

Fachhochschule Köln
Fakultät für Informations-, Medien- und Elektrotechnik
Institut für Medien- und Phototechnik
Masterstudiengang Medien- und Bildtechnologie

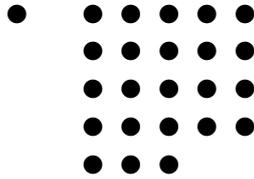
Masterarbeit

Entwicklung einer Softwareschnittstelle für objektive Analyse der Videobildqualität

Borys Golik
Matr.-Nr. 11038740

Erstgutachter: Prof. Dr.-Ing. Klaus Ruelberg,
Fachhochschule Köln
Zweitgutachter: Dipl.-Ing. Dietmar Wueller,
Image Engineering

19. Februar 2010



Cologne University of Applied Sciences
Faculty of Information, Media and Electrical Engineering
Institute of Media and Imaging Technology
Degree program of Master of Engineering in Media and Imaging Technology

Master Thesis

Software Interface for Video Image Quality Analysis

Borys Golik
Matriculation No. 11038740

First Reviewer: Prof. Dr.-Ing. Klaus Ruelberg,
Cologne University of Applied Sciences
Second Reviewer: Dipl.-Ing. Dietmar Wueller,
Image Engineering

February 19, 2010

Zusammenfassung

Titel: Entwicklung einer Softwareschnittstelle für objektive Analyse der Videobildqualität

Autor: Dipl.-Ing. Borys Golik

Gutachter: Prof. Dr. Klaus Ruelberg, Dipl.-Ing. Dietmar Wueller

Zusammenfassung: Einer der zentralen Aspekte in der Videoproduktion ist die Frage nach der Bildqualität. Im Bezug auf Messtechnik birgt die Umstellung auf digitale Systeme viel Potenzial, aber auch neue Fragestellungen. Die Entwicklung und die Umsetzung einer Softwareschnittstelle, die eine objektive softwarebasierte Auswertung der Bildqualität ermöglicht, sind ein wesentlicher Bestandteil dieser Arbeit. Der schriftliche Teil beschreibt ausgewählte Gebiete der Messtechnik im Hinblick auf die Farbmessung und erläutert die Implementierung der Software.

Stichwörter: Video, Bildqualität, Messtechnik

Sperrvermerk: Die vorgelegte Arbeit unterliegt keinem Sperrvermerk.

Datum: 19. Februar 2010

Abstract

Title: Software Interface for Video Image Quality Analysis

Author: Dipl.-Ing. Borys Golik

Reviewer: Prof. Dr. Klaus Ruelberg, Dipl.-Ing. Dietmar Wueller

Abstract: Image quality has always been one of the most important aspects in video engineering. With regard to quality assessment, the advent of digital technologies entails a great potential on the one hand, on the other hand it reveals previously unknown issues. The major challenge of this master thesis was the development and implementation of a software interface for the objective analysis of the video image quality. This manuscript gives an overview of selected topics of the measurement technology in matters of colorimetry. Furthermore, the development and the implementation of the software is discussed and documented.

Keywords: Video, Image Quality, Measurement Technology

Remark of closure: The thesis is not closed.

Date: February 19, 2010

Contents

1. Introduction	7
2. Measurement Basics	9
2.1. Image Quality in Still Imaging and Video	9
2.2. Units and Levels	11
2.3. Measurement and Monitoring	14
2.3.1. Waveform Monitor	14
2.3.2. Vectorscope	15
2.3.3. Histogram	18
2.3.4. Gamut	19
3. Colorimetry in Video Applications	25
3.1. The CIE Colorimetric System	25
3.1.1. Spectrum and Tristimulus Values	25
3.1.2. Color Matching Functions and Standard Observers	27
3.1.3. Color Temperature	32
3.1.4. Standard Illuminants	34
3.2. Color Spaces	38
3.2.1. Definition	38
3.2.2. Digital Image Color Workflow	38
3.3. RGB Encoding	40
3.3.1. Computer Graphics	40
3.3.2. Video Applications	43
3.4. Luma and Chroma Encoding	44
3.4.1. Luma and Chroma	44
3.4.2. Standard Definition TV	45

3.4.3. High Definition TV	48
3.4.4. Conversion between SD- and HDTV	49
3.5. Legal and Valid Signals	50
4. Software Design	53
4.1. Programming Language and Tools	53
4.2. IE-Analyzer	55
4.3. DeckLink Adaptor	57
4.3.1. C++ Framework	57
4.3.2. Image Acquisition	61
4.4. Video Module	63
4.4.1. Principle	63
4.4.2. Graphical User Interface (GUI)	65
4.4.3. Live Color Comparison	67
4.4.4. Integration into IE-Analyzer	74
5. Results	77
5.1. Test conditions	77
5.2. OECF and Noise	78
5.3. Resolution	80
5.4. Color	82
6. Conclusion and Outlook	85
Bibliography	86
Appendices	
A. Tables	93
B. Screenshots	95
C. Classes (UML)	102
Remarks	109

1. Introduction

Human vision is a highly complex process. Though it has its physical and perceptual limitations it is one of the most important sources of information about our environment. We are surrounded by images in various media that continuously strive for more realistic pictures. Thus, image quality assessment has become a very important discipline in the recent past. It covers the entire processing chain from capture, storage and transmission to reproduction of image data.

The majority of the broadcasters all over the world are in the process of adopting digital technologies, many have already changed over. Digital transmission via satellite and cable were introduced a fairly long time ago. Terrestrial digital broadcasting proved its ability in a number of locations around the world. The same tendency can be observed in photography, where digital still cameras have been replacing their analog ancestors over the past two decades. For the traditional analog video imaging there are a lot of well-established performance standards. The results of the measurement of signal parameters according to these standards highly agree with the perception of image quality. While these standards are still important today, this relationship diminishes due to several effects having their seeds in digital signal processing, such as compression artifacts. The amount and visibility of these artifacts is highly content-dependent which makes image quality measurement become a challenge.

Subjective viewing experiments provide reliable scoring of perceived quality. However, these measurements are complex and time-consuming which makes them very expensive. Furthermore, they can not be automated or used for

on-line quality monitoring. Objective approaches based on pixel-wise error measurements, such as mean squared error (MSE) or peak signal-to-noise ratio (PSNR), can be integrated into hardware devices. This equipment often includes traditional measurement instruments, such as a waveform monitor or a vectorscope. Operating only on pixel-by-pixel basis, the error-related measurement neglects the impact of the image content on the perception of artifacts. Assessment methods modeling the characteristics of the human visual system (HVS) could provide higher correlation between objectively measured parameters and subjective sensation.

