optimizations regarding the perceivable difference between two colors [Sharma et al., 2004], because the geometric distance is not always sufficient.

For example, CIEDE2000 makes the assumption that L in the color spaces L*a*b or L*u*v is not the correct perceived luminance of a color. This is a major difference to its forerunners CIE94 and CIE76. CIEDE2000 is complex to work with, because it involves a greater number of calculations than CIE76 and CIE94. Works on implementing this complex formula have been published by [Lindbloom, 2007], [Johnson and Fairchild, 2003] and [Sharma et al., 2004].

I follow the step-by-step calculation as shown by [Sharma et al., 2004] and the equations for calculation by [Lindbloom, 2007] to explain the color difference method CIEDE2000 between two colors, $L_1a_1b_1$ and $L_2a_2b_2$.

The final equation is given in equation 2.38 [Lindbloom, 2007], the symbols and indices have to be read as follows: L stands for lightness, C for chroma, H for hue angle, R for rotation, S for weighting functions and K for parametric weights which are usually unknown and therefore set to $K_L = K_S = K_H = 1.0$. For calculation purposes all angles have to be assigned in degrees.

$$\Delta E_{CIEDE2000} = \sqrt{\left(\frac{\Delta L'}{K_L S_L}\right)^2 + \left(\frac{\Delta C'}{K_C S_C}\right)^2 + \left(\frac{\Delta H'}{K_H S_H}\right)^2 + R_T \left(\frac{\Delta C'}{K_C S_C}\right)\left(\frac{\Delta H'}{K_H S_H}\right)} \tag{2.38}$$

The first step consists of the calculation of the intermediate variables $C'_{1,2}$ and $h'_{1,2}$ as given in equations 2.39-2.48 [Lindbloom, 2007]. The calculations of $C'_{1,2}$ includes a correction of the a-axis denoted by a factor $(1+G)$ which compensates for correct color difference perception at low chromaticity values [Johnson and Fairchild, 2003]. Lindbloom [Lindbloom, 2007] suggests to use the method *atan2* for the implementation of tan^{-1} in equations 2.47 and 2.48 because this function avoids possible divisions by zero.

$$C_1 = \sqrt{a_1^2 + b_1^2} \tag{2.39}$$

$$C_2 = \sqrt{a_2^2 + b_2^2} \tag{2.40}$$

$$\overline{C} = (C_1 + C_2)/2 \tag{2.41}$$

$$G = 0.5\left(1 - \sqrt{\frac{\overline{C}^7}{\overline{C}^7 + 25^7}}\right) \tag{2.42}$$

$$a'_1 = a_1(1+G) \tag{2.43}$$

$$a'_2 = a_2(2+G) \tag{2.44}$$

$$C'_1 = \sqrt{a_1'^2 + b_1^2} \tag{2.45}$$

$$C'_2 = \sqrt{a_2'^2 + b_2^2} \tag{2.46}$$

$$h'_1 = \begin{cases} tan^{-1}(b_1/a'_1) & tan^{-1}(b_1/a'_1) \geq 0 \\ tan^{-1}(b_1/a'_1) + 360^\circ & tan^{-1}(b_1/a'_1) < 0 \end{cases} \tag{2.47}$$

$$h'_2 = \begin{cases} tan^{-1}(b_2/a'_2) & tan^{-1}(b_2/a'_2) \geq 0 \\ tan^{-1}(b_2/a'_2) + 360^\circ & tan^{-1}(b_2/a'_2) < 0 \end{cases} \tag{2.48}$$

The second step is the calculation of $\Delta L'$, $\Delta C'$ and $\Delta H'$ as given in the equations 2.49-2.52 [Lindbloom, 2007]. $\Delta H'$ has to be calculated differently depending on the orientation of the hue-angle difference $h'_2 - h'_1$.

$$\Delta L' = L_2 - L_1 \tag{2.49}$$

$$\Delta C' = C'_2 - C'_1 \tag{2.50}$$

$$\Delta H' = 2\sqrt{C'_1 C'_2}\sin(\Delta h'/2) \tag{2.51}$$

$$\Delta h' = \begin{cases} h'_2 - h'_1 & |h'_2 - h'_1| \leq 180^\circ \\ h'_2 - h'_1 + 360^\circ & |h'_2 - h'_1| > 180^\circ; h'_2 \leq h'_1 \\ h'_2 - h'_1 - 360^\circ & |h'_2 - h'_1| > 180^\circ; h'_2 > h'_1 \end{cases} \tag{2.52}$$

The third step is to calculate the weighting functions S_L, S_C and S_H, and the rotation R_T [Johnson and Fairchild, 2003]. The weighting functions 2.53-2.55 were introduced to adjust color difference perception between lightness, chroma und hue [Johnson and Fairchild, 2003]. The weighting function for hue angles S_H also includes a chroma term T [Johnson and Fairchild, 2003]. The rotation R_T [Lindbloom, 2007] addresses an issue of CIE L*a*b. This issue is that blue regions are highly non-linear in regard to the perceived color differences.

$$S_L = 1 + \frac{0.015(\overline{L}' - 50)^2}{\sqrt{20 + (\overline{L}' - 50)^2}} \tag{2.53}$$

$$S_C = 1 + 0.045\overline{C}' \tag{2.54}$$

$$S_H = 1 + 0.015\overline{C}'T \tag{2.55}$$

$$\overline{L}' = (L_1 + L_2)/2 \tag{2.56}$$

$$\overline{C}' = (C_1' + C_2')/2 \tag{2.57}$$

$$T = 1 - 0.17\cos(\overline{H}' - 30^\circ) + 0.24\cos(2\overline{H}') + 0.32\cos(3\overline{H}' + 6^\circ) - 0.20\cos(4\overline{H}' - 63^\circ) \tag{2.58}$$

$$R_T = -2R_C\sin(2\Delta\Theta) \tag{2.59}$$

$$R_C = \sqrt{\frac{\overline{C}'^7}{\overline{C}'^7 + 25^7}} \tag{2.60}$$

$$\Delta\Theta = 30\exp\left\{-\left(\frac{\overline{H}' - 275^\circ}{25}\right)^2\right\} \tag{2.61}$$

$$\overline{H}' = \begin{cases} (h_1' + h_2' + 360^\circ)/2 & |h_1' - h_2'| > 180^\circ \\ (h_1' + h_2')/2 & |h_1' - h_2'| \leq 180^\circ \end{cases} \tag{2.62}$$

2.6 Additional conversion formulas

The formulas (2.63-2.67)[Lindbloom, 2007] show the conversion from XYZ to sRGB color space. Like the conversion from spectral data to CIE XYZ before, it also involves the use of a reference white. Usually this conversion formulas are used to calculate values for displaying purposes.

$$\begin{pmatrix} r \\ g \\ b \end{pmatrix} = [M_{sRGBD65}]^{-1} * \begin{pmatrix} X \\ Y \\ Z \end{pmatrix} \tag{2.63}$$

$$[M_{sRGBD65}]^{-1} = \begin{pmatrix} 3.2404542 & -1.537385 & -0.4985314 \\ -0.9692660 & 1.8760108 & 0.0415560 \\ 0.0556434 & -0.2040259 & 1.0572252 \end{pmatrix} \tag{2.64}$$

$$R = \begin{cases} 12.92r & r \leq 0.0031308 \\ 1.055r^{1.0/2.4} - 0.055 & r > 0.0031308 \end{cases} \tag{2.65}$$

$$G = \begin{cases} 12.92g & g \leq 0.0031308 \\ 1.055g^{1.0/2.4} - 0.055 & g > 0.0031308 \end{cases} \tag{2.66}$$

$$B = \begin{cases} 12.92b & b \leq 0.0031308 \\ 1.055b^{1.0/2.4} - 0.055 & b > 0.0031308 \end{cases} \tag{2.67}$$

Besides matrix 2.64 there are other conversion matrices for different reference white sources. Several can be looked up in [Lindbloom, 2007].

2.7 CIE standard illuminants

The CIE has normed a selection of light sources as standard illuminants to be used as references for color comparison and color applications. These standard illuminants are given with their color temperature, which defines a color by its black body radiation. A so called black body is an ideal object that absorbs all incident light and emits light caused by thermal radiation. The color temperature is the surface temperature of this object [Nassau, 2001]. The CIE standard illuminants are [Wyszecki and Stiles, 2000]:

- CIE A. This standard illuminant corresponds to a gas filled coiled tungsten lamp with a color temperature of 2856K.
- CIE B. This standard illuminant represents sunlight at noon with a color temperature of 4874K.
- CIE C. This standard illuminant represents average daylight with a color temperature of 6774K.
- CIE D65. This standard illuminant was derived from CIE C and represents daylight with a color temperature of 6500K.

The standard illuminants above are all non-fluorescent light sources. The difference between non-fluorescent and fluorescent light is illustrated in figure 2.13. It can be observed that a fluorescent light source differs from a non-fluorescent light source because it contains peaks in its spectral distribution. The handling of these peaks is an important aspect of any spectral data representation method.

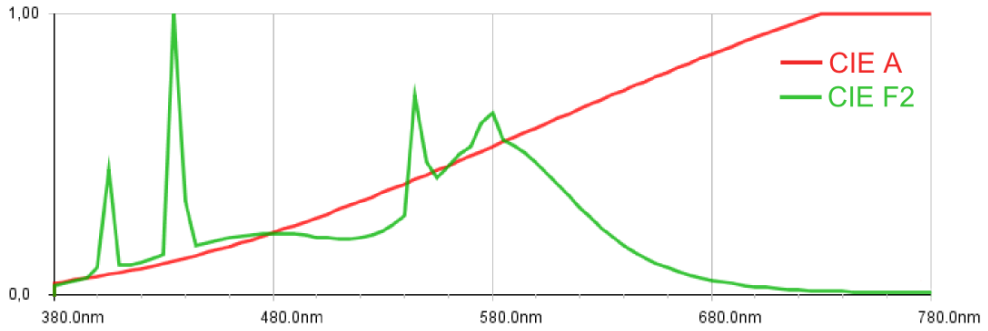

Figure 2.13: This image shows two CIE illuminants, A and $F2$ (both were scaled to a maximum value of 1.0 for displaying purpose). The difference between the two illuminants is clearly visible. $F2$ contains several peaks which interrupt its otherwise smooth distribution.

The mentioned peaks are usually much smaller than the sampling interval used in an application. This causes a distortion (or aliasing) which results in a noticeable color difference compared to a given reference. This hails from the violation of the Nyquist criteria which usually prevents the loss of information [Hearn and Baker, 2004]. More details on this topic are reviewed in chapter 4.

The CIE also published standard illuminants for fluorescent light sources [Hunt, 1991], which were primary designed for industrial applications. The CIE fluorescent illuminants are:

- CIE F1-F6. These standard fluorescent light sources include: F1 (daylight at 6430K), F2 (cool white light like the light of a common fluorescent lamp at 4230K), F3 (white light at 3450K), F4 (warm white light at 2940K), F5 (daylight at 6350K) and F6 (white light at 4150K).

- CIE F7-F9. These broad band fluorescent light sources which were designed for better color rendering properties include: F7 (daylight at 6500K), F8 (light at 5000K) and F9 (cool white light at 4150K)

- CIE F10-F12. These narrow band fluorescent light sources can be mixed together to produce highly efficient white light.

Literature suggests to use the fluorescent light sources F2,F7 and F11 for testing, because each of them represents a typical one of their kind [HunterLAB, 2005]. This is the main reason why I will use them for testing color representation methods which include the use of fluorescent light sources.

Apart from color system which are used to specify a color, there are color ordering systems like the natural color system [Hunt, 1991]. Color ordering systems are used for checking, selecting and measuring color in a wide spectrum of applications. They are designed to be easy to use but impose difficulties when using two different color ordering systems together [Hunt, 1991].

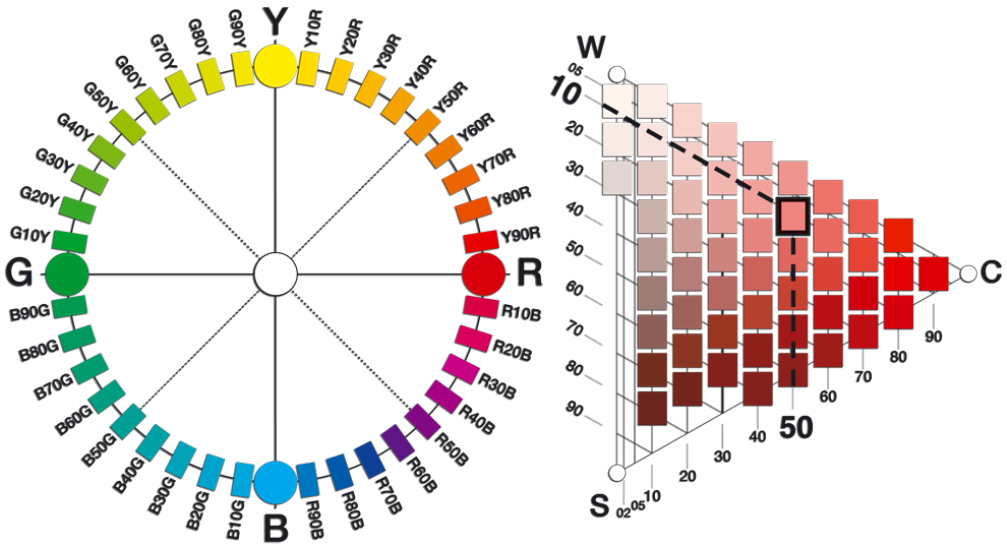

Figure 2.14: NCS color wheel (left image) and an example for a NCS color triangle (right image) at Y90R. The left image shows the NCS color wheel with its 10 percent transitions between the colors yellow, green, red and blue. The image on the right shows a NCS color triangle taken from the color Y90R (this caption corresponds to a hue composed of ninety percent red with ten percent yellow). This triangle shows the variations of the selected color Y90R and different white/black combinations. The selected color is at blackness ten percent and chromaticness fifty percent. All mentioned color information drawn together results in the exact definition of a color in the NCS system. The example on the right can be expressed in standardized NCS notation as *NCS S 1050 - Y90R*.

2.8 Natural color system (NCS)

The NCS (natural color system) was developed in Sweden. It is a three dimensional color system [Hunt, 1991] which is based on the complementary color theory of Hering [Hering, 2007] and represents colors by a combination of four unique hues. Hering proposed to use four unique hues and their transitions (red to green and blue to yellow), which relate to human color perception, instead of three (red, green and blue). These hues are called unique because they cannot be represented by a combination of other colors [Hunt, 1991].

The basic colors of NCS are white (W), black (S), yellow (Y), green (G), red (R) and blue (B). In this system gray tones are labeled as neutral (N). The perceived fraction of the basic color is expressed as a percentage value. The whole NCS data set was obtained by observation of around 60,000 test subjects and contains about 16,000 NCS combinations. For the final NCS atlas this number was truncated by clipping the numbers to 10 percent steps. This resulted in about 2,000 color combinations. Due to missing color pigmentation 1,930 entries remained in the final NCS atlas [Hunt, 1991].

CHAPTER 3

Wavelength dependent effects

Computer graphics (CG) is used in many different applications in science, art, engineering, business, industry, medicine, entertainment, education and training. Generating realistic scenes has become more and more sophisticated in recent years, especially in fields were photo-realism counts. For example, in the final stages of product design, CAD/CAM applications give an outlook on the appearance of the final workpiece by applying realistic lighting and surface conditions. It is also used for advertising (e.g. of automobiles) and to show both architects and their clients how a building will look in reality [Hearn and Baker, 2004].