Due to high performance, recent computer systems provide an opportunity of video signal processing at relative low costs. Analyzing digital image frames as a whole offers a technical benefit over the evaluation of single video lines. The major task of this work is to develop and implement a software interface between video input hardware and the IE-Analyzer, which is a software product made by Image Engineering for the purpose of standard-compliant image quality analysis of digital imaging devices. Furthermore, several measurement and monitoring utilities, such as waveform monitor and live color comparison, should provide a foundation for an objective software-based quality assessment method. The development and realization of these tools is also a part of this work. Furthermore, this manuscript gives a brief overview of selected aspects of the video measurement technology and colorimetry.

2. Measurement Basics

2.1. Image Quality in Still Imaging and Video

Imaging systems in video and photographic applications are generally concerned with pictures meant for humans. From the moment the picture is captured to the moment it is presented to an observer different kinds of image degradation can emerge. Capturing device, signal processing, transmission channel and reproduction equipment contribute to the distortion of the original scene. Admittedly, not every corruption affects perceived image quality to the same extent. Due to the properties of the human visual system the sense of picture quality is related to spatial and temporal image information, lighting and ambient conditions. Taking this into consideration, image quality metrics are supposed to estimate the degree of image degradation at different stages of processing with relation to the HVS.

The subjective approach is based on averaging opinions of many test persons. ITU recommendation ITU-R BT.500 (2002) describes a methodology for the subjective assessment of the quality of television pictures. Even though these methods suit the requirements of scientific researches and experiments, subjective quality assessment is not feasible in engineering practice causing great cost and being time consuming. Batch processing and signal monitoring are neither possible.

The purpose of the methods for objective image quality assessment is to provide metrics for a reliable and repeatable estimation of image quality as perceived by humans. Theoretically, this approach offers a way to monitor

and adjust image quality, to assist with improving algorithms and to benchmark different systems (Wang and Bovik, 2006). The objective approaches can coarsely be divided into *full reference* (FR), *reduced reference* (RR) and *no-reference* (NR) methodologies. Full reference methods depend on the availability of a reference signal or image that is assumed to be free of degradation. To be more precise, image similarity or fidelity rather than image quality is measured by this means. Reduced reference approach extracts certain features from the reference image, that are then compared to the corresponding features of the image to be analyzed. Intended for data transmission, this method saves bandwidth reducing the amount of auxiliary image data to comparatively few features. If a reference image is not available, no-reference assessment approach is supposed to figure out and quantify the extent of degradation. Though a human observer can easily accomplish this task, having a sense of how the image should look like based on his experience or prior knowledge, it becomes a great challenge for a software algorithm.

Different measurement procedures have been described in international standards for specific quality assessment in photographic applications, as well as in European norms (EN) for measurements of video equipment. They are generally based upon the usage of reference signals or images in the form of reflective or transparent test charts. On this account, only full reference image quality assessment methods are discussed hereafter. Moreover, only measurements of source signals are discussed, disregarding transmission channel and signal layer conformity of image data.

International standards define camera settings and test conditions for the measurements. Though standard conform image quality assessment provides a number of characteristics of capturing systems, it still might be important to relate objective measurement results to the real world scenes outside the lab. The specifications to be analyzed can be divided into obligatory, recommended and optional categories (Wueller, 2006). Mandatory aspects include opto-electronic conversion function OECF and modulation transfer function MTF (or spatial frequency response SFR). Other important properties of a system can be derived from OECF (white balancing, dynamic range, utiliza-

tion of available digital values and noise characteristics) and from MTF/SFR (limiting resolution, sharpness, quality of image stabilization and auto-focus accuracy). Recommended and optional measurements refer to such phenomena as distortion, shading (vignetting), chromatic aberration, color rendition and defective pixels.

In order to provide a reference, some selected standards have to be mentioned. For digital still-imaging applications, the International Organization for Standardization (ISO) has specified norms for OECF measurement (ISO 14524, 2009), resolution (ISO 12233, 2009) and noise (ISO 15739, 2003). Previously mentioned methods for the subjective assessment of the quality of television pictures are described in ITU-R BT.500 (2002). The European Committee for Electrotechnical Standardization (CENELEC) has published four parts of a standard concerning measurements for video cameras (PAL/SECAM/NTSC). EN 61146-1 (1997) deals with non-broadcast single sensor devices, EN 61146-2 (1998) with professional two- or three-sensor cameras, EN 61146-3 (1998) with non-broadcast camera recorders and EN 61146-4 (1999) is concerned with the automatic functions of video cameras and camera recorders.

2.2. Units and Levels

Every video line is a time-multiplex of active video, containing image information, and blanking interval. Together they yield a line period. Active image is defined as the set of all active lines. Signal level during the blanking interval is used as zero level reference for other video levels.

According to standards, active parts of a video line are constrained to certain specified voltage or digital value limits. *Reference black level* is referred to the lowest allowed peak black level in a video signal, while *reference white level* corresponds to a specified maximum limit for white peaks. Reference white level, sometimes referred to as *nominal white level* or simply *white level*, is

used as a 100 % reference to calibrate gains and other settings of measurement devices (Jack and Tsatsulin, 2002).

Analog video standards specify the peak-to-peak range of a video signal to be 1 V. However, there are currently two different picture-to-sync ratios. NTSC systems operate with the ratio of 10:4, whereas PAL systems, as well as HDTV, employ the 7:3 ratio. Referring to the $1 V_{PP}$ signal excursion, nominal white level in 10:4 ratio systems is approximately 714 mV, while in systems utilizing 7:3 picture-to-sync ratio it is set to 700 mV (Steinberg, 1997). To make the handling more convenient, the Institute of Radio Engineers (IRE) — the predecessor of the Institute of Electrical and Electronics Engineers (IEEE) — has introduced *IRE units* for the NTSC standard. The level difference between blanking interval and reference white — the picture excursion — was specified to be 100 IRE, thus 1 IRE unit corresponds to 7.14 mV.

According to the NTSC-M standard specification and the USA national standard, reference black level in North America has an offset to blanking interval of 7.5 percent of the picture excursion, which corresponds to 7.5 IRE or 53.6 mV (Steinberg, 1997). This offset is referred to as *setup* or *pedestal*. As a result, an NTSC system with setup has nominally 92.5 IRE units from reference black to reference white. Due to problems in maintaining precise black level reproduction, 7.5 % setup has been abolished from modern video systems. All variants of 576i and HDTV systems feature *zero setup* instead (Poynton, 2007). Concerning the setup issue, Japan employs a hybrid NTSC form — NTSC-J. As the picture-to-sync ratio is still 10:4, zero setup was adopted in 1985. As a result, there are now three different level standards for analog video interfaces (see Figure 2.1).

In digital video applications it is necessary to retain signals occasionally exceeding reference signal values. This may occur as a consequence of filter processing. For that reason *headroom* and *footroom* have been introduced. The Radiocommunication Sector of the International Telecommunication Union (ITU-R) has released a recommendation for studio encoding parameters of digital television (ITU-R BT.601, 2007). It defines luma channel (and R' , G'

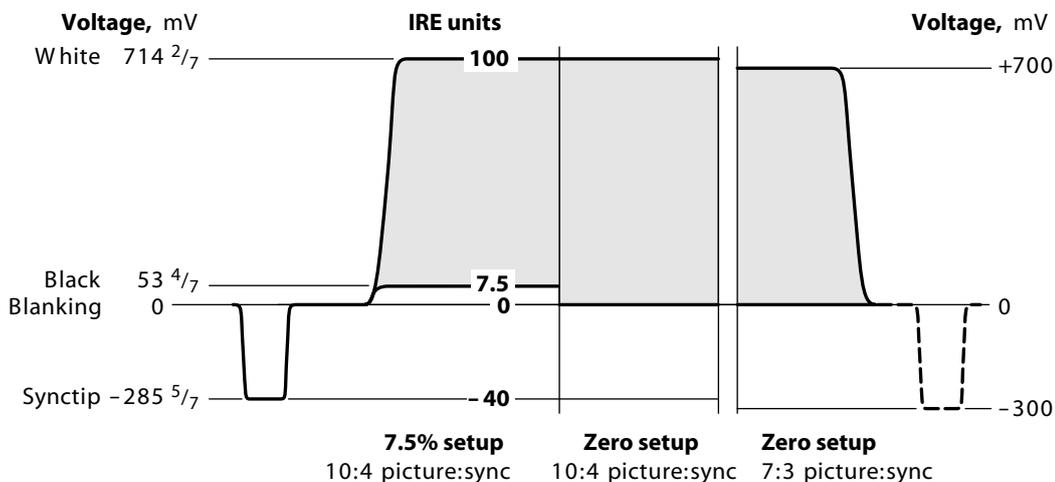


Figure 2.1.: Comparison of 7.5% and zero setup (Poynton, 2007).

or B' channels, when no chroma subsampling was employed) to have a value range of 220 quantization levels with the reference black level corresponding to quantization level 16 and the reference white corresponding to level 235. Color difference signals have to be encoded with an offset, so that zero signal level corresponds to digital level 128, and to range between levels 16 and 240 (see Figure 2.2).

In a 10-bit environment, the reference values are multiplied by a factor of 4. Consequently, the luma and RGB signals have a value range between quantization levels 64 and 940, while color difference signals can take values between levels 64 and 960. All signals may occasionally exceed their limits. The same quantization levels are also used in the Recommendation ITU-R BT.709-5, which describes parameter values for the HDTV standards for production and international program exchange (ITU-R BT.709, 2002).

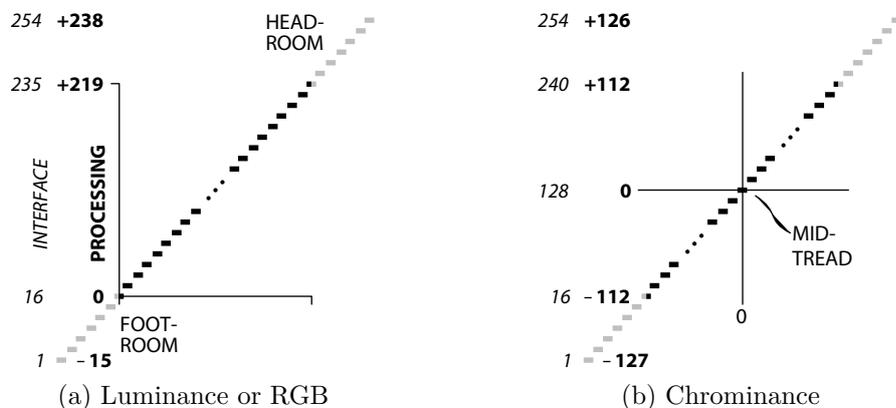


Figure 2.2.: Footroom and headroom in an 8-bit domain. Codes 0 and 255 are preserved as a timing reference (Poynton, 2007).

2.3. Measurement and Monitoring

2.3.1. Waveform Monitor

Oscilloscopes specially designed for video signals are referred to as waveform monitors (WFM). They display signal levels with respect to time and allow, in contrast to studio monitors or television sets, for objective measurement of video signals.

Originally, WFMs were analog measurement instruments utilizing cathode ray tubes (CRT). The vertical voltage is driven by the amplified video signal, whereas the horizontal time axis is driven by a sawtooth signal adjustable by different presets. They define whether to display a single video line (H), two successive lines (2H), complete video frame (V) or two video frames (2V). When a composite video signal is used, built-in low-pass filter allows to change the frequency response characteristics of the amplifier to toggle the display between *luma* (below approx. 2 MHz) and *chroma* (over approx. 2 MHz). Without the low-pass filtering, a *flat* representation is possible. In case of a component or $R'G'B'$ signal, the *parade* and *overlay* modes are available. Luma and two chroma signals are displayed side by side in parade mode and accordingly superimposed in the overlay mode (Schmidt, 2005).

The NTSC graticule is generally scaled in IRE units and extends from approximately -40 to 120 IRE in 10 IRE steps. Black level setup is labeled at the 7.5 IRE voltage. The PAL graticule is scaled in volts and can extend from 0 to approx. 1.2 V or from -0.3 to 1 V in 0.1 V increments.

Except for video trigger circuits and the ability to use external reference for sync signal, classic waveform monitors closely resemble oscilloscopes. With the advent of digital video, WFMs gained numerous additional features and capabilities. Many of them can display video pictures, check color gamut, measure physical properties of the bit stream using jitter and eye pattern display or check for the auxiliary data, such as embedded audio or meta data. Furthermore, modern waveform monitors have abandoned CRTs. As a convenience to the operator, colored liquid crystal displays (LCDs) are mostly used instead. A hardware module called *rasterizer* reproduces the behavior of a CRT display creating a raster signal for this purpose. Some waveform monitors do not feature any display at all, they can be connected to a VGA display instead.

2.3.2. Vectorscope

Since the phase angle, responsible for the hue, can not be easily measured in a high-frequency subcarrier signal by means of a WFM, vectorscopes are used for chrominance measurements. A Vectorscope is basically an oscilloscope working in X-Y mode, i.e. the horizontal deviation is driven by a second test signal instead of a sawtooth wave. It accepts standard television or video signals as input and has a specialized graticule for dedicated measurements though. When a composite analog signal is applied, internal filters and demodulators are used to obtain the color difference components P_B and P_R . The horizontal deviation is then driven by P_B , while P_R drives the vertical axis. In digital applications, C_B and C_R signals are plotted against each other instead.