The ultimate aim of this applications is to generate images which trigger the same impression for the viewer as the scene in reality would [Devlin et al., 2002]. To achieve this, a broad spectrum of issues to simulate reality have to be addressed, such as the correct simulation of shapes, colors (and illumination), motion and natural phenomena [Sun, 2000]. The group of natural phenomena is further distinguished into optical and geometric effects.

An important subgroup of optical effects involves wavelength dependent light/surface interactions. These interactions are influenced by the spectral appearance of the involved lights and materials. Spectral rendering techniques are used to implement these effects and enable an application to yield plausible results. The term spectral in this context describes everything which involves wavelength dependency [Devlin et al., 2002].

This chapter contains a review of the most common spectral effects: fluorescence and phosphorescence, refraction and dispersion, scattering, interference and diffraction.

3.1 Fluorescence and phosphorescence

Glassner [Glassner, 1995a] gives an example for fluorescence in his book, a fluorescent electric bulb. The electric bulb emits absorbed energy at an a wavelength which is usually in the visible spectrum as visible light. The fluorescent material in this example is the gas in the bulb.

Fluorescence is a material property which occurs within molecular structures [Devlin et al., 2002]. It describes an effect which becomes visible when light is absorbed at

one wavelength and re-emitted at a different, usually longer, wavelength. An important property of this effect is the time span between absorbance and re-emittance. This time span is usually smaller than 10^-8 seconds for fluorescence. The emitted energy has no orientation and is uniform in any direction [Nassau, 2001].
Phosphorescence is a similar effect to fluorescence, but incurs a longer time span than the mentioned 10^-8 seconds. This effect occurs when light is slowly emitted by a phosphorescent material during the transition from a higher to a lower energy state. The emitted light is also not oriented and uniformly distributed in any direction [Glassner, 1995a]. Examples for phosphorescence include "glow-in-the-dark" materials [Nassau, 2001] and the cathode ray tube (CRT) monitor [Glassner, 1995a] where the screen of the CRT is coated with phosphors for each color. These phosphors are targeted and shot at by one of three electron guns (one for each color; red, green and blue). The phosphors absorb the electrons and re-emit them over time (in this example as light in the visual spectrum) [Glassner, 1995a].

3.2 Refraction and dispersion

Refraction usually happens during the transition of electromagnetic waves between two materials. This transition results in a deceleration and redirection of the involved electromagnetic waves. With an exception to the angle of incidence of incident light at 90 degrees this also results in a bending of the electromagnetic waves. The property of a material which describes its refractional behavior is described by Snell's Law (equation 3.1) [Nassau, 2001] and the refractive index. An example for refraction in nature is the bending of light when light enters a transparent glass filled with water. The effect becomes visible if an object is partially put into the water. From outside, the object can be perceived as bended under the surface of the water. This perception is caused by refraction. Another example for refraction is the distortion of the setting sun. There, the perceived appearance of the sun is distorted by air with varying refractive indices depending on its height above the ground [Lynch and Livingston, 1995].

$$\sin \Theta_r = \frac{\eta_i}{\eta_r} \sin \Theta_i \tag{3.1}$$

Equation 3.1 contains the refractive indices of the materials η_i and material η_r. The angle Θ_r describes the angle of refraction, and Θ_i the angle of incidence [Hearn and Baker, 2004]. Some example values for refractive indices are: $\eta = 1.0$ for vacuum, $\eta = 1.52$ for ordinary crown glass, $\eta = 1.33$ for water and $\eta = 1.31$ for ice [Hearn and Baker, 2004].

With the exception of vacuum the refractive index of a material depends on the wavelength of the incident light and the temperature of the material. Refraction or the deceleration of electromagnetic waves occurs not uniformly along the electromagnetic spectrum and also depends on the orientation of the inner structure of the material (e.g. crystalline structures)[Nassau, 2001]. The wavelength dependent behavior of refraction is known as dispersive refraction or as **dispersion.**

Isaac Newton discovered dispersion when he conducted his prism experiment (see figure 2.1 for details). He observed that a white beam of light can be split in its spectral colors by

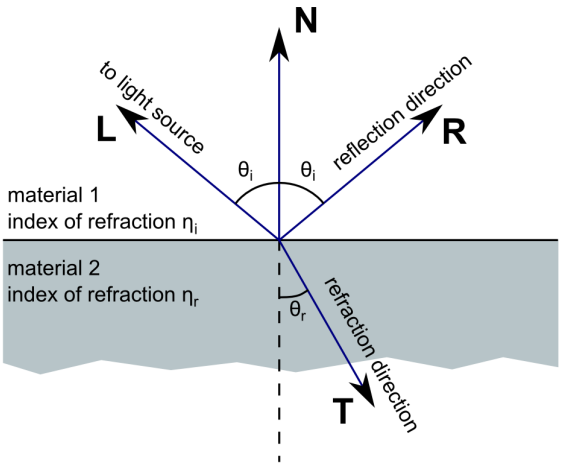

Figure 3.1: This image shows the elements involved in refraction: the refractive indices η_i and η_r of both materials, the angle of incidence Θ_i and the angle of refraction Θ_r. The angle of refraction Θ_r is determined by Snell's Law which is given in equation 3.1. The vector N denotes the normal vector of the point on the surface [Hearn and Baker, 2004].

a prism made of glass. Within the material (glass) of the prism the refractive index decreases with increasing wavelength. This wavelength dependency causes the prism to split the white light into its spectral components [Nassau, 2001]. Examples for dispersion in nature are the fire in cut gems, the colors of the rainbow and the green flash [Nassau, 2001]. The green flash is a green spot which becomes briefly visible before sunrise or after sunset at an unobstructed horizon above the sun [Lynch and Livingston, 1995].

3.3 Scattering

When an electromagnetic wave hits a small particle the particle may interact with this wave. The resulting effects are summarized as scattering. The particle size d in relation to the wavelength determines which type of scattering takes place. If the particle size d is much smaller than the involved wavelength λ ($d < \lambda/10$) then the occurring scattering is called Rayleigh scattering. If the particle is around the size of the involved wavelength ($d \approx \lambda$) Mie scattering appears instead of Rayleigh scattering. If the particles are much larger than the involved wavelength ($d >> \lambda$) geometric scattering appears at the edges of the particles which results in the effect called diffraction [Haferkorn, 2002].

Rayleigh scattering

Lord John William Rayleigh (1842-1919) analyzed the effects caused by scattering, such as the color of the blue sky. He defined Rayleigh scattering and stated that scattering not only occurs with small particles, but also appears in slightly impure materials and liquids. These impurities cause a variation in the refractive index of the material and thus cause scattering.

The intensity I of the scattered light is given in equation 3.2 [Nassau, 2001]. The equation is a relation of the intensity of the scattered light I_s and the incident light I_0. The relation is proportional to the inverse of the fourth power of the wavelength λ. This equation describes the fact that the lower the wavelength of the light the more scattering occurs when light travels through a medium [Lynch and Livingston, 1995].

$$\frac{I_s}{I_0} = \frac{constant}{\lambda^4} \tag{3.2}$$

The scattered light is polarized, which relates to the nature of light [Lynch and Livingston, 1995]. The directional characteristics also depends on the size and form of the scattering particles [Nassau, 2001].

Examples for Rayleigh scattering are the color of the sky and the color of the sun perceived from earth's surface. The distance sunlight travels through atmosphere determines the parts of light which are scattered by air molecules. If the distance is short, short-wave light is scattered out and leads to the blue color of the sky and the light yellow colored sun. The color of the sun results from the absence of blue in the direct light from the sun [Nassau, 2001].

At sunrise and in the evening the distance is much longer and thus causes scattering of longer wavelengths. This results in the characteristic red color of the sun. Figure 3.2 illustrates atmospheric scattering and the resulting color of the sky and the sun. Ultraviolet light is the strongest scattered part of sunlight, thus scattering is the reason why the risk for sunburn is highest at noon and lower in the evening [Nassau, 2001].

Mie scattering

If the particle size is around the wavelength ($d \approx \lambda$), than Mie scattering occurs instead of Rayleigh scattering. This type of scattering was discovered by Gustav Mie (1868-1957) and takes place when spherical particles with sizes between $d < \lambda$ and $d < 2\lambda$ scatter electromagnetic waves [Nassau, 2001]. Mie scattering appears with a much stronger forward component than Rayleigh scattering. The scattering behavior itself is very complex to calculate, because it is influenced by the geometry and refractive index of the scattering particles. When scattering particles are relatively large only so called white scattering will appear, which can be seen in nature as the white of fog and low clouds. Small water droplets scatter the light and create these phenomena [Nassau, 2001].

3.4 Interference and diffraction

When two light rays run closely parallel to each other they may interfere with each another. This **interference** is caused by light's electromagnetic wavelike behavior (see figure 3.3) and may result in constructive enhancement or destructive refinement [Nassau, 2001].

Augustin-Jean Fresnel (1788-1827) conducted experiments which resulted in the so called interference patterns. Figures 3.4 and 3.5 show the design and exaggerated result of one of his experiments. He discovered that his experiment only worked with one light source to be used as

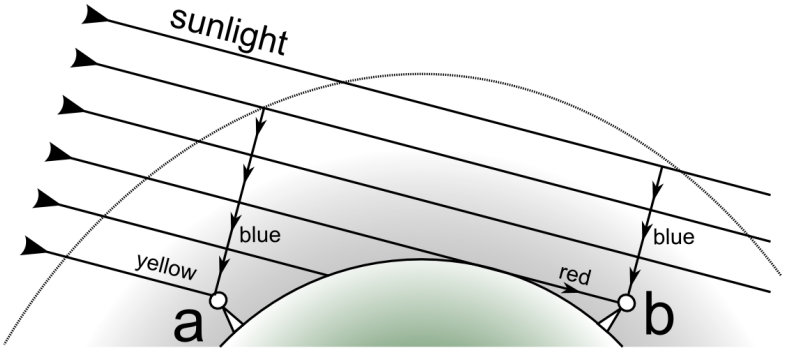

Figure 3.2: This image shows a schematic display of atmospheric scattering. The image contains two observers a and b. The atmosphere is represented by the space between the surface and the gray colored line. In case of observer a the sky appears widely blue because of the short distance light travels through the atmosphere. For observer a the sun looks yellow because of absence of blue in the direct sunlight. This changes for observer b. The direct light of the sun travels around a thirty two times [Lynch and Livingston, 1995] the greater distance through atmosphere than for observer a. This results in much more scattering which results in a red sun for observer b. The sky above her appears blue, but the nearer in the direction of the sun she looks the color changes to characteristic colors of the sky in the evening or at dawn. The red or yellowish color of the sun at these daytimes depends on the presence of small particles, like dust or ash, in the atmosphere [Nassau, 2001, Lynch and Livingston, 1995].

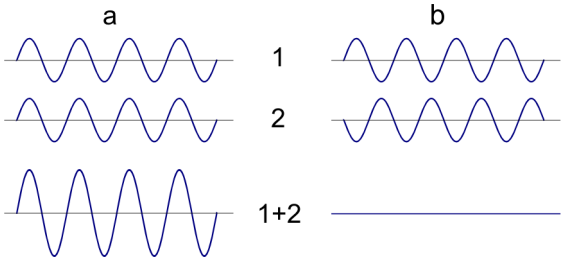

Figure 3.3: The image above shows two results for interference. Depending on the phase difference interference results in constructive enhancement a or destructive refinement b [Nassau, 2001].

source for the two rays of light. When he used an additional ordinary light source the experiment lead to no interference patterns. The reason for this is, that an additional light source contains more incoherent light than one split light source. The ability to use an additional light source for interference experiments changed with the development of lasers, which are capable to produce monochromatic coherent light [Nassau, 2001].

Examples for visible effects of interference are the rainbow colors on a soap bubble or on a thin film of oil on water, the coloration of bird feathers, or color-changing ink on bank notes. Figure 3.6 and 3.7 illustrate so called thin film interactions [Nassau, 2001].

Diffraction occurs when light is spread at the edges of an object. In any case it involves interference, although interference does not always include diffraction. An example for diffraction is an experimental setup consisting of a large plate with a small hole and a light source

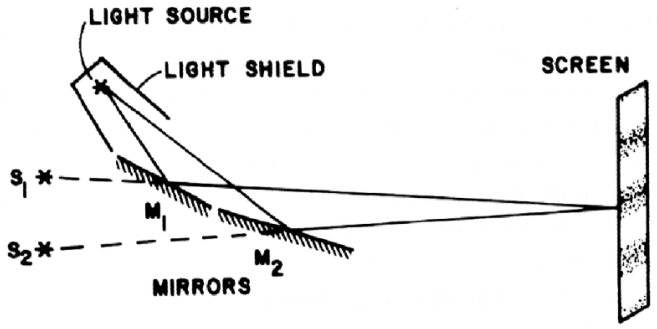

Figure 3.4: This image shows Fresnel's two mirror experiment. The used light source produces two light rays, s_1 and s_2, by redirecting them with two mirrors at a screen. There, the light rays lead to interference pattern, which are displayed in figure 3.5 [Nassau, 2001].

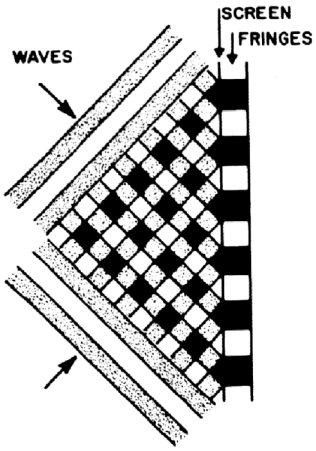

Figure 3.5: This image shows interference patterns produced by the waves of two light rays. The angle between the two light rays is exaggerated for displaying purposes [Nassau, 2001].

directed at the hole. The light passes through the small hole, but the edges of the hole cut into the wavefront of the incident light. This bends the light which results in a spreading of the light behind the plate. If a small screen is put behind the hole, an observer could observe diffraction patterns on the screen [Nassau, 2001].

Diffraction can be observed at the surface of the rear side of a CD or DVD, where small bumps or cavities cause diffraction, interference and the resulting characteristic play of color [Nassau, 2001].

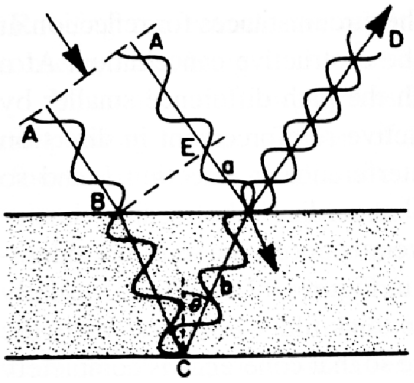

Figure 3.6: Interference. This image shows thin film interactions involving a material with constant thickness. There, the incident light rays are reflected from the back and front surfaces of the thin film. The phase difference in D is dependent on the material thickness and the refractive index of the material. When the incident light is white light, some wavelengths vanish and others remain. The resulting colors are called Newton's interference colors. Figure 3.7 shows a similar setup, but with varying material thickness [Nassau, 2001].