The graticule represents all the colors inside a circle, neutral gray is placed in the center (see Figure 2.3). For a given color point, the saturation is represented by its distance from the center of the circle, whereas the angle around the center

2. Measurement Basics

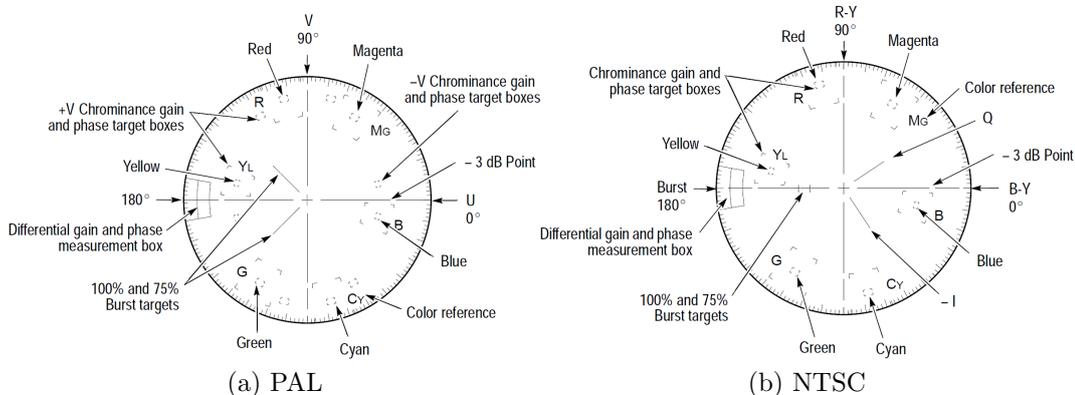


Figure 2.3.: Vectorscope graticule (Tektronix Inc., 1998).

stands for the color hue. For this reason, gain difference between a color and its target box indicates a saturation fault and a different phase angle hints at a wrong hue. Six target boxes related to the color bars test pattern define the reference colors. Their tolerance limits amount to 5% of signal gain and 3% of phase angle (Schmidt, 2005).

In video and television applications two different types of color bar test patterns are used. According to their amplitudes, they either refer to 75% or to 100% color bars. In order to provide proper margins for each of them, two sets of target color boxes would be necessary. Since older vectorscopes had a fixed silkscreen graticule, a device calibration was employed to ensure correct size of the tolerance boxes. The calibration is carried out using a short color burst line, which is first moved to the middle of the screen to match the neutral gray and then rotated to be aligned with the burst target line on the graticule. NTSC color burst has a phase of 180° , while the alternating phase in PAL causes burst signal to appear at 135° and 225° at even and odd lines respectively (see Figure 2.3). Finally the vectorscope's gain is adjusted, so that the burst signal either matches the 75% or the 100% mark.

Waveform monitor and vectorscope can both be used to visualize *differential gain* (dG). It shows the change of color saturation at different luma levels. Normally, luma does not affect the amplitude of chroma signal. Therefore,

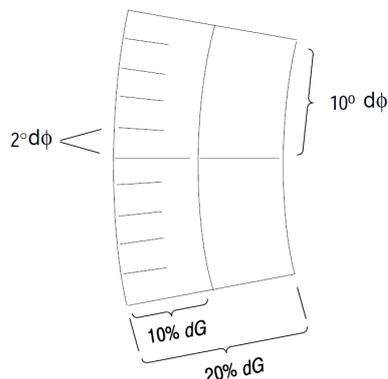


Figure 2.4.: Differential gain and phase measurement box on a vectorscope graticule (Tektronix Inc., 1998).

smaller dG values indicate better color rendition of a system. For this purpose many devices offer a target box on a graticule. It is located at the 100 IRE line on WFMs and at the outer edge of the P_B axis on the left on vectorscopes (see Figure 2.3). Vectorscopes also provide a facility to measure *differential phase* (dP or $d\phi$). While differential gain denotes changes in the chroma amplitude depending on the luma level, differential phase is related to variations in the chroma phase or color hue (Jack, 2005). Either a five- or ten-step modulated staircase signal or a modulated ramp can be used as input to measure differential gain and phase. Further information on differential gain and phase can be found in Poynton (2007).

Just to touch on the functional range of vectorscopes, audio applications should be mentioned. Vectorscopes can be used to display the difference between two audio channels in stereo applications, provided that one channel drives the horizontal and the other one the vertical deflection. Any stereo separation will produce a Lissajous curve in contrast to mono signals, resulting in a straight line with the slope of one.

Many waveform monitors introduce vectorscope features nowadays, hence stand-alone devices have become rare. Multifunctional devices, utilizing computer graphics to display signals, are still referred to as waveform monitors.

2.3.3. Histogram

In digital image processing, a histogram is a statistical representation of image data. In contrast to subjective visual evaluation complicated by the adaptation ability of the HVS, histogram offers an objective method to judge the distribution of pixel values in the image. Usually plotted as a bar chart it illustrates the frequency of occurrence of each intensity level (Bovik, 2000). In order to create a histogram the overall amount of pixels for each intensity level in the image is calculated. Multichannel images are typically split into single channels and their histograms are presented separately. In addition to the filter design applications in digital image processing, histograms are often used for the purpose of visualization.

Histograms do not contain any spatial information about the image, thus no image reconstruction from a given histogram is possible. In general, a vast amount of images with exactly the same histogram can be synthesized by spatially shuffling pixels of the initial picture.

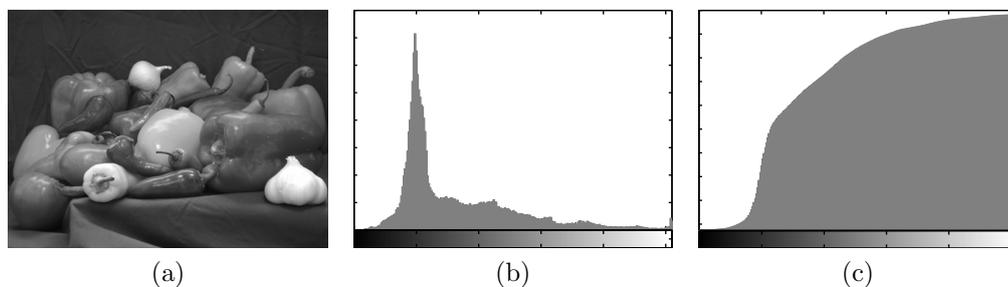


Figure 2.5.: Histogram representation: (a) original image, (b) histogram, (c) cumulative histogram.