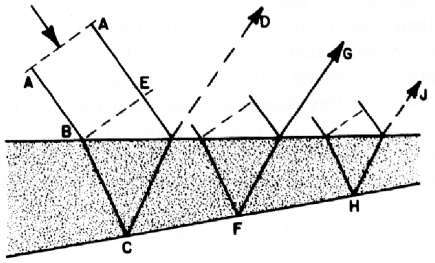

Figure 3.7: Interference. In contrast to the setup in figure 3.6 this setup results in different colors caused by varying material thickness. This illustrates how the play colors occurs on a soap bubble, because there, the thickness of the soap film is usually not uniform and changes over time [Nassau, 2001].

CHAPTER 4

Digital signal processing

Techniques of digital signal processing (DSP) allow to exactly reproduce signals, to save them without information loss and to reduce noise in these signals [Doblinger, 2007]. The focus of DSP is on the work with discrete time signals and their manipulation in discrete-time or frequency domain. In a discrete-time domain a signal is given as an indexed $x[n]$ list of real or complex numbers, which describe a signal sampled at a uniform sampling rate. A multiplication of two signals in discrete-time domain is defined as an element-wise product of the signal values [Hayes, 1999]. The formal definition of an infinite digital signal in discrete-time domain is [Oppenheim et al., 1999]:

$$x = \{x[n]\} \quad -\infty < n < \infty \tag{4.1}$$

Fourier transformations allow to represent signals as a weighted sum of sine waves in the frequency domain. For discrete time signals the discrete Fourier transformation is used instead [Blinn, 1989].

In DSP the two terms sampling and reconstruction are commonly used. Sampling usually describes the transition from a continuous signal to a discrete signal [Hearn and Baker, 2004]. Reconstruction describes the transition from a discrete signal to a continuous reconstructed signal [Mitchell and Netravali, 1988].

Usually a digital signal is obtained from an analog signal or another digital signal by sampling. In the theory of sampling and its applications the Nyquist criterion [Hearn and Baker, 2004] plays an important role. The violation of this criterion usually introduces errors related to the sampling process, which are known as aliasing.

Aliasing is distinguished in pre-aliasing due to poor sampling and the absence of a prefilter, and post-aliasing which results from to poor reconstruction [Mitchell and Netravali, 1988].

4.1 Nyquist criterion

Bryant [Bryant et al., 2009] summarizes the Nyquist criterion, which is very important in theory and application of sampling, in his book as follows:

> *"A signal with a maximum frequency f_a must be sampled at a rate $f_s > 2f_a$ or information about the signal will be lost because of aliasing. [...] Aliasing occurs whenever $f_s < 2f_a$ [...] To prevent aliasing, it can be shown that the signal must be sampled at least twice as fast as the highest frequency component."* [Bryant et al., 2009]

The Nyquist criterion is important for my work because in some cases (e.g. fluorescent light sources) spectral data contains very narrow signal peaks. This is not problematic if a method works with densely sampled data and a high number of sample points. However, this is seldom the case. Usually, to minimize calculation and storage costs the number of sample points is reduced. This reduction leads to a violation of the Nyquist criterion. Spectral data and working with them usually incorporates sampling rate changes. These involve reconstruction and thus may result in postaliasing of the data. Both sampling and reconstruction involve approximations, which may introduce further aliasing [Pharr and Humphreys, 2004].

To prevent further aliasing and to introduce no additional distortions, Oppenheim and Hayes [Oppenheim et al., 1999, Hayes, 1999] suggest to use the so called *ideal sampling rate conversion* (see figure 4.2 for details) involving a low pass filter for reconstruction of the signal. A low pass filter filters out high frequencies in the signal. This is done before resampling, otherwise the high frequencies would reappear after sampling as errors and distortions [Blinn, 1989].

However, the application of a low pass filter may not always reduce aliasing, and a wrongfully designed filter may blur the data and introduce new errors [Jähne, 2005]. Theoretically, ideal anti-aliasing requires to filter a function. Usually, the data is given as an indexed list of discrete sample points and not as a function in a time domain. This results in the fact that aliasing cannot be entirely removed [Blinn, 1989].

4.2 Digital filters

In DSP a digital filter is a time-discrete system which is defined by the application of a, usual mathematical, transformation T on any given input sequence $x[n]$ to create an output sequence $y[n]$ [Hayes, 1999]. This output may be smoothened, sharpened, noised-reduced, sampled or other wisely transformed. Formally this is given as [Oppenheim et al., 1999]:

$$y[n] = T\{x[n]\} \tag{4.2}$$

Any time discrete system is distinguished by its access of the input data (linear vs. nonlinear) and whether the system is time variant or time invariant [Doblinger, 2007]. Furthermore, in addition to these, a filter is also distinguished by its impulse response, which determines if it is a recursive filter (infinite impulse response, IIR) or its opposite (finite impulse response, FIR). A digital filter is classified by its frequency response, which describes which frequencies should be eliminated and which frequencies are preserved. There are four filter types:

- **Low pass filters.** Filters belonging to this type eliminate frequencies above a given cutoff frequency ω_c. Examples for low pass filters are given in section 4.5 [Hayes, 1999].

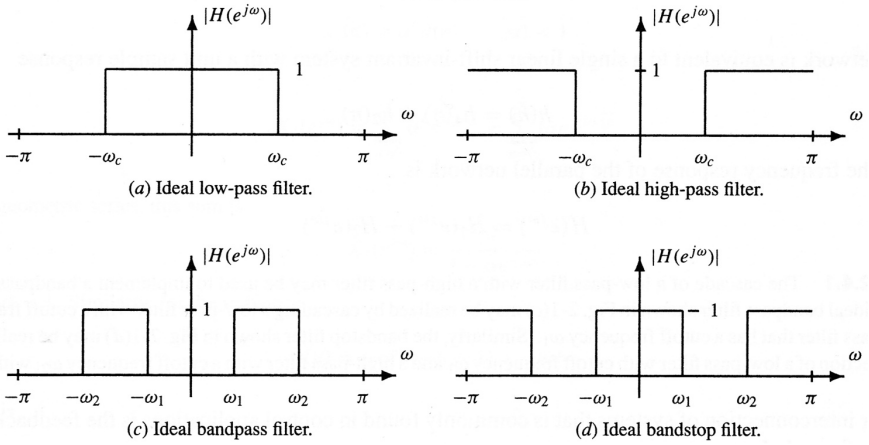

Figure 4.1: Frequency response of ideal filter types. The depicted filters are (a) ideal low pass filter, (b) ideal high pass filer, (c) ideal band pass filter and (d) ideal band stop filter. The cut-off frequency is denoted by ω_c (ω_1 and ω_2 are frequency band limits). Frequency responses of digital filters are obtained by applying Fourier analysis on the filters[Hayes, 1999]. Ideal filters have no limits, thus practical applications require a limiting (windowing) of any given filter [Theußl et al., 2000].

- **High pass filters.** These filters are the opposite of low pass filters. They eliminate frequencies below a given cutoff frequency ω_c [Hayes, 1999].
- **Band pass filters.** Filters of this type preserve frequencies within a given frequency range between ω_1 and ω_2 [Hayes, 1999].
- **Band stop filters.** This kind of filters eliminate frequencies within a given frequency range between ω_1 and ω_2 [Hayes, 1999].

The ideal versions of these filters are depicted in figure 4.1. Since ideal filters have no limits, practical application requires a limiting (windowing) of any given filter. This limiting causes distortions called leakage or ringing artifacts. To counter these filters where the filter function is multiplied with another function, are used [Theußl et al., 2000]. A windowed filter function is formally defined as in equation 4.3 [Hayes, 1999].

$$w[n] = \begin{cases} f[n] & if\ n\ within\ window\ borders \\ 0.0 & if\ n\ not\ within\ window\ borders \end{cases} \tag{4.3}$$

Filters, such as a filter window $w[k]$, are applied on time-discrete data by applying it by convolution [Hayes, 1999]. The convolution theorem is hereby defined as

$$y(n) = \sum_{k=0}^{N-1} w(k)x(n-k) \tag{4.4}$$

or generally as

$$y(n) = w(k) * x(n) \tag{4.5}$$

In frequency space this convolution is equivalent to a multiplication of the Fourier coefficients of the filter and the signal [Hayes, 1999].

The selection of a suitable filter follows the goals of digital filter design which aims to select a filter with a good frequency response, and to set the window $w[k]$ as small as possible. The realization of these goals is always a compromise between these contradicting aims [Oppenheim et al., 1999].

Moreover, the selection of a suitable filter is usually also a trade off between sharpness (preservation of details) and introduction of further aliasing [Turkowski and Gabriel, 1990]. The application of filters may result in one or more of the following artifacts: sample-frequency ripple, anisotropic effects, ringing, blurring and aliasing [Mitchell and Netravali, 1988]. These are in detail:

- **(Post)aliasing** is aliasing due to spectral leakage, i.e. high frequent data reappears as distortions in low frequency regions in a reconstructed signal [Pharr and Humphreys, 2004].
- **Sample frequency rippling** is a form of post aliasing which appears as an oscillation of distortions at the sampling frequency in a reconstructed signal [Marschner and Lobb, 1994].
- **Blurring or smoothing**, although desired to reduce signal noise, may also eliminate desired details of the signal [Marschner and Lobb, 1994].
- **Ringing and overshooting** result from discontinuities and manifest themselves as echoes of the signal throughout the reconstructed signal [Mitchell and Netravali, 1988].
- **Anisotropy** describes the asymmetric behavior of the absence or presence of artifacts in the reconstructed signal [Marschner and Lobb, 1994].

As mentioned, it is possible to transfer discrete data from time domain to frequency domain by applying the discrete Fourier transformation (DFT).

4.3 Discrete Fourier transformation (DFT)

The discrete Fourier transformation (DFT) transforms $(N + 1)$ uniform sampled points to frequency space and utilizes the approximation formulas given in equations 4.6 and 4.7 [Bronstein et al., 2001, p.951].

The resulting spectral coefficients, which otherwise would result in a complex number, are split in to a_k, which represents the real part, and b_k, which represents the imaginary part. The value of a_0 always results in the mean average of x_v and b_0 is always equal 0.0.

$$x_v = vh \;\; (v = 0, 1, \ldots, N), \;\; h = \frac{2\pi}{N} \tag{4.6}$$

$$DFT : a_k \approx \tilde{a}_k = \frac{2}{N}\sum_{v=1}^{N} f(x_v)\cos kx_v,\ \ b_k \approx \tilde{b}_k = \frac{2}{N}\sum_{v=1}^{N} f(x_v)\sin kx_v\ \ (k = 0, 1, 2, \ldots, n) \tag{4.7}$$

To transform the Fourier coefficients a_k and b_k back to a discrete time domain the inversion of the Fourier transformation (IDFT) has to be applied. The trigonometric polynomials in equation 4.8 and 4.9 depict this inversion.

The first polynomial in equation 4.8 transforms the data from frequency space to discrete-time domain and satisfies the interpolation condition $\tilde{g}_1(x_v) = f(x_v)$ [Bronstein et al., 2001]. This condition ensures that the IDFT returns the original values at the positions of the initial values when all spectral coefficients are used.

$$IDFT : \tilde{g_1}(x) = \frac{1}{2}\tilde{a}_0 + \sum_{k=1}^{n-1}(\tilde{a}_k \cos kx + \tilde{b}_k \sin kx) + \frac{1}{2}\tilde{a}_n \cos nx \tag{4.8}$$

For the application in this work, the second polynomial given in equation 4.9 is used since it transforms the data under the condition of minimal quadric error. The variable N denotes the number of initial data points used for calculation and n the number of used spectral coefficients. This polynomial approximates the function $f(x)$ under the requirement of $m < n$ and results in a minimal quadric error sum [Bronstein et al., 2001] of the reconstructed values at any position.

$$IDFT : \tilde{g_2}(x) = \frac{1}{2}\tilde{a}_0 + \sum_{k=1}^{m}(\tilde{a}_k \cos kx + \tilde{b}_k \sin kx) \tag{4.9}$$

Working with data in frequency space is different than in the time domain. For example, the convolution with a filter is done by multiplication in frequency space [Hayes, 1999]. Representation methods for spectral data like Sun's composite method [Sun et al., 1999, Sun, 2000] utilize the frequency space to compress the spectral information. However, using data in frequency space incorporates that transformations between the two domains are necessary. The growth of complexity of a DFT or IDFT is $O(n^2)$ which means that in order to perform a DFT or IDFT on a given signal the effort to do so increases at a quadric rate, which is not very efficient.

Sophisticated improvements such as the fast Fourier transformation speed this up to $O(nlog(n))$ [Bronstein et al., 2001]. In addition Libraries for implementation such as the FFTW present convenient ways to speed up Fourier transformations [Frigo and Johnson, 2009].

4.4 Ideal sampling rate conversion

Ideal sampling rate conversion reduces postaliasing introduced by resampling [Hayes, 1999]. Its concept, which incorporates the usage of a low pass filter for reconstruction, is depicted in figure 4.2.

For example, a spectral data set consists of $N = 42$ entries. The desired data set should consist of $N^* = 16$ entries. To introduce no further aliasing the sampling rate should be changed

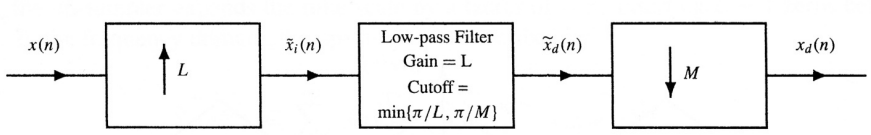

Figure 4.2: Ideal sampling rate conversion by an integer factor L/M to avoid further introduction of aliasing. This is done in three steps. The first step is to increase the sampling rate (up-sampling) by an integer factor L. The second step is to reconstruct the data with a low pass finite input response filter (FIR-filter) with a cutoff-frequency $\omega_c = \min(\pi/L, \pi/M)$. The last step is to decimate (down-sampling) the data by an integer factor M [Hayes, 1999].

by

$$\frac{L}{M} = \frac{N}{N*} = \frac{16}{42} = \frac{8}{21} \tag{4.10}$$

The usage of a rational factor composed of two integers prevents further distortions in the resampling process which would otherwise be introduced by the approximation of decimal numbers [Hayes, 1999]. The used low-pass filter should have a cutoff frequency ω_c, given by

$$\omega_c = \min(\frac{\pi}{L}, \frac{\pi}{M}) = min(\frac{\pi}{8}, \frac{\pi}{21}) = \frac{\pi}{21}. \tag{4.11}$$

Up-sampling and down-sampling work as shown in the following pseudo code snippet:

```
1 x[N]; // source data
2 u[M*L]; // up-sampled data
3
4 for(i=0; i < N; i++) u[i * L] = x[i]; // up-sampler
5
6 ... // interpolate u by low pass filter
7
8 y[L]; // down-sampled data
9 for(i=0; i < M; i++) y[i] = u[i * M]; // down-sampler
```

The following section is concerned with the low pass filters used in this work.

4.5 Low pass filters

The ideal low pass filter is the $sinc$ filter. It is characterized by the $sinc$ function, defined as [Oppenheim et al., 1999]:

$$sinc(x) = \begin{cases} if \;\; x \neq 0.0 & \sin(\pi x)/(\pi x) \\ if \;\; x = 0.0 & 1.0 \end{cases} \tag{4.12}$$

In practice, windowed filters are used instead of the ideal filter. A windowed filter is defined by its filter window width and the filter extent. The first describes the range where the filter function returns other values than 0.0 (as in equation 4.3) and the second defines how many values affect the filter result [Pharr and Humphreys, 2004].

The following subsections contain the filters used in this work for filtering spectral data. All filter functions are displayed as a function of $w(x, ...)$ which is a small deviation from the representation of the filter functions in the used literature. A plot containing all filter functions used in my work is given in figure 4.3.

Box filter

The box filter also known as rectangular window is a simple filter which weighs all samples by the same weight 1.0. Although it is the worst possible filter since it introduces strong spectral leakage, it's very fast and easy to implement. Thus it is the favorite used filter for testing purposes. The window function w is given in equation 4.13 [Theußl, 1999]. τ describes the width of the window.

$$w(x,\tau) = \begin{cases} 1 & if\ |x| \leq \tau \\ 0 & else \end{cases} \tag{4.13}$$

Tent filter

The tent filter or Bartlett window [Theußl, 1999] is an advancement to the box filter. It weighs the samples by different weights, depending on the distance to the center of the filter. The nearer a sample is to the center the higher is its weight [Pharr and Humphreys, 2004]. The window function w is given in equation 4.14 [Theußl, 1999]. τ describes the width of the window.

$$w(x,\tau) = \begin{cases} 1 - \frac{|x|}{\tau} & if\ |x| < \tau \\ 0 & else \end{cases} \tag{4.14}$$

Welch filter

The Welch window is similar to the Bartlett window. The difference between the two window functions is that the weights in the Welch window fall off at a quadric rate [Theußl, 1999]. The filter function is given in equation 4.15 [Theußl, 1999].

$$w(x,\tau) = \begin{cases} 1 - \left(\frac{x}{\tau}\right)^2 & |x| < \tau \\ 0 & else \end{cases} \tag{4.15}$$

Gaussian filter

The Gaussian filter is well known in image processing and utilizes the Gaussian bell curve. This filter usually blurs the results. For implementation it is important to subtract the values at the window borders from the filter function which is performed by the term $-\exp^{-\alpha w^2}$ in equation 4.16. This is done to satisfy the condition that a windowed filter has the values of 0.0 at its borders. The value α controls the decline of the bell curve. A small α results in a small decline and a high α in a strong decline of the curve. The filter function is given in equation 4.16 [Pharr and Humphreys, 2004, p.358].

$$w(x, \alpha) = \exp^{-\alpha x^2} - \exp^{-\alpha w^2} \tag{4.16}$$

Mitchell filter

This filter is based on the general form of the symmetric cubic filter with six parameters of freedom. Mitchell and Netravali [Mitchell and Netravali, 1988] reduced the six parameters to two parameters b and c to control the appearance of the filter. The filter equation is given in equation 4.17. Parameters in the filter function of $b = 1.0, c = 0.0$ lead to an appearance as a cubic B-spline, $b = 0.0$ to cardinal cubics and $b = 0.0, c = 0.5$ to the Catmull-Rom spline.

This filter may contain negative filter values which usually sharpen the result. Mitchell and Netravali suggest parameters which satisfy $b + 2c = 1$ to make a compromise between ringing, blurring, and anisotropy [Mitchell and Netravali, 1988]. They also provided a two-dimensional plot of subjective observations, how the filter behavior changes with different values of b and c. This plot is given in figure 4.4.

$$w(x, b, c) = \frac{1}{6} \begin{cases} (12 - 9b - 6c)|x|^3 + (-18 + 12b + 6c)|x|^2 + (6 - 2b) & |x| < 1 \\ (-b - 6c)|x|^3 + (6b + 30c)|x|^2 + (-12b - 48c)|x| + (8b + 24c) & 1 \leq |x| < 2 \\ 0 & otherwise \end{cases} \tag{4.17}$$

Lanczos filter

The Lanczos filter is based on the $sinc$ function. It offers one of the best trade offs between preserving details, introduction of aliasing and ringing artifacts [Turkowski and Gabriel, 1990]. The filter function is given in equation 4.18 [Pharr and Humphreys, 2004]. The variable τ determines the number of $sinc$ cycles.

$$w(x, \tau) = \frac{\sin \pi x/\tau}{\pi x/\tau} \tag{4.18}$$

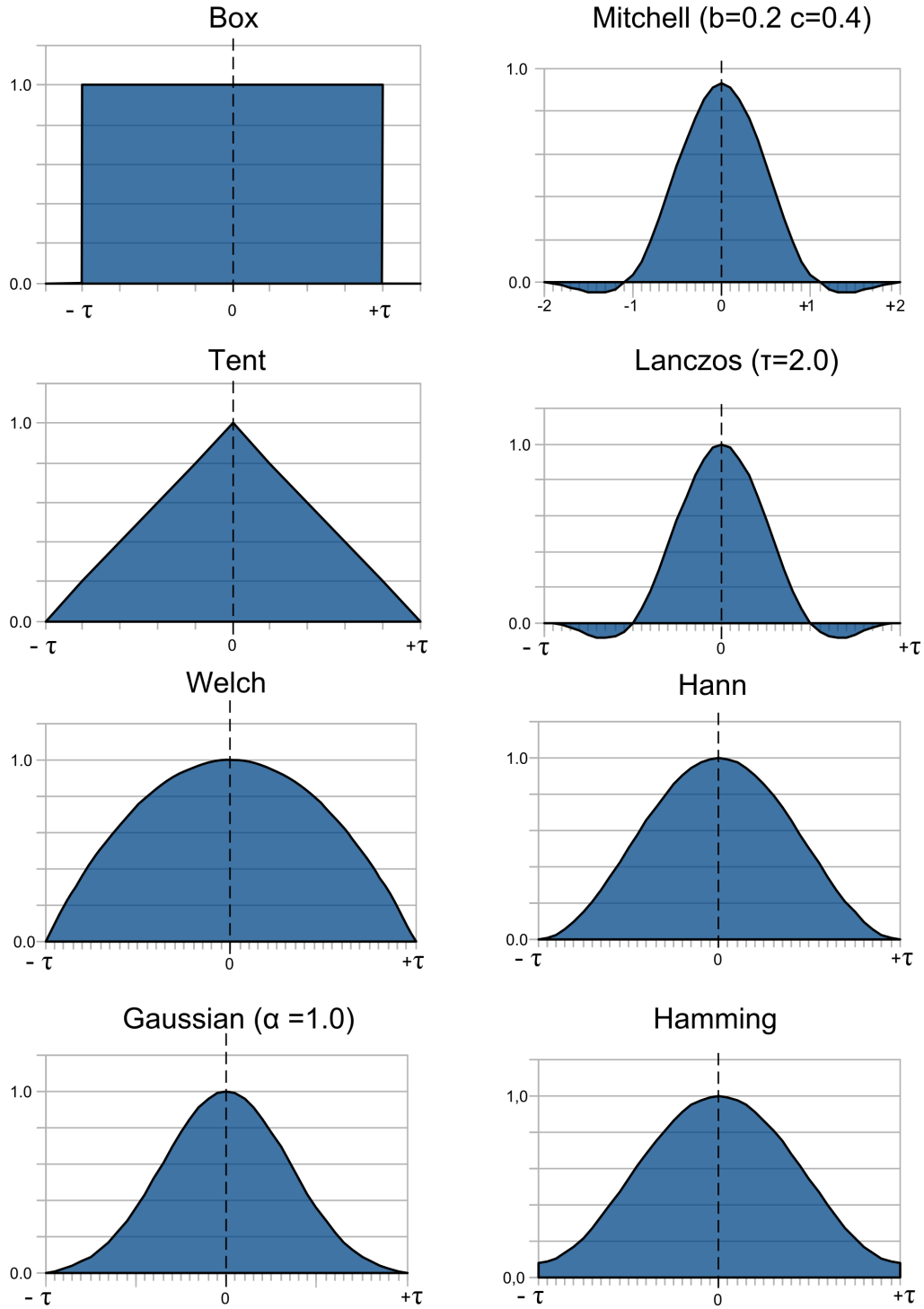

Figure 4.3: This image is an overview over all low pass filters used in this work.

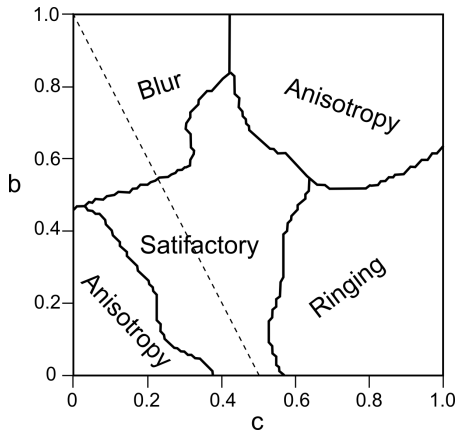

Figure 4.4: Two dimensional plot of the subjective parameter behavior of the Mitchell filter. [Mitchell and Netravali, 1988]. The plot shows the dominance of different filtering errors for given parameters b, c and the parameter line $b + 2c = 1$.

Hann and Hamming filter

Both filters are based on the cosine and are similar to each another, but differ in the value α. The Hann window utilizes $\alpha = 0.5$ and the Hamming window sets this value to $\alpha = 0.54$. The latter is problematic because it introduces discontinuities at the borders of the window function [Theußl, 1999].

$$w(x, \alpha, \tau) = \begin{cases} \alpha + (1 - \alpha) \cos\left(\pi \frac{x}{\tau}\right) & |x| < \tau \\ 0 & else \end{cases} \tag{4.19}$$

CHAPTER 5

Rendering of wavelength dependent effects

Realistic image synthesis ultimately aims at generating realistic images with the computer. An observer should experience the same visual sensation when she looks at the real scene and the computer generated image. In other words, the synthetic image has to be physically and perceptually equivalent when compared with reality [Devlin et al., 2002]. To achieve this aim the following aspects have to be addressed: representation of geometric shapes, generation of authentic colors and kinetics, and the correct simulation of natural phenomena [Sun, 2000]. Natural phenomena include interactions between lights, materials and surfaces. A subgroup of the natural phenomena is defined in physics by spectral interactions, thus they are named *spectral effects* [Devlin et al., 2002].

Methods of digital image synthesis can be described by Kajiyahs rendering equation (equation 5.1). The equation defines the visual sensation of an observer by the propagation of light rays and basically summarizes how much light a point x receives from another point x'. Furthermore the equation includes the influence of light transport from a third point x'' to x'. A digital image synthesis method is defined by the parts of the rendering equation that can be satisfied. Methods like path tracing satisfy the whole equation. The rendering equation does not cover spectral effects [Kajiya, 1986], but is useful to demonstrate the concept behind digital image synthesis.

$$I(x, x') = g(x, x') \left[\epsilon(x, x') + \int_S \rho(x, x', x'') I(x', x'') dx'' \right] \tag{5.1}$$

The rendering equation is composed of the following terms [Kajiya, 1986]:

- $I(x, x')$. This term is called *transport intensity* and describes the intensity of the light which reaches point x and originates from point x'.

- $g(x, x')$. This geometry term represents *shadowing* by geometry. If the direct path from point x to point x' is blocked than this term is equal to 0.0.
- $\epsilon(x, x')$. This term describes how much intensity point x' emits in the direction of point x.
- $\rho(x, x', x'')$. This *scattering term* represents how much light is scattered from point x'' at point x' and, from there, to point x.
- Integration $\int_S$ describes a summation of all light which is scattered at point x. This covers all influence on the intensity at an hemisphere at point x. This integral and further recursive invocations of $I(x', x'')$ are the reasons why analytic solutions are not suitable to solve the rendering equation.

Ray tracing (also known as Whitted ray tracing) does not satisfy all parts of the rendering equation, does not include integration, and the recursion exists only for perfect mirrors. Ray tracing has not the capabilities to produce diffuse reflections or to provide satisfying specular reflections from less glossy surfaces [Whitted, 1980]. Ray tracing is reviewed in the next section because it is often used and very useful to provide an overview on how digital image synthesis actually works.

5.1 Whitted ray tracing

This method was introduced by Turner Whitted [Whitted, 1980]. The concept of ray tracing is basically an inversion of an other approach of digital image synthesis. This approach was to track the light starting from the light source throughout the scene. This is a very costly approach because the majority of the light in a scene is not reaching the observer. Whitted took the inversion of this approach and added his main contribution, a recursion, which resulted in Whitted's ray tracing. In Ray tracing, as the name indicates, rays are emitted by the observer and are tracked throughout a scene. At each position where a ray hits an obstacle the influence of existing light sources and geometric shadowing are taken into account. Whitted ray tracing does not end there, because the reflected rays are traced again like the initial ray. This recursive process only ends for a particular ray if it hits no other obstacle. The list which contains the reflection path for each ray is stored in a tree structure. This tree structure is used to calculate the ray/surface intersections which finally result in the intensity (color) of one ray that reaches the observer. The composition of one point in this tree structure is given in equation 5.2 [Whitted, 1980].

$$I = I_\alpha + k_d \sum_{j=1}^{j=ls} (\bar{N} \cdot \bar{L_j}) + k_s S + k_t T \tag{5.2}$$

In equation 5.2 I stands for the intensity which is calculated for one point in the tree structure of the ray/object intersections and is composed of [Whitted, 1980]:

- The intensity of ambient light I_α

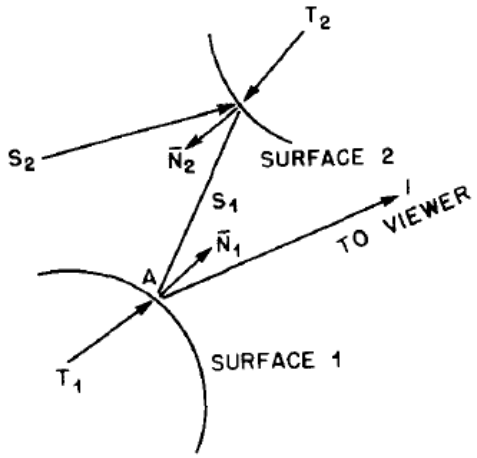

Figure 5.1: The image shows an example of ray/surface interactions which influence the intensity reaching the observer I from point A. Every $\bar{N}$ denotes the normal vector of a surface. The rays labeled T describe the contribution of intensity to the final ray resulting from transmission (refraction). The rays contributing intensity from specular reflection are labeled S. The displayed path is finally resolved in each point as a recursion 5.2 [Whitted, 1980].

- The intensity contributed by each light source $\bar{L_j}$ calculated by the dot product of the ray in direction to the light source $\bar{L_j}$ and the normal vector of the surface $\bar{N}$. Shadowing in direction $\bar{L_j}$ is taken into account.

- The intensity composed of the specular reflection in direction of the direct mirror reflection S.

- The intensities T_1,T_2 contributed by transmission (refraction) through the material.

- The coefficients k_d, k_s, k_t represent diffuse, specular and transmissive material properties which usually incorporate an approximation to physical laws.

The procedure is repeated for all pixels of the target image [Whitted, 1980]. Usually more than one ray is calculated for a pixel to improve image quality by utilizing digital filtering techniques [Shirley and Morley, 2003]. Usually, before a hit of a ray at a surface is processed, a ray is checked against a bounding volume to prevent unnecessary calculations. These bounding volumes are usually realized as spheres. These spheres are also used for the testing of a ray tracing application, because there, the surface and the bounding volume are the same object [Shirley and Morley, 2003].

For the interested reader I suggest the books of Shirley [Shirley and Morley, 2003], Glassner [Glassner, 1995a], [Glassner, 1995b] and Pharr [Pharr and Humphreys, 2004], which describe the implementation of a working rendering application and its backgrounds.

The result of any rendering technique is, in a manner of speaking, the color for each pixel of the resulting image. The way in which colors are represented in a rendering application determines its capabilities in regard to spectral effects.