There is no perfect histogram. Histogram's shape can vary in many different ways depending on the image content. However, some conclusions concerning image properties can be drawn by its means. Histograms shifted to one side indicate either exposure issues or a low-/high-key scene. The more intensity values are used, the higher is the image contrast, thus narrow distributions stand for low contrast pictures. If a histogram exhibits "holes", the amount

of actually used intensity values is reduced. This low dynamic may be caused e.g. by contrast enhancement, when single histogram bins are pulled apart. Significant peaks on the distribution's boundaries may expose clipping effects. Compression artifacts in GIF files (holes) and JPEG files (spread out of single histogram lines) can also be revealed by means of histograms. Since normalized histograms are empirical estimates of pixel value distributions, every statistical property, such as mean value, variance or standard deviation, can be investigated on them. More information on interpreting histograms can be found in Burger and Burge (2006).

The *cumulative histogram* (Figure 2.5c) should be briefly mentioned to complete the overview. This type can be used by some image operators involving histograms, such as histogram equalization. A cumulative histogram is derived from an ordinary histogram, its particular value is the sum of all the preceding values in the original histogram up to the specified position. Hence, cumulative histograms are monotonically increasing functions (Burger and Burge, 2006).

2.3.4. Gamut

Human vision system is trichromatic by nature. Three different kinds of color photoreceptor cone cells on the retina allow humans to perceive electromagnetic radiation between 400 nm and 700 nm as color sensation. Video applications also generally deal with colors consisting of three components. Obtained from the scene, these components — mostly red, green and blue — can be transformed into another representation, optimized for certain purposes, such as processing, transmission and recording. The extent of colors that can be handled by an imaging system, i.e. captured by a camera or displayed on a monitor, is referred to as the color gamut.

Gamut specification is usually given by its primary chromaticities and a white point. Using additive reproduction, every color inside a gamut can be reproduced by a linear combination of its primaries. White point is the color that results adding primary chromaticities in equal quantities. Projected

onto the CIE chromaticity diagram, as explained in Section 3.1.2, gamuts of trichromatic systems form a triangle with the vertices representing primary chromaticities (see Figure 2.6).

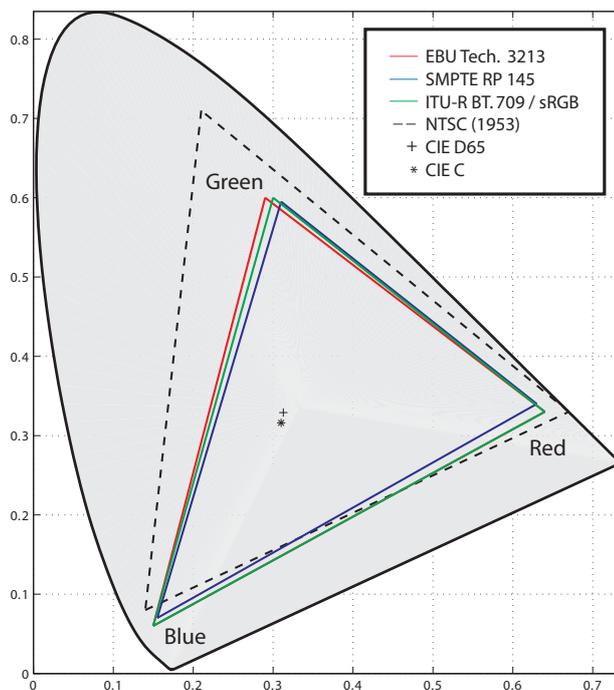


Figure 2.6.: RGB primaries of video standards.

In 1953, NTSC primaries and the white reference were specified according to phosphors used in CRT displays at that time. Due to the requirements for brighter displays, NTSC primaries have been superseded by EBU, SMPTE and ITU-R BT.709 primaries. When the European Broadcasting Union (EBU) introduced the PAL standard in the 1960s, phosphor technology was improved considerably. Primaries specified in EBU Tech. 3213 in 1975 are still used today in 576i systems, while 480i systems refer to SMPTE RP 145 primaries (Poynton, 2007). In consequence of a political compromise, ITU has retained EBU's red and blue primary chromaticities for HDTV, while green was averaged over EBU and SMPTE greens (ITU-R BT.709, 2002). Later, ITU-R BT.709 values have been adopted in the sRGB standard. Table A.1 on page 93 provides an outline on the specified primaries for the video and television standards. Although NTSC primaries from 1953 have no practical

use today, they are included in the Table A.1 and in Figure 2.6 for the sake of completeness.

Although visualizing color gamuts as projection onto the CIE chromaticity diagram is a common practice, evaluating colors against a gamut is misleading using this representation. While the chromaticity diagram is a two-dimensional graph, colors are three-dimensional quantities. When viewed in two dimensions, a color gamut, usually shaping a triangle, may enclose a certain color. Taking the third dimension — the luminance — into consideration, the gamut boundary forms an irregular shape. From this point of view it is apparent that this color can be out-of-gamut (Figure 2.7).

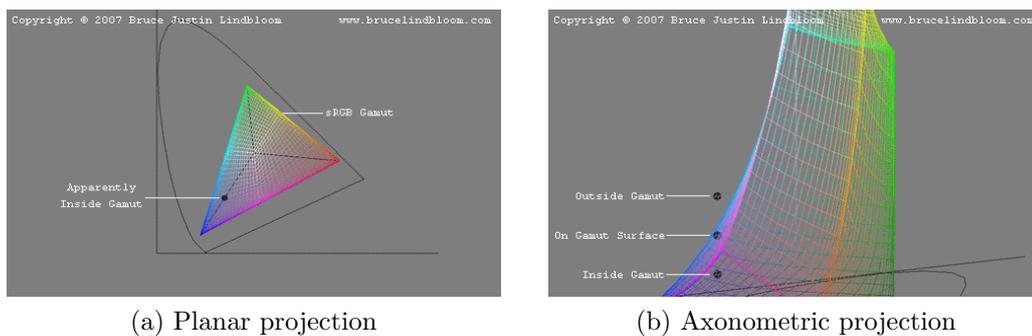


Figure 2.7.: Pitfall of judging in-gamut colors using CIE chromaticity diagram (Lindbloom, 2007).

For similar visualization reasons, waveform monitors and vectorscopes are not suitable for judging color gamut. Nevertheless, monitoring and measuring gamut violations is an important issue. When colors lie outside of a gamut, they can not be represented by a given system. Depending on the gamut mapping strategy or rendering intent the color is either mapped to another in-gamut color or clipped. Monitoring the color gamut is necessary to avoid the negative effect of out-of-gamut colors on the visual impact of a video picture.