5.2 Traditional color representation

Traditionally colors are represented by a RGB triplet of color values of the basic colors red, green and blue. These represent colors by 8 bit per color. This representation is also used to define light sources and material properties. Color calculations are multiplications of two RGB triplets, which is very fast and easy to implement [Johnson and Fairchild, 1999].

However, the RGB representation has the major disadvantage that color information is already lost before any calculation or conversion process has taken place. The reason for this loss is shown in the science based representation of colors as spectral data: colors relate to an electromagnetic energy distribution along the visible spectrum [Nassau, 2001] with usually considerable more than three sample points. Thus, three values cannot accurately represent this distribution. A further disadvantage of this representation is that RGB is device dependent, which means that the result varies from device to device (e.g. ink jet printer vs. TFT-flat panel). Moreover, RGB cannot display metamerism, because two objects with the same RGB values will look the same under any lighting condition [Peercy, 1993].

These disadvantages of RGB [Peercy, 1993] lead to the necessity of a more accurate color representation where colors also retain their spectral properties. A representation which satisfies this allows to reproduce spectral effects like fluorescence, diffraction, dispersion and polarization [Devlin et al., 2002].

Rendering applications using spectral representation methods are called “spectral rendering” which encompasses realistic image synthesis and spectral data representations [Devlin et al., 2002].

To illustrate how the representation of colors influences the resulting image and how spectral color representation improves the result, I rendered two scenes with the rendering toolkit ART, developed at the technological university of Vienna at the institute for computer graphics and algorithms.

The first scene in figure 5.3 shows a rendering of the sky from a fish-eye perspective. The image on the right of figure 5.3 shows the difference between traditional RGB representation and a spectral representation. The second scene given in figure 5.2 shows a chessboard scene under the influence of sky color and refraction.

5.3 Spectral rendering

In a spectral rendering application colors have to be represented by their spectral distribution or a compressed form of it. The main issue is to represent high frequent details like narrow signal peaks [Devlin et al., 2002]. This leads to a trade-off between speed and accuracy.

Techniques such as point sampling for instance, with sampling intervals higher than 10nm and base functions have problems with narrow signal peaks which usually occur in fluorescent

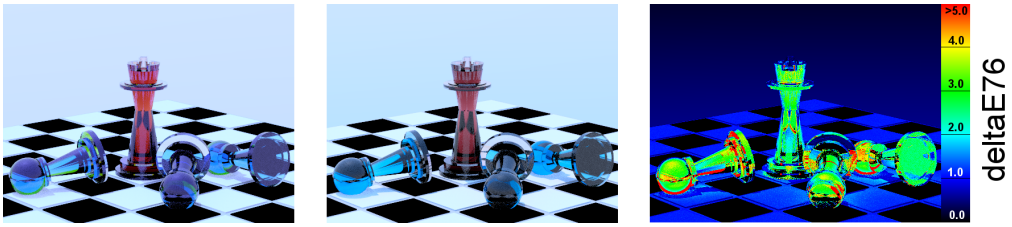

Figure 5.2: Different color representations: chessboard under skylight. The image shows three different images of the same scene. The image on the left was rendered utilizing RGB color representation. The image in the middle was rendered using a spectral data representation by Riemann summations (see section 5.4 for details). The image on the right shows the color difference between the two rendered images according to the color difference formula CIE76. The largest difference occurs within the refraction in the material of the chess figures. A noticeable difference is visible as a green hue in the refraction of the pawns.

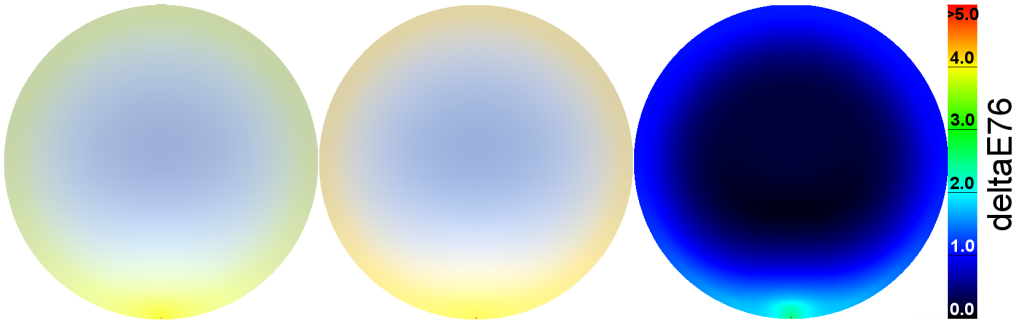

Figure 5.3: Different color representations: sky dome rendered from a fish-eye perspective. The image shows three different images of the same scene. The image on the left was rendered utilizing RGB color representation. The image in the middle was rendered using a spectral data representation by utilizing Riemann summations (see section 5.4 for details). The image on the right shows the color difference between the two rendered images according to the color difference formula CIE76. The largest color difference occurs at the horizon and around the sun and is visible as more greenish hue at the horizon in the left image.

light sources because of aliasing due to the violation of the Nyquist theorem. Adaptive or composite methods try to deal with this issue [Devlin et al., 2002].

Several spectral data representation methods exist in literature. They can be distinguished by their underlying concept:

- **Time-space based methods.** Methods utilizing this concept represent a spectra by discrete points sampled at a constant (uniform) sampling distance [Peercy, 1993]. Point sampled spectra or Riemann summations belong to this category. Riemann summations differ from point sampling because they represents the spectral data by partial integrals of its wavelength distribution [Peercy, 1993].

- **Linear methods.** This approach utilizes a different representation of the data. Methods like principal component analysis (PCA) are used to obtain the eigenvectors and eigen-

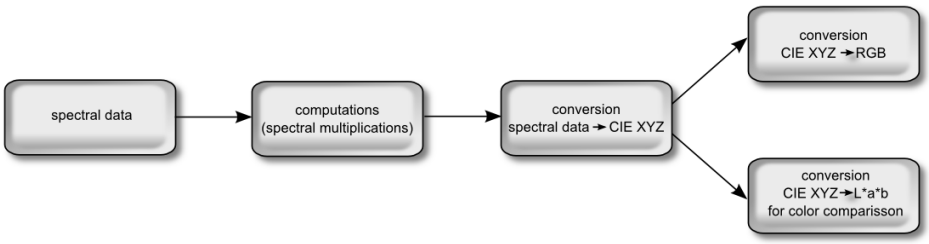

Figure 5.4: Abstract rendering pipeline based on [Hall, 1999]. This pipeline includes all steps of a rendering pipeline, but neglects all calculation steps related to the presence of geometry.

values of the data [Johnson and Fairchild, 1999] [Peercy, 1993]. These base functions are used to represent the spectral data.

- **Frequency-space based methods.** Methods utilizing this approach represent spectral data in frequency space by its Fourier coefficients which are obtained by Fourier transformation [Sun et al., 1999].

- **Composite methods.** This approach was introduced by Sun [Sun, 2000] and separates the representation of spectral data into two representation tasks. The first task is the representation of the non problematic distribution of wavelengths (low frequent signals) by a compact method. The second task is the separate representation of the problematic details (high frequent signals) [Sun et al., 1999] [Sun, 2000].

In addition to these methods, it is also important to distinguish between the purpose of the chosen representation method. Methods may be designed for accurate data storage, fast calculation speed or compactness. In my work I focus on representations concerning rendering applications and their accuracy. I assume that the abstract rendering pipeline given in figure 5.4 represents a rendering application.

5.4 Methods for spectral rendering and their challenges

In this section spectral data representation methods are described in detail.

Point sampling

Point sampling utilizes linear interpolation techniques for reconstruction and is a widely used method to represent signals or sequences of numerical data. The data is represented by uniform sampled values in an indexed list. Every sampled value in this list represents an energy level at a discrete wavelength.

Sun [Sun, 2000] summarizes the advantages and disadvantages of this method in his work. He states that the method is very accurate if the sample interval is kept low and thus a lot of points are used for data representation. Precision usually decreases when a higher sample interval is used.

However, the method has severe problems when representing fluorescent data, because the width of the details occurring in this data is usually far below the sampling distance [Sun, 2000]. According to digital signal processing theory [Oppenheim et al., 1999] this introduces errors due to aliasing because the width of the details violates the Nyquist criterion (see section 4.1 for details). Therefore, Sun concludes that point sampling is very useful for non-fluorescent data [Sun et al., 1999].

Peercy [Peercy, 1993] claims that a low number (around 4) of data points is sufficient for basic applications. However to accurately represent detailed data, a high number of sampling points is needed. The storage costs of this method corresponds to the number of used sample points. Calculation complexity grows linearly with the number of used sample points ($O(n)$) which is very efficient.

Multiplication of two spectra represented by point sampling is realized as an element wise multiplication of two arrays containing the sample points. This assumes that the sample points are given in the same sampling distance and scope. If the data differs in these properties it is necessary to apply sampling rate conversions and clipping to adjust the data.

Riemann summations

This method is similar to point sampling, but each point in the indexed list representing a spectrum is a partial sum of the area under its spectral distribution. Riemann summations are calculated by numeric integrals in a constant interval [Bronstein et al., 2001]. Each integral can be expressed as in equation 5.3. y describes the partial sums within an interval $[a, b]$. The function f represents the spectral distribution usually given as an indexed list of discrete values.

$$y = \int_a^b \bar{f}(\lambda) d\lambda \approx \Delta\lambda \sum_a^b \bar{f}(\lambda) \tag{5.3}$$

The limits of the summation a and b are given by the selected sampling distance. Peercy [Peercy, 1993] claims that Riemann summations with $n+2$ partial sums are accurate for the first $2n + 2$ spectral coefficients of the corresponding Fourier series of the data. A given function can be well described with Riemann summations if only a small number of spectral coefficients is sufficient to describe the function. This means, that a given spectrum only consists of low-frequent data, which is the case with most non-fluorescent data were no high frequent details are present. Peercy [Peercy, 1993] claims that about 4 partial sums are normally sufficient. The storage costs and calculation complexity growth are similar to point sampling.

Riemann summations are not very well suited to handle complex cases, such as fluorescent data. The presence of high frequency data requires more spectral coefficient to accurately represent it by its Fourier coefficients. However, Riemann summations work better in these cases than point sampling, because they better preserve the area under the function of the data which has a major impact on accuracy. This can be seen when formulas for the calculation of XYZ are taken into account where the area under the function plays an important role as the result of the integrals in equation 2.11, 2.12 and 2.13. Multiplication of two spectra represented by Riemann summations is realized in the same way as in methods utilizing point sampling as an element wise multiplication of two arrays containing the partial sums.

Linear methods

These methods are called linear methods because they combine linear sets of base functions (eigenvectors and eigenvalues) to represent a spectra [Devlin et al., 2002]. In this work, they are not used, because they require prior knowledge of the interactions of light and surfaces occurring in a scene [Peercy, 1993] and are only mentioned for completeness.

The mentioned prior knowledge is built by drawing a set of base functions together. These base functions describe any spectral distribution occurring in the scene, even spectra resulting from interactions of light, materials and surfaces. This set is calculated utilizing predictions about spectral interactions in the scene. Peercy [Peercy, 1993] clearly states, that this method is only effective if prior knowledge of a scene is available for base function candidate generation. Principal component analysis (PCA) or singular value decomposition (SVD) is used to calculate the eigenvalues and eigenvectors needed. The eigenvectors represent the base function of the initial data set and are used according to their eigenvalue. This is done by starting to eliminate the eigenvectors with the smallest eigenvalues because they are less important to reproduce the data than eigenvectors with a high eigenvalue [Stahel, 2008]. The number of used eigenvectors (dimensions) can be reduced to achieve data compression.

Frequency-space based methods

Fourier transformations allow the representation of time-discrete data in frequency space as a summation of sine waves. This representation can also be utilized to represent a spectrum. Usually spectral data is stored as spectral coefficients which are split into their real and complex parts. Sun [Sun, 2000] utilizes frequency-space based methods to represent low frequent data of any given spectrum for the use with his composite method. Frequency-space based methods introduce the need to calculate Fourier transformations and their inversion.

To obtain the Fourier coefficients from discrete data, calculation methods such as discrete Fourier transformation (DFT) or fast Fourier transformation (FFT) have to be used [Bronstein et al., 2001]. The DFT method has been introduced in chapter 4 in section 4.3.

Composite methods

The basic idea of composite methods is to divide the spectral data into low frequent and high frequent data. This idea was introduced by Sun [Sun, 2000]. He argues that the smooth data can be easily handled by representing it with a small amount of Fourier coefficients of its Fourier series. The high frequent data, which contains the peaks or in other words, regions with a high relative gradient, is represented by point sampling.

The division of the spectral data into a low and a high frequent signal can be achieved by dividing it according to the relative slope between two sample points. This is done by processing the initial spectral data, which is usually available as detailed point sampled data. Deville and his colleagues [Deville et al., 1994] describe an algorithm for this division in their work. Their spectrum segmentation algorithm splits peaks from the spectrum if a certain gradient threshold is reached. The result of this algorithm is a low frequent spectral distribution and a separate list containing high frequent details.

Sun's spectral data representation [Sun et al., 1999] requires that multiplication of spectra must happen in the discrete-time domain as point-sampled data (by applying IDFT on the Fourier representation) [Sun et al., 1999]. This is due the fact that working at the same time in discrete-time and frequency domain is too complex. The re-sampling leads to an increased growth of complexity of this method of $O(n^2)$ [Sun, 2000]. The calculation effort of this method can be reduced by utilizing the Shannon theorem for the re-sampling distance. There are also faster implementations of the Fourier transformation such as FFT (Fast Fourier Transformation) [Bronstein et al., 2001] which reduces the growth of complexity to $O(n \log n)$. The limitation of FFT is that it can only work with a number of sample points of the initial time-discrete data equal to numbers with the basis of two.

By separating the problematic peaks from the smooth data this method can accurately represent fluorescent data with only a few values and is therefore a very compact method. However, the calculation costs are high compared to other methods. The storage costs of this method is composed of the used spectral coefficients a_k and b_k and the costs for the storage of the high frequent data which depends on the spectral data.

Testing spectral data representation methods

In literature, different setups are used for testing spectral data representation methods. Peercy [Peercy, 1993] uses a test scene with 4 different colors chosen from the Gretag Macbeth color checker [McCamy et al., 1976]. He evaluated the different methods with a small selection of fluorescent light sources.

Sun and his colleagues [Sun et al., 1999] compared their method with standard RGB representations. They also utilize a comparison between their method and point sampled representations. Their data set consists of a selection of 7 colors from the Gretag Macbeth color checker [McCamy et al., 1976] and several different fluorescent light sources.

Rougeron and Péroche [Rougeron and Péroche, 1997] use the CIE standard illuminants C, D65, F2 with the whole twenty-four colors from the Gretag Macbeth color checker.

Sun [Sun, 2000] tests his method in his work using the whole Gretag Macbeth color checker and the light sources CIE A,B,C and CIE F1-F12. All of the tests have in common that they utilize all, or a reduced number, of colors obtained from the Gretag Macbeth color checker.

In my work two sets of initial test data are used. The first consists of spectral data from the CIE standard illuminants and is composed of the spectral data of the standardized light sources A,B,C and F2,F7,F11. According to [HunterLAB, 2005], this reduced data-set is sufficient for testing methods in color applications. The data of the CIE illuminants was sampled at a distance of 5.0nm for wavelengths ranging from 380nm to 780nm.

The second data set consists of spectral data taken from the NCS atlas. The whole color atlas was scanned with a spectroradiometer and sampled at a distance of 10.0nm ranging from wavelengths starting at 380nm an ending at 730nm. For this work, the data gathered from the spectroradiometer was slightly adjusted because some entries had to be removed due to errors in the scanning process. Finally, 1,927 entries remained in the data set.

The Gretag Macbeth [McCamy et al., 1976] color checker with its 24 entries is utilized for some comparison purposes where a high number of entries could not be displayed or handled

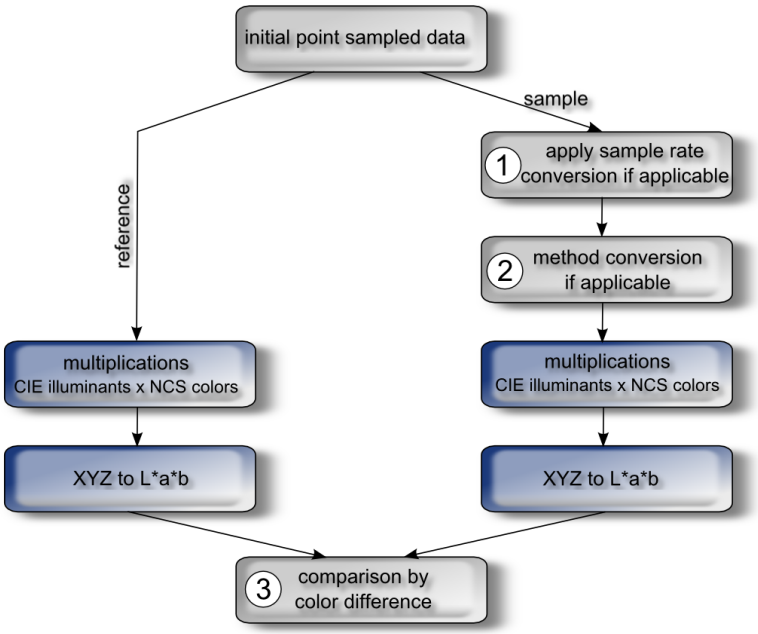

Figure 5.5: This image shows the approach used to measure the accuracy of a chosen method. Steps 1 and 2 denote the major difference in the computation between the reference data and the sample used for comparison. Step 1 covers sampling rate conversions, and step 2 includes the conversion to the respective representation method. The final step 3 represents the color comparison and is done by applying the CIEDE2000 color difference formula to the sample and the reference.

appropriately. This data was sampled at a distance of 5.0nm ranging from wavelengths of 380nm to 780nm.

For my work an application based approach was chosen to perform the necessary calculations for color comparison. The application created for this work operates with spectral data sets and calculates color differences for them for selected testing scenarios. It also computes statistical key figures and is able to visualize spectral data and calculation results. The used comparison mechanism is visualized in figure 5.5.

To compare several methods of data representation the program calculates multiplications of spectral data and light sources, which corresponds to multiplications that occur in a spectral rendering applications. The results of these calculations are compared with results of a reference set consisting of multiplications of densely sampled initial data. The program generates a report for any combination of colors and illuminants.

The color difference formula CIEDE2000 (see section 2.5 for details) is used to compare the results of the different representation methods and the results of the reference calculated from initial data.

The program calculates the following statistical characteristics and saves them to a file: mean average, standard deviation, median, quantiles of 2.5%, 25%, 75% and 97.5%. These character-

istics are used to represent the results of the calculations as box-plots which occur in the chapter "Results".

Besides doing calculations, the program also has the option to generate reports of the results. Furthermore, individual color differences are printed in these reports. Two reports of this kind where used to produce figure A.1 in the appendix.

5.5 Enhancements to Riemann summations

While working with the methods to represent spectral data I had the idea to apply the principle of Sun's composite method [Sun, 2000] in another way. Instead of using Fourier series, my variation uses Riemann summations to represent the smooth parts of a spectrum. Using Riemann summations has the great advantage of avoiding costly algorithms as the discrete Fourier transformation (compare: Methods based on a composite approach, 5.4). Both, Sun's and my method utilize separation of the initial data by relative gradients as shown by Deville and his colleagues [Deville et al., 1994]. After this step my method differs in concept because it stores the smooth parts of the spectra as Riemann summations. It utilizes the fact that if no peaks are present in both spectra involved in a multiplication a point wise multiplication as in standard Riemann summations implementations is used.

CHAPTER 6

Results

In this chapter I want to present the results of my work. This chapter is divided into two sections. The first section is on the comparison of different methods to represent spectral data. The second section is concerned with the impact of the application of low pass filters during the preparation of the data on the results.

6.1 Test scenarios

I divided the scenarios I used in my work into two groups. The first group was designed for the comparison of different methods for spectral data representation. The second group was designed for the comparison of different low pass filter setups. Within these two scenarios a used method is labeled and evaluated according to the following properties:

- Representation methods and their abbreviations: point sampling [PS], Riemann summations [RS], Sun's composite method [CS] and enhanced Riemann summations [RE]. The relative slope for peak detection is fixed at 30 percent for Sun's composite method and 20 percent for enhanced Riemann summations.

- Number of sample points or spectral coefficients used: This property corresponds to the compactness of a method and is given as a number after the abbreviation of the used method. This number should be interpreted as follows for point sampling and Riemann summations: any number of sample points includes start- and endpoints. This convention does not refers to the number of sample points without start- and endpoint which is used by [Peercy, 1993]. For any other method the number corresponds to the amount of used spectral coefficients and represents all a_k and b_k (real and imaginary part of the complex number) without b_0, because b_0 is always equal to zero. A number of spectral coefficients of 5 means, that $a_{0,1,2}$ and $b_{1,2}$ are used for calculation, or in other words the first three Fourier coefficients.

- Type of the used illuminants: Fluorescent illuminants [F] (CIE F2,F7 and F11) or non-fluorescent illuminants [NF] (CIE A,B and C).
- Used low pass filter during sampling rate conversion. This property describes the selected filter used during sampling rate changes. Possible values are: none, box, tent, Gauss, Lanczos, Mitchell, Hann window, Hamming window and Welch window.
- The extent of the filtering window: This property is only used if filters are applied and describes how many sample points usually influence the filter. Sample values which are not within 380nm and 780nm are counted as equal to 0.0. Possible values for this property are uneven integers ranging from number 3 to 21.
- Filter setup. This property is only used if filters are applied and contains the variables which control the appearance of some of the used filters, like α for the Gaussian filter, τ for the Lanczos filter and b, c for the Mitchell filter.

In the first group of scenarios no filters are applied. Sampling rate conversions, if applicable, are applied by linear interpolation and no ideal sampling rate conversion process is used. In this group the following properties vary: representation method, number of sample points and type of the used illumination.

6.2 Method comparison

This section is on finding an answer to the question of how accurate methods for the representation of spectral data work with a large data set? I distinguish between the type of the used light source, between non-fluorescent light sources (CIE A,B,C, labeled as NF) and fluorescent ones (CIE F1,F2,F7, labeled as F). To measure how accurate a method works I refer to the statistical key figures which were computed in the method comparison process.

Method comparison with non-fluorescent data

In this subsection I present the results of all test scenarios designed to answer how accurate a method works with non-fluorescent data. The statistical characteristics of the results are displayed in the table 6.1. The table was sorted according the mean color difference of multiplications of light sources with the NCS color data. The used color difference formula was CIEDE2000. Image 6.1 shows the results of all methods for non-fluorescent light sources as a box plot. The box plots depict the quantiles and the median of the results [Stahel, 2008].

The statistical figures, especially the mean average in table 6.1 indicates that all methods work well with non-fluorescent data. The majority of the test results is below an average mean color difference of 1.0.

The results indicate, that the NCS seems to incorporate data which triggers the peak detection algorithm used by the composite and enhanced Riemann method. This can be seen in table 6.1 when the results of the methods using Riemann summations [RS] are compared to the enhanced method [RE]. If there is no fluorescent data, both methods should yield the same results.

color difference CIEDE2000							
	mean average	**standard deviation**	**median**	**Q2.5**	**Q25**	**Q75**	**Q97.5**
PS32	0.149	0.100	0.123	0.023	0.067	0.218	0.379
RE32	0.159	0.090	0.143	0.038	0.083	0.222	0.355
RS32	0.194	0.114	0.172	0.045	0.094	0.276	0.438
CS15	0.395	0.478	0.315	0.118	0.241	0.372	1.258
PS16	0.605	0.396	0.510	0.103	0.286	0.875	1.499
RE16	0.631	0.362	0.584	0.078	0.340	0.882	1.373
RS16	0.756	0.463	0.663	0.137	0.380	1.084	1.759
CS11	0.866	0.918	0.625	0.138	0.390	1.012	3.216
CS7	1.262	1.151	0.836	0.252	0.574	1.544	4.658
PS8	2.295	1.465	1.994	0.403	1.047	3.347	5.565
RS8	2.622	1.435	2.425	0.540	1.487	3.600	5.677
RE8	2.799	1.679	2.527	0.335	1.480	3.925	6.359
CS5	5.605	3.336	5.122	0.972	3.089	7.416	12.643

Table 6.1: Comparison of all test scenarios, designed to answer how accurate a method works with non-fluorescent data. The table is sorted according the mean average of the color difference of each result. The used color difference formula is CIEDE2000. The abbreviations are read as follows: point sampling (PS), Riemann summations (RS), enhanced Riemann summations (RE), composite method(CS), quantiles with respective percent value (Q2.5, Q25, Q75 and Q97.5). All methods with a color difference below 1.0 are printed in bold font. A visualization as a box plot is given in figure 6.1. The number behind the abbreviation stands for the number of sample points (e.g. PS16: point sampling with 16 sample points) or spectral coefficients (e.g. CS5: composite methods with the first 3 spectral coefficients pairs a_k and b_k without b_0).

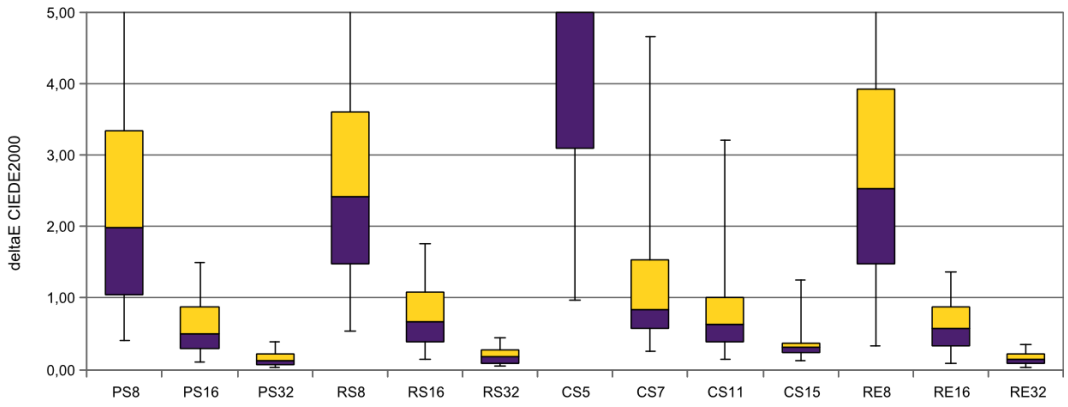

Figure 6.1: Comparison of methods tested with non-fluorescent data as a box-plot. The numbers which were used to create this box-plot are displayed in table 6.1.

	color difference CIEDE2000						
	mean average	**standard deviation**	**median**	**Q2.5**	**Q25**	**Q75**	**Q97.5**
CS15	0.316	0.222	0.293	0.099	0.199	0.358	0.816
RE32	0.474	0.379	0.341	0.064	0.191	0.667	1.484
CS11	0.719	0.491	0.637	0.134	0.394	0.915	1.997
RE16	1.160	0.612	1.056	0.249	0.709	1.512	2.579
CS7	1.205	1.060	0.789	0.209	0.433	1.629	3.778
RS32	1.642	0.885	1.536	0.281	0.918	2.289	3.529
RS16	2.026	1.033	1.880	0.436	1.249	2.682	4.350
PS32	2.649	1.311	2.501	0.723	1.546	3.642	5.341
CS5	4.608	3.296	3.745	0.629	2.002	6.538	12.697
RS8	4.976	2.497	4.679	1.337	2.998	6.483	10.596
RE8	5.179	3.095	4.506	1.013	2.921	6.836	12.858
PS8	10.669	4.677	9.973	3.877	7.113	13.556	22.262
PS16	11.198	5.281	10.451	3.792	6.776	14.843	22.683

Table 6.2: Comparison of all test scenarios designed to answer how accurate a method works with fluorescent data. The table is sorted according the mean average of the color difference of each result. The used color difference formula is CIEDE2000. The abbreviations are read as follows: point sampling (PS), Riemann summations (RS), enhanced Riemann summations (RE), composite method(CS), quantiles with respective percent value (Q2.5, Q25, Q75 and Q97.5). All methods with a color difference below 1.0 are printed in bold font. A visualization as a box plot is given in figure 6.2. The number behind the abbreviation stands for the number of sample points (e.g. PS16: point sampling with 16 sample points) or spectral coefficients (e.g. CS5: composite methods with the first 3 spectral coefficients pairs a_k and b_k without b_0).

However, this is not the case. Thus, it can be concluded that the NCS color atlas contains peaks with a relative slope greater than 20 percent.

The relation between median and mean average shows that all results are not evenly distributed, since the median lies below the mean average in every depicted case. This indicates a skewed distribution of the results.

As a whole, point sampling as the simplest of the introduced methods with the lowest growth of complexity of $O(n)$ works very well in this scenario. This proves the summary regarding point sampling from Sun [Sun, 2000]. This method should be the method of choice if any coverage of fluorescent data is not needed.

The next subsection contains the test scenarios including fluorescent data.

Method comparison concerning fluorescent data

The results incorporating the usage of fluorescent data are presented in table 6.2 and visualized as a box plot in figure 6.2.

The presence of peaks in the fluorescent data has a noticeable impact on the results of the test scenarios. Thus, the results from table 6.2 and 6.1 are quite different. Except for Sun's composite

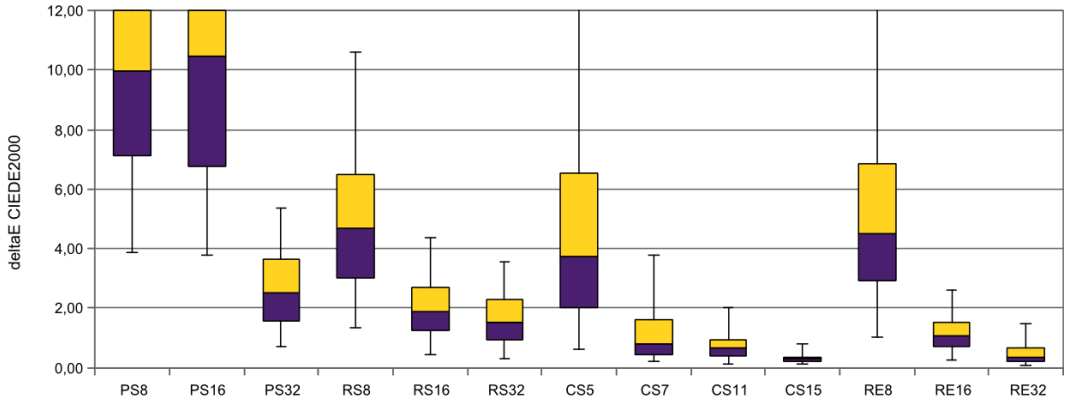

Figure 6.2: Comparison of methods tested with fluorescent data as a box-plot. The numbers which were used to create this box-plot are displayed in table 6.2.

method they are usually inferior to the results obtained by the tests using non-fluorescent data.