In video applications, the term gamut is often used referring to the allowed encoding range of the signals, particularly in the $R'G'B'$ domain. A signal is considered illegal if it exceeds permitted limits. An invalid signal may remain

within allowed range, i.e. it can be legal, but it will produce an illegal output when converted into another encoding system. Illegal and invalid signals both cause *gamut violations* or *gamut errors*. Different encoding systems will be covered in Section 3.3 RGB Encoding on page 40 and Section 3.4 Luma and Chroma Encoding on page 44. Legal and valid signals will be discussed in detail later in Section 3.5 on page 50.

Tektronix has introduced different signal displays for visualizing gamut errors. For component signals *diamond* and *split diamond plots* indicate color gamut violations within non-linear $R'G'B'$ domain (Figure 2.8). Blue and green components form the upper diamond, while red and green signals are represented in the lower one. The vertical axis of the upper part is driven by the sum of blue and green signals, the horizontal one by their difference. The lower diamond is composed of red and green signals in the same manner. Gamut violation becomes apparent, when signal plot exceeds one or both diamond shapes (Tektronix Inc., 2002).

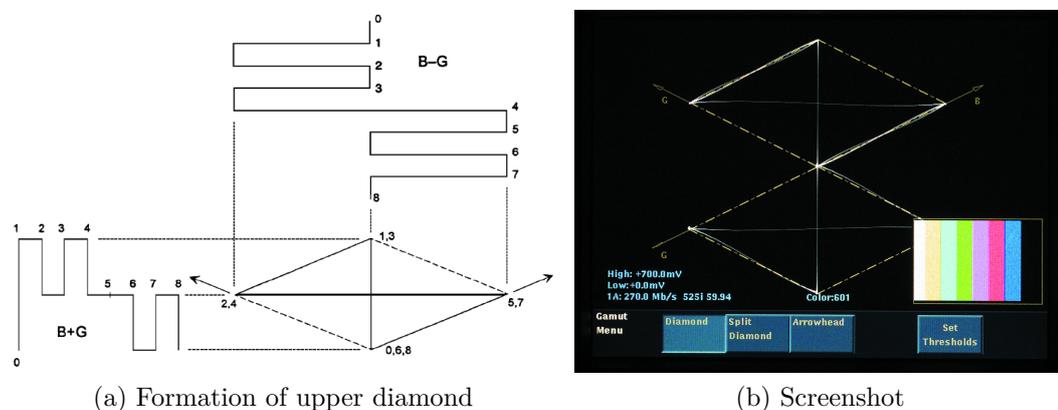


Figure 2.8.: Tektronix diamond display (Tektronix Inc., 2002).

Arrowhead plot displays luma on the vertical axis, while chroma subcarrier gain from the composite signal drives the horizontal axis (Figure 2.9). In this manner it is possible to determine, whether the signal will remain within gamut after PAL or NTSC encoding, without using composite encoder (Tektronix Inc., 2002).

2. Measurement Basics

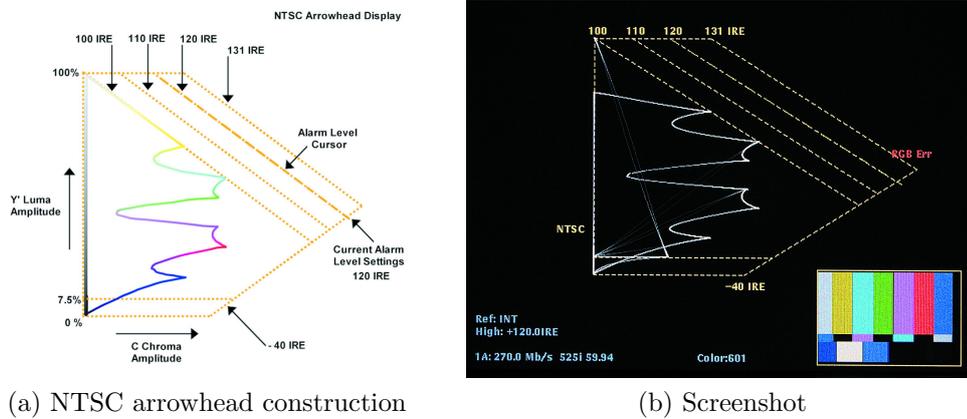


Figure 2.9.: Tektronix arrowhead display (Tektronix Inc., 2002).

While a component vector plot on a vectorscope only displays the color difference signal and an additional WFM is necessary for simultaneously monitoring of luma, *lightning plot* integrates both representations into a single diagram (Figure 2.10). The upper half of the plot represents P_B signal with respect to luma, while the lower part shows P_R signal versus inverted luma. Using color bars test signal it is possible to adjust levels on a composite signal observing the effects at a glance (Tektronix Inc., 2002).

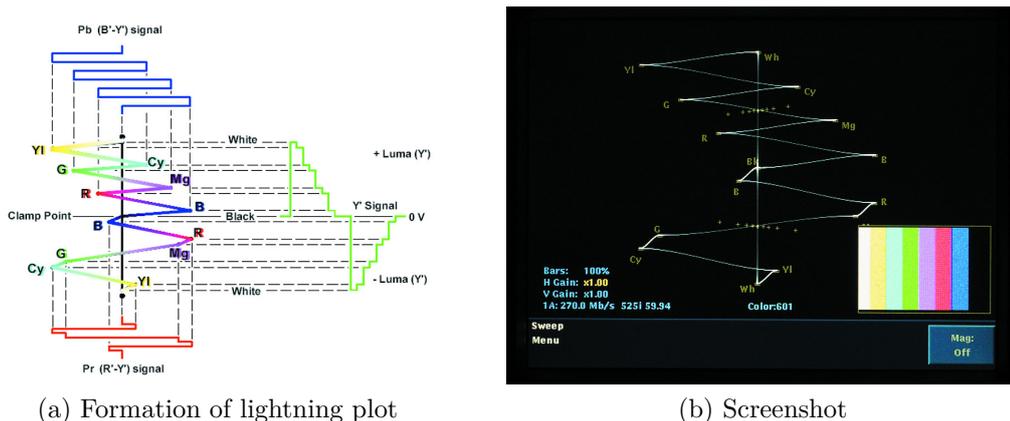


Figure 2.10.: Tektronix lightning display (Tektronix Inc., 2002).

Videotek, a subsidiary of Harris Corporation, has introduced *iris display* to judge color gamut of digital signals (Figure 2.11). It consists of a circular dia-

2. Measurement Basics

gram, where each composite color sample is mapped onto a dot pair. First dot consists of luma value plus half the color saturation, the other one is composed of luma minus half the saturation. Minimum and maximum allowed values are indicated by graticule circles. Gamut violation is noticeable, when the signal plot exceeds the circular boundaries. More information on iris display can be found in the application note by Harris (Harris Corporation, 2005).