All results in table 6.2 and 6.1 show the same connection between median and mean average. The median is usually below the corresponding mean average, which indicates that the distribution of the results is shifted.

The results in table 6.2 indicate the following conclusions. Composite methods which incorporate the separate handling of low frequent and high frequent data work best with fluorescent data. Such methods are the composite method of Sun [Sun, 2000] and the enhanced Riemann summations. The results prove that this separation is a valid course of action. In cases that a separation is not applicable the results indicate that the methods using Riemann summations work best. The results also display the dramatically reduced accuracy of point sampling when working with fluorescent data.

The final choice of which method to use when working with fluorescent data is heavily dependent on calculation costs and storage requirements. Accurate point sampling needs a lot of sample values. In contrast, composite methods include a higher algorithmic effort but require fewer sample values to store a spectrum.

6.3 Impact of low pass filters on point sampling

This section is on the results with low pass filters used during the sampling rate conversion of the initial data. Only results from point sampling are analyzed because other methods are not affected by the usage of filters since they usually do not include any sampling rate conversion process.

Filter comparison with non-fluorescent light sources

The results presented in this section are divided into results incorporating the usage of non-fluorescent light sources and those incorporating fluorescent light sources. Results are sorted

according to mean average.

Table A.1 (located in the appendix) contains the results for scenarios using non-fluorescent light sources. The mean average indicates that the Mitchell filter and the Lanczos window are good choices. If only a small number of sample points is used, every filter used for sampling rate conversion improves the results. They tend do be very wide when a small number of sampling points is used. This filter behavior becomes visible when the filter extent is varied. Results of different filter extents are depicted in figures 6.3 and 6.4.

Results change when higher numbers of sample points are used. However, only the Mitchell and Lanczos filter yield improved results. Any other filter worsens the average mean and the standard deviation to a small degree in all test scenarios. In contrast, with scenarios which incorporate a small number of sample points, filters usually yield relatively the best results when their filter extent is small.

Filter comparison with fluorescent light sources

The scenarios concerning fluorescent light sources yield different results. In their cases, the application of a low pass filter during sample conversion usually improves the result of any test scenario. Results in table A.2 (located in the appendix) indicate that any low pass filter improves the result and that the Mitchell and Lanczos filter yield the best results. The filter behavior is similar to the results regarding non-fluorescent light sources. If a low number of sample points is used the filter extent should be very wide to achieve good results.

The behavior of the filter parameters shows major changes when compared to non-fluorescent scenarios. As depicted figures 6.3 and 6.4 almost any filter shows a optimal setup when used with 8 or 16 sample points, which may be obtained by choosing the setup at minimal average color difference.

Filter comparison in detail

When I review a specific filter I will refer to the numbers in tables A.1 and A.2 (located in the appendix), and the results depicted in figures 6.3 and 6.4 for the interpretation of the results.

Box filter

The box filter, despite its simplicity, yields adequate results when used with scenarios with fluorescent light sources. The results are good when a low number of sample points is given. This changes when a higher number of sample points is present. Although the results improve as well, they are inferior to the results of other filters used with the same data.

Tent filter

The tent filter yields similar results to the box filter. Its behavior for different extents when used with fluorescent light sources and a number of sample points greater than 8 follows a smoother curve than the box filter. The tent filter is the best choice when using a filter with only one parameter.

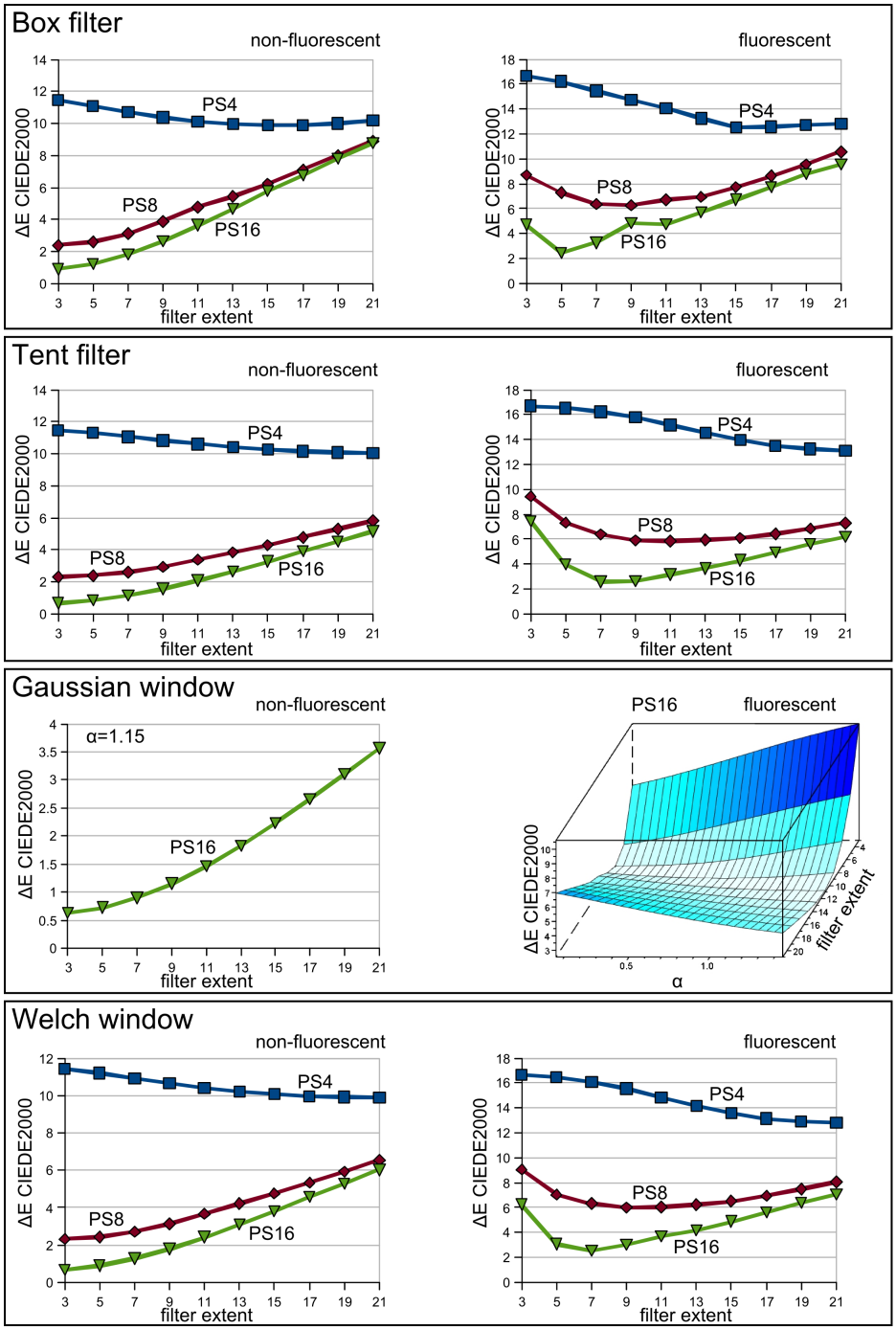

Figure 6.3: Impact of preprocessing on point sampling. The mean averages of point sampling involving fluorescent and non-fluorescent light sources were added together for visualization purposes. The application of a suitable filter improves the results of point sampling, involving fluorescent light sources, but reduces the accuracy of the results involving non-fluorescent light sources.

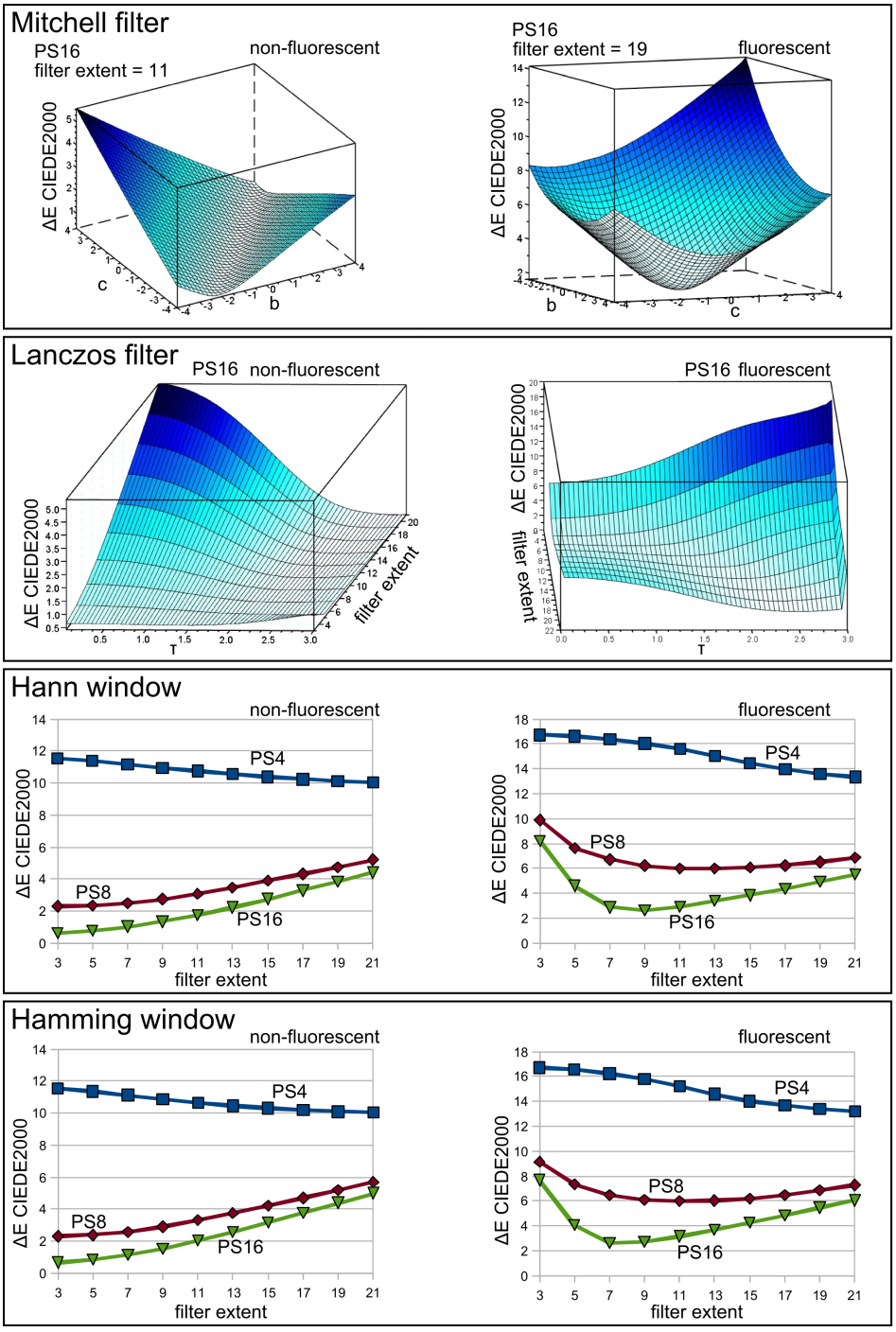

Figure 6.4: Impact of preprocessing on point sampling. The mean averages of point sampling involving fluorescent and non-fluorescent light sources were added together for visualization purposes. The application of a suitable filter improves the results of point sampling involving fluorescent light sources, but reduces the accuracy of the results involving non-fluorescent light sources.

Welch window

The Welch window yields similar results to the box filter and shows a smooth parameter behavior similar to the tent filter. In some cases, the results of the application of the Gaussian filter with a small parameter α tend to converge to the results of the Welch window.

Gaussian window

The Gaussian filter is different from the previous filters because it introduces an additional parameter α which controls the width of the Gaussian bell. The filter itself is a good choice when used with scenarios incorporating fluorescent light sources. It tends to settle at a very low α which results in a filter window similar to the Welch window. If this is the case it is inferior to the Welch window. However, there is an exception when 8 sample points are used for calculation. In this case, the parameter results in a typical Gaussian bell and the filter yields better results than the Welch window.

Mitchell filter

The Mitchell filter performs very well in any given scenario. Due to the existence of total 3 parameters (filter extent, b, c) there are a lot of possible parameter combinations. The average error in scenarios using fluorescent light sources indicate the existence of a minimum color difference(figure 6.4). The Mitchell filter is the best choice of all introduced filters and improves even results of scenarios without fluorescent light sources, although the selection of optimal parameters seems to be a complex task.

Lanczos filter

The Lanczos filter yields the second best results in any given scenario. Due to only two parameters, there are fewer possible parameter combinations than for the Mitchell filter. The two parameters tend to yield similar results in different combinations. It seems that the filter extent is connected to the parameter τ. A higher extent seems to be connected to a higher τ in order to yield better results.

Hann and Hamming window

The two windows behave similar to the Welch window and show a similar parameter behavior. The discontinuities of the Hamming window seem to negatively affect the results in comparison to the Hann window.

CHAPTER 7

Conclusion

The results of this work make it possible to obtain the following answers to the question, how well methods for representing spectral data work regarding their accuracy:

1. When working with a digital image synthesis application it is important to know if it will have to deal with fluorescent data or not. This knowledge has a major impact on the choice of the method which should be selected for the representation of spectral data. This point is emphasized when looking at the results of different methods in chapter 6. Almost any used representation method is sufficient for the representation of non-fluorescent data, but this does not hold true for fluorescent data.

2. If an application only works with non-fluorescent data, point sampling is sufficient to represent spectral data.

3. If an application works with both fluorescent and non-fluorescent data a composite approach to yield accurate results should be used.

4. The results prove that handling fluorescent data requires either complexer methods such as Sun's composite method [Sun et al., 1999] or a higher number of sample points to yield suitable results. Composite methods like [Sun, 2000, Sun et al., 1999] result in the highest accuracy.

A composite approach, as first introduced by Sun [Sun, 2000], includes the separation of smooth parts and peaks of a spectrum. Deville and his colleagues [Deville et al., 1994] describe an algorithm for this separation utilizing the relative gradient between sample points.

The second research question is on the impact of the usage of low pass filters during sampling rate conversion on the methods presented in this work. The results indicate:

1. That there is no improvement for methods based on Riemann summations or composite approaches. This relates to the absence of any sampling rate conversion.

color difference CIEDE2000				
	none	**Mitchell**	**Lanczos**	**Tent**
PS16	11.198	1.644	1.804	2.598
		-85.32%	-83.89%	-76.80%

Table 7.1: This table summarizes the impact of the ideal sampling rate conversion process on the mean average of point sampling with 16 sample points and test scenarios including fluorescent light sources.

2. That the ideal sampling rate conversion process including a low pass filter has a positive impact on the results. Table 7.1 depicts the improvements for point sampling with 16 sample points and scenarios concerning fluorescent light sources. The Mitchell filter yields the best result, but introduces 3 parameters (filter extent, b and c). The Lanczos filter yields the second best results and only introduces 2 parameters (filter extent and tau). Finally the Tent filter yields the best results for filters with only one parameter (filter extent).