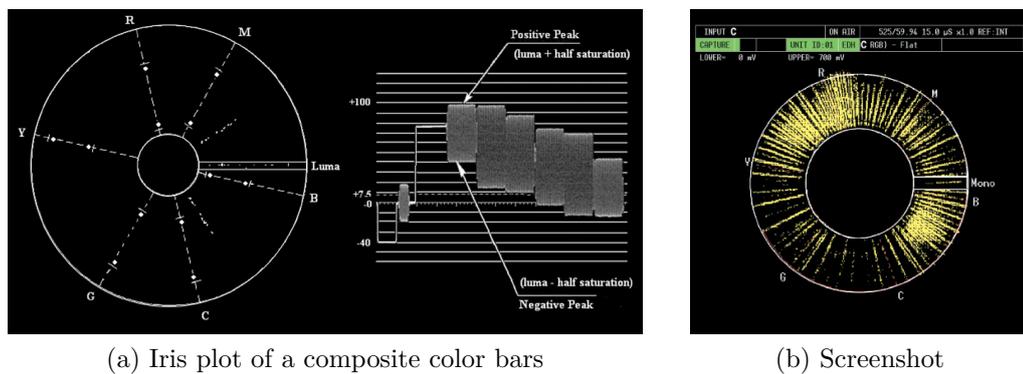


Figure 2.11.: Videotek iris gamut display (Harris Corporation, 2005).

3. Colorimetry in Video Applications

3.1. The CIE Colorimetric System

3.1.1. Spectrum and Tristimulus Values

The term *light* describes the spectral part of electromagnetic radiation that is visible to the human eye. Since it exhibits properties of both quanta and waves (wave-particle duality), and different parts of the light spectrum correspond to waves of different lengths, it is convenient to define spectral colors using wavelengths, typically given in nanometers. Denoted in such a manner, the part of the electromagnetic spectrum containing radiation that provokes a light sensation ranges from about 380 nm to 780 nm (Eichler et al., 1993).

Color of the emitted light can be specified by its *spectral power distribution* (SPD). Since the luminance of a color is mostly handled separately, absolute spectral power distributions are often normalized to unity or 100 at 560 nm (Berns, 2000) and form *relative SPDs* as illustrated in Figure 3.1a. A relative SPD describes the relative radiant power emitted by the source at each wavelength or band of wavelengths over the light spectrum. Illuminated objects absorb certain parts of illuminant's spectrum, reflecting or transmitting remaining spectral components that, processed by the HVS, result in a color sensation. Since the color of a reflective or transmissive surface is strictly associated with the illuminant's spectrum, *spectral reflectance*, respectively *trans-*

3. Colorimetry in Video Applications

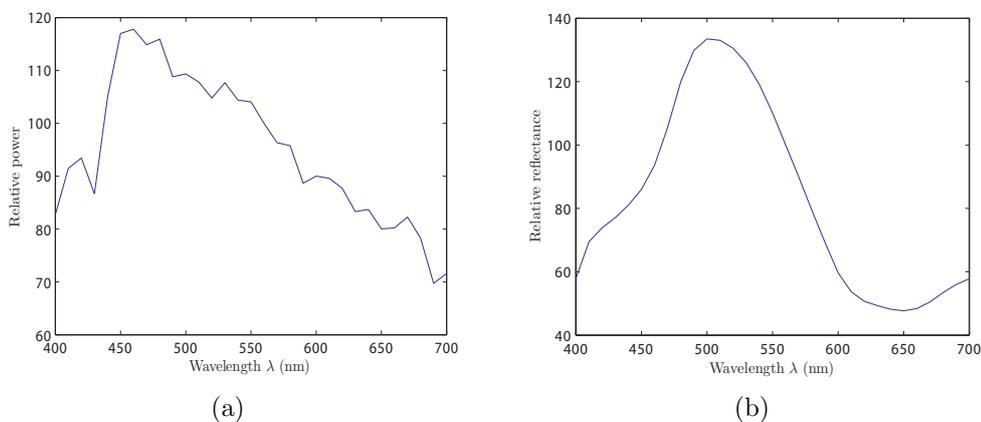


Figure 3.1.: Spectral representations: (a) SPD of the standard illuminant CIE D65, (b) spectral reflectance curve of the patch F1 (bluish green) of the X-Rite ColorChecker.

mittance curve provides detailed description of color’s properties (Figure 3.1b). It represents the amount of reflected or transmitted light at each wavelength as a percentage of the incident light.

Two materials with surfaces exhibiting the same spectral properties will cause identical color sensation under the same viewing conditions. Though different illuminants would affect the perceived color, the colors among themselves will still remain alike. They will have the same appearance observed by color-blind persons and animals as well. It seems natural to reproduce a color replicating its spectral reflectance curve. The two common techniques — the micro-dispersion method and the Lippmann method — are both photographic. Even though they provide a spectrally correct color reproduction, their feasibility for general use is very limited on account of bulky and expensive equipment and extremely long exposure times (Hunt, 2004). In the engineering practice, digital representation of spectral properties is also inconvenient because of the large volume of encoded data.

The idea of trichromacy in the human color vision can be traced back to the 17th century (Hunt, 2004). Subsequent physiological research has substantiated this assumption discovering cone and rod cells — photoreceptor cells

in the human retina. Cones are less sensitive to the light than the rod cells, but they allow the perception of color. Physiological research has proved the existence of three different types of absorbing pigments in the different types of the cone cells (Wyszecki and Stiles, 2000). One pigment absorbs long-wave light (L-cones), two others medium- and short-wave light (M- and S-cones). As a reaction on the absorption, cone cells generate electrical signals which, after running through complex preprocessing by other retinal cells, reach the brain via the optic nerve. Processed in the brain, these nerve impulses finally evoke the sensation of color. Owing to this knowledge, the idea to describe colors by three variables only, instead of dealing with their spectra, appears reasonable. *Colorimetry* is the science that describes and quantifies colors reducing the spectral power distributions to their physiological counterparts — *tristimulus values*. Therefore, the basic concept of trichromatic reproduction is: if suitable spectral conversion functions are used, three numerical values are necessary and sufficient to specify a color (Poynton, 2007).

Two colors with different SPDs but the same tristimulus values cause the same color sensation. This effect is referred to as *metamerism* or *metameric match*. Since colors of reflective and transparent surfaces also depend on the properties of the illuminant, it depends on the spectral properties of the light source, whether the illuminated metameric pair looks alike or not. Although much more feasible, metameric match is more sensitive to the changes in illumination than spectral reproduction.