3. That the choice of the ideal filter setup is not trivial, since the optimal filter parameters depend on the used spectral data.

7.1 Summary of contributions

Regarding the first question on the accuracy of existing methods my work contributes:

1. A ranking of existing methods based on a large spectral data set and the CIE illuminants.

2. An enhancement to Riemann summations derived from a composite approach which neglects the need for Fourier transformations and their complex calculations when applying a composite approach. The major advantages of my method are that it may revert to standard Riemann summations if no handling of high frequent data is necessary, and that it does not require complex algorithms.

With the results of the tests concerning sampling rate conversions my work contributes the following:

1. My results show that there is a positive influence on the ideal sampling rate conversion process on point sampling involving fluorescent data.

2. My work shows that the Mitchell, Lanczos and Tent filter are good choices for the preparation of spectral data.

3. Detailed results concerning the usage of a low-pass filter for different filters and filter extents for the use with point sampling.

7.2 Future work

After the numeric comparisons in my work, I think an implementation of a small rendering application could be a valuable asset. The results of my work could thus be verified by generating test scenes incorporating spectral effects. Furthermore, this application could be used to measure performance and perceived accuracy of spectral representation methods.

Future work could also focus on the exploitation of prior knowledge on the colors involved in a scene [Peercy, 1993]. This knowledge may influence the selection and configuration of spectral data representation methods.

Another approach for the usage of sampling rate conversions including low pass filter may include the question which filter setup would be suitable for the representation of one given spectrum. It should be possible to obtain properties to decide which filter parameters should be used for sampling rate conversion.

APPENDIX A

Appendix

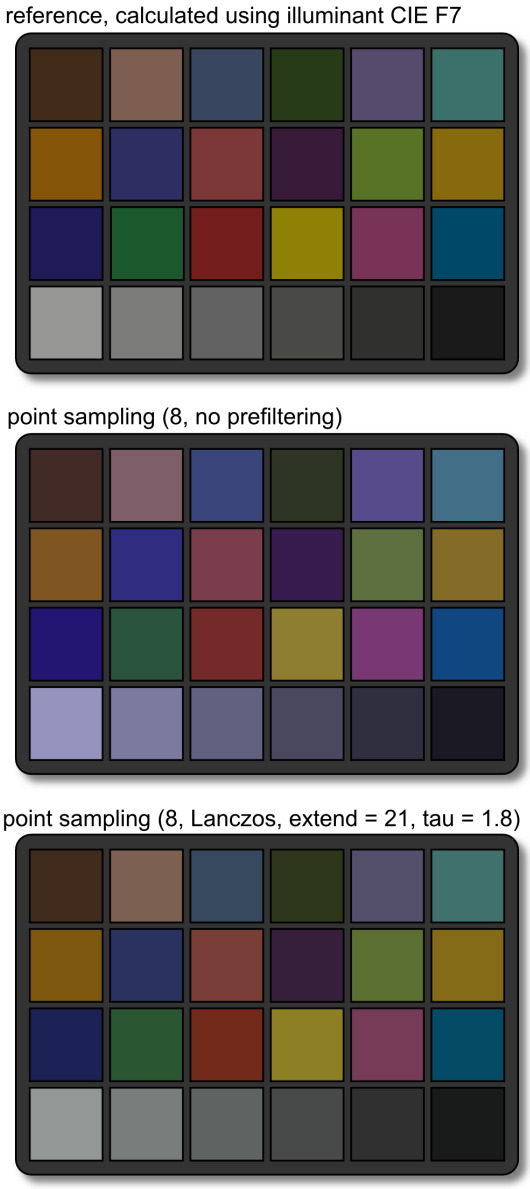

Figure A.1: Illustration using the Gretag Macbeth color checker chart [McCamy et al., 1976] to depict the impact of ideal sampling rate conversion. The first image from above shows the reference for comparison. The second image shows the same color checker filled with colors generated by utilizing point sampling and no ideal sampling rate conversion. The last image shows the results utilizing point sampling and ideal sampling rate conversion utilizing a low-pass filter. The mean average color difference calculated by the color difference formula CIEDE2000 improves from 8.189 to 2.251 (-74.18%).

color difference CIEDE2000

scenarios using non-fluorescent light sources				
		no filter	**box**	**tent**
point sampling	mean	11.656	9.891	10.032
4 sample values	σ	7.117	5.321	5.456
	parameters		extent = 15	extent = 21
point sampling	mean	2.295	2.403	2.320
8 sample values	σ	1.465	1.498	1.444
	parameters		extent = 3	extent = 3
point sampling	mean	0.605	0.903	0.657
16 sample values	σ	0.396	0.562	0.403
	parameters		extent = 3	extent = 3
		Welch	**Gaussian**	**Mitchell**
point sampling	mean	9.901	9.908	9.801
4 sample values	σ	5.285	5.299	5.298
	parameters	extent = 21	extent = 21	extent = 21
			$\alpha = 0.05$	$b = 2.0$, $c = -0.4$
point sampling	mean	2.327	2.303	1.535
8 sample values	σ	1.436	1.445	0.931
	parameters	extent = 3	extent = 3	extent = 13
			$\alpha = 0.90$	$b = -0.2$, $c = 2.0$
point sampling	mean	0.675	0.623	0.362
16 sample values	σ	0.406	0.399	0.198
	parameters	extent = 3	extent = 3	extent = 11
			$\alpha = 1.15$	$b = 1.0$, $c = 2.0$
		Lanczos	**Hann**	**Hamming**
point sampling	mean	9.949	10.047	10.038
4 sample values	σ	5.382	5.557	5.467
	parameters	extent = 21	extent = 21	extent = 21
		$\tau = 0.05$		
point sampling	mean	1.954	2.296	2.311
8 sample values	σ	1.207	1.432	1.442
	parameters	extent = 17	extent = 3	extent = 3
		$\tau = 2.70$		
point sampling	mean	0.415	0.620	0.659
16 sample values	σ	0.267	0.388	0.411
	parameters	extent = 17	extent = 3	extent = 3
		$\tau = 2.65$		

Table A.1: This table contains the best results for test scenarios concerning non-fluorescent light sources per low pass filter.

color difference CIEDE2000				
	scenarios using fluorescent light sources			
		no filter	**box**	**tent**
point sampling	mean	17.783	12.533	13.112
4 sample values	σ	6.335	5.378	5.182
	parameters		extent = 15	extent = 21
point sampling	mean	10.669	6.282	5.838
8 sample values	σ	4.677	3.174	3.188
	parameters		extent = 9	extent = 11
point sampling	mean	11.198	2.424	2.598
16 sample values	σ	5.281	1.207	2.160
	parameters		extent = 5	extent = 7
		Welch	**Gaussian**	**Mitchell**
point sampling	mean	12.839	12.864	12.681
4 sample values	σ	5.143	5.150	5.215
	parameters	extent = 21	extent = 21	extent = 21
			$\alpha = 0.05$	$b = 2.0, c = -1.0$
point sampling	mean	6.006	5.960	4.885
8 sample values	σ	3.448	3.326	2.286
	parameters	extent = 9	extent = 15	extent = 11
			$\alpha = 1.30$	$b = 1.0, c = -2.4$
point sampling	mean	2.542	2.543	1.644
16 sample values	σ	1.456	1.509	0.793
	parameters	extent = 7	extent = 7	extent = 19
			$\alpha = 0.05$	$b = -0.4, c = 1.0$
		Lanczos	**Hann**	**Hamming**
point sampling	mean	13.010	13.327	13.206
4 sample values	σ	5.189	5.269	5.186
	parameters	extent = 21	extent = 21	extent = 21
		$\tau = 0.05$		
point sampling	mean	5.837	5.992	6.003
8 sample values	σ	3.338	3.259	3.346
	parameters	extent = 21	extent = 13	extent = 11
		$\tau = 1.80$		
point sampling	mean	1.804	2.599	2.636
16 sample values	σ	1.113	1.575	2.117
	parameters	extent = 21	extent = 9	extent = 7
		$\tau = 2.50$		

Table A.2: This table contains the best results for test scenarios concerning fluorescent light sources per low pass filter.

Bibliography

[Blinn, 1989] Blinn, J. F. (1989). Jim Blinn's corner- What We Need Around Here Is More Aliasing. *IEEE Computer Graphics and Applications*, 9:75–79.

[Bronstein et al., 2001] Bronstein, I. N., Semendjajew, K. A., Musiol, G., and Mühlig, H. (2001). *Taschenbuch der Mathematik*. Verlag Harri Deutsch, 5th edition.

[Bryant et al., 2009] Bryant, J., Kester, W., and Sheingold, D. (2009). *Digital signal processing*, chapter ADCs, DACs, and Sampling Theory, pages 1–122. Newnes.

[Deville et al., 1994] Deville, P. M., Merzouk, S., Cazier, D., and Paul, J.-C. (1994). Spectral data modeling for a lighting application. *Comput. Graph. Forum*, 13(3):97–106.

[Devlin et al., 2002] Devlin, K., Chalmers, A., Wilkie, A., and Purgathofer, W. (2002). Star report on tone reproduction and physically based spectral rendering. In *Eurographics 2002*. Eurographics Association.

[Doblinger, 2007] Doblinger, G. (2007). *Zeitdiskrete Signale und Systeme*. Schlembach.

[Frigo and Johnson, 2009] Frigo, M. and Johnson, S. G. (2009). FFTW: Fastest Fourier Transform in the West. avaiable online:http://www.fftw.org/ (last access: 28.02.2010).

[Glassner, 1995a] Glassner, A. S. (1995a). *Principles of Digital Image Synthesis - Volume One*. Morgan Kaufmann, 1st edition.

[Glassner, 1995b] Glassner, A. S. (1995b). *Principles of Digital Image Synthesis - Volume Two*. Morgan Kaufmann, 1st edition.

[Haferkorn, 2002] Haferkorn, H. (2002). *Optik*. Wiley.

[Hall, 1999] Hall, R. (1999). Comparing spectral color computation methods. *IEEE Comput. Graph. Appl.*, 19(4):36–45.

[Hayes, 1999] Hayes, M. H. (1999). *Schaum's outline of theory and problems of digital signal processing*. McGraw-Hill.

[Hearn and Baker, 2004] Hearn, D. and Baker, M. P. (2004). *Computer Graphics with OpenGL*. Pearson Prentice Hall.

[Hering, 2007] Hering, E. (2007). *Zur Lehre vom Lichtsinne : sechs Mittheilungen an die Kaiserl. Akademie der Wissenschaften in Wien*. VDM, Müller.

[Hunt, 1991] Hunt, R. W. G. (1991). *Measuring Colour*. Ellis Horwood, West Sussex, 2nd edition.

[HunterLAB, 2005] HunterLAB (2005). Equivalent white light sources and cie illuminants. *Insight on Color*, 17(5).

[Jähne, 2005] Jähne, B. (2005). *Digitale Bildverarbeitung*. Springer, Berlin.

[Johnson and Fairchild, 1999] Johnson, G. M. and Fairchild, M. D. (1999). Full-spectral color calculations in realistic image synthesis. *IEEE Comput. Graph. Appl.*, 19(4):47–53.

[Johnson and Fairchild, 2003] Johnson, G. M. and Fairchild, M. D. (2003). A top down description of S-CIELAB and CIEDE2000. *Color Research & Application*, 28(6):425–435.

[Kajiya, 1986] Kajiya, J. T. (1986). The rendering equation. In Evans, D. C. and Athay, R. J., editors, *SIGGRAPH '86: Proceedings of the 13st annual conference on Computer graphics and interactive techniques*, pages 143–150. ACM.

[Lindbloom, 2007] Lindbloom, B. J. (2007). Useful color equations. avaiable online: http://www.brucelindbloom.com/ (last access: 04.12.2009).

[Lynch and Livingston, 1995] Lynch, D. K. and Livingston, W. (1995). *Color and Light in Nature*. Cambridge University Press.

[Marschner and Lobb, 1994] Marschner, S. R. and Lobb, R. J. (1994). An Evaluation of Reconstruction Filters for Volume Rendering. In *Proceedings of the IEEE Visualization '94 conference*.

[McCamy et al., 1976] McCamy, C. S., Marcus, H., and Davidson, J. G. (1976). A color-rendition chart. *Journal of Applied Photographic Engineering*, 2(3):95–99.

[Mitchell and Netravali, 1988] Mitchell, D. P. and Netravali, A. N. (1988). Reconstruction Filters in Computer Graphics. *Computer Graphics*, 22(22):221–228.

[Nassau, 2001] Nassau, K. (2001). *The Physics and Chemistry of Color: The Fifteen Causes of Color*. Wiley Series in Pure and Applied Optics.

[Oppenheim et al., 1999] Oppenheim, A. V., Schafer, R. W., and Buck, J. R. (1999). *Discrete-time signal processing*. Prentice Hall.

[Peercy, 1993] Peercy, M. S. (1993). Linear color representations for full speed spectral rendering. In *SIGGRAPH '93: Proceedings of the 20th annual conference on Computer graphics and interactive techniques*, pages 191–198, New York, NY, USA. ACM.

[Pharr and Humphreys, 2004] Pharr, M. and Humphreys, G. (2004). *Physically Based Rendering: From Theory to Implementation*. Morgan Kaufmann.

[Rougeron and Péroche, 1997] Rougeron, G. and Péroche, B. (1997). An adaptive representation of spectral data for reflectance computations. In Dorsey, J. and Slusallek, P., editors, *Rendering Techniques*, pages 127–138. Springer.

[Sharma et al., 2004] Sharma, G., Wu, W., Dalal, E. N., and Celik, M. U. (2004). Mathematical discontinuities in CIEDE2000 color difference computations. In *Color Imaging Conference*, pages 334–339. IS&T - The Society for Imaging Science and Technology.

[Shirley and Morley, 2003] Shirley, P. and Morley, R. K. (2003). *Realistic Ray Tracing*. A K Peters.

[Smith, 2001] Smith, A. R. (2001). Digital paint systems: An anecdotal and historical overview. *IEEE Ann. Hist. Comput.*, 23(2):4–30.

[Stahel, 2008] Stahel, W. A. (2008). *Statistische Datenanalyse*. Vieweg, 5th edition.

[Sun, 2000] Sun, Y. (2000). *A spectrum-based framework for realisitic image synthesis*. PhD thesis, Simon Fraser University.

[Sun et al., 1999] Sun, Y., Drew, M. S., and Fracchia, F. D. (1999). Representing spectral functions by a composite model of smooth and spiky components for efficient full-spectrum photorealism. In *PMCVG '99: Proceedings of the 1999 IEEE Workshop on Photometric Modeling for Computer Vision and Graphics*, page 4, Washington, DC, USA. IEEE Computer Society.

[Theußl, 1999] Theußl, T. (1999). Sampling and reconstruction in volume visualization. Master's thesis, Technical University of Vienna.

[Theußl et al., 2000] Theußl, T., Hauser, H., and Gröller, E. (2000). Mastering windows: improving reconstruction. In *Volviz*, pages 101–108.

[Turkowski and Gabriel, 1990] Turkowski, K. and Gabriel, S. (1990). *Graphics Gems I*, chapter Filters for Common Resampling Tasks, pages 147–165. Academic Press.

[Whitted, 1980] Whitted, T. (1980). An improved illumination model for shaded display. *Communications ACM*, 23(6):343–349.

[Wyszecki and Stiles, 2000] Wyszecki, G. and Stiles, W. S. (2000). *Color Science*. Wiley Series in Pure and Applied Optics, 2nd edition.