3.1.2. Color Matching Functions and Standard Observers

In order to develop a colorimetric system describing colors with the tristimulus values, several investigations were carried out in the 1920s. William D. Wright and John Guild independently conducted color matching experiments in 1928-1929, respectively 1931. Their results laid the foundation for the CIE RGB color space that, later in 1931, provided the basis for the CIE XYZ color space. According to the *Grassmann's laws of additive color matching* (Berns, 2000), three primary light sources with variable intensities can be combined to evoke

3. Colorimetry in Video Applications

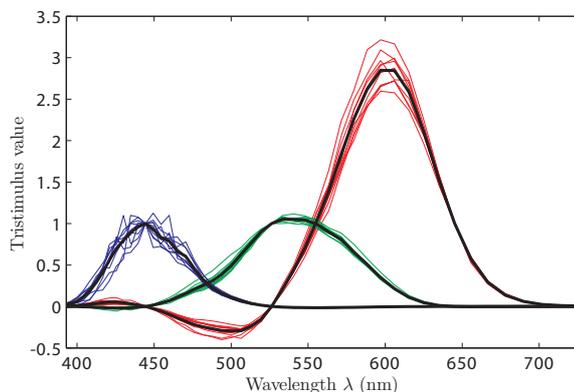


Figure 3.2.: Individual 2° RGB color matching functions according to Stiles and Burch (1955) for 10 subjects (University College London, 2010). Black line gives the mean value.

a specific color sensation. Wright and Guild asked the test persons to match the test color adjusting the intensities of three differently colored beams. For the experiments, the circular test field of 2 degrees of the visual angle was split into reference and test area and surrounded by darkness.

Matching some high saturated spectral colors was not possible using this method, but when the test color was desaturated with a certain amount of one primary beam, it was possible to reproduce it with the remaining two beams. In this case the amount of the desaturated color primary was assumed to be negative. Later experiments of Walter S. Stiles and Jennifer M. Burch in 1955 disclosed similar effects (Figure 3.2).

Although Wright used monochromatic light primaries, while Guild utilized wide-band sources, the correspondence between their experiments after the linear transformation into a common set of primaries was extremely good (Fairchild, 1998). These standard RGB primaries were defined in 1931 by the CIE as monochromatic light at 700.0 nm for red, 546.1 nm and 435.8 nm for green and blue respectively. The values for green and blue were chosen at these positions since they represent two of the spectral peaks of the mercury vapor discharge that could easily be reproduced at that time. The red value at 700 nm, which was difficult to reproduce in 1931 as a monochromatic beam,

was chosen because the HVS color perception is rather stable in this area of spectrum with respect to small wavelength variations. As mentioned before, a 500 nm sample spectral color can only be matched using an additive mixture of the green and blue primaries when a given amount of the red primary is used to desaturate the sample.

Using the CIE $\bar{r}(\lambda)$, $\bar{g}(\lambda)$ and $\bar{b}(\lambda)$ *color matching functions* (CMFs) it is possible to obtain the tristimulus color values for a given spectrum by multiplication with the amount of the spectral power at each wavelength (Grassmann's proportionality) and subsequently integrating the resulting curves (Grassmann's additivity) (Fairchild, 1998). Equations 3.1 describe this calculations for a stimulus with the spectral power distribution $\Phi(\lambda)$ within the range $[\lambda_1, \lambda_2]$.

$$\begin{aligned} R &= \int_{\lambda_1}^{\lambda_2} \Phi(\lambda)\bar{r}(\lambda)d\lambda \\ G &= \int_{\lambda_1}^{\lambda_2} \Phi(\lambda)\bar{g}(\lambda)d\lambda \\ B &= \int_{\lambda_1}^{\lambda_2} \Phi(\lambda)\bar{b}(\lambda)d\lambda \end{aligned} \tag{3.1}$$

For self-luminous stimuli, $\Phi(\lambda)$ represents their spectral radiance respectively their relative SPD. For reflective surfaces, $\Phi(\lambda)$ is defined as the product of the spectral reflectance function $R(\lambda)$ and the relative SPD of the light source $S(\lambda)$. For transparent materials $\Phi(\lambda)$ is calculated in the same way using their spectral transmittance curves $T(\lambda)$.

$$\begin{aligned} \Phi(\lambda)_{reflective} &= R(\lambda)S(\lambda) \\ \Phi(\lambda)_{transparent} &= T(\lambda)S(\lambda) \end{aligned} \tag{3.2}$$

Although it was possible to describe colors with their tristimulus values based on the RGB color matching functions, two concerns led to the development of a new colorimetric system. Since the red, green and blue CMFs exhibit both positive and negative values, design of related spectrophotometers is very complex. Furthermore, the CIE standard luminosity function $V(\lambda)$, defined

formerly in 1924, was based on other experiment conditions than the color matching experiments. Due to this fact, colors with the same tristimulus values could have different photometric values. In 1931 the CIE defined a new set of primaries for the standard 2° observer called X, Y and Z with the corresponding spectral weighting functions $\bar{x}(\lambda)$, $\bar{y}(\lambda)$ and $\bar{z}(\lambda)$. The XYZ color matching functions were defined numerically, exhibiting only non-negative values. The CIE standard luminosity function is incorporated and equals the $\bar{y}(\lambda)$ function. Equations 3.3 and 3.4 describe the transformation between the RGB and XYZ models (Berns, 2000).

$$\begin{bmatrix} \bar{x}_{\lambda=380} & \cdots & \bar{x}_{\lambda=780} \\ \bar{y}_{\lambda=380} & \cdots & \bar{y}_{\lambda=780} \\ \bar{z}_{\lambda=380} & \cdots & \bar{z}_{\lambda=780} \end{bmatrix} = \begin{bmatrix} 0.49000 & 0.31000 & 0.20000 \\ 0.17690 & 0.81240 & 0.01063 \\ 0.00000 & 0.01000 & 0.99000 \end{bmatrix} \cdot \begin{bmatrix} \bar{r}_{\lambda=380} & \cdots & \bar{r}_{\lambda=780} \\ \bar{g}_{\lambda=380} & \cdots & \bar{g}_{\lambda=780} \\ \bar{b}_{\lambda=380} & \cdots & \bar{b}_{\lambda=780} \end{bmatrix} \cdot \begin{bmatrix} n_{\lambda=380} \\ \cdots \\ n_{\lambda=780} \end{bmatrix} \quad (3.3)$$

$$n_{\lambda} = V(\lambda) / [0.17690 \bar{r}_{\lambda} + 0.81240 \bar{g}_{\lambda} + 0.01063 \bar{b}_{\lambda}] \quad (3.4)$$

The experiments leading to the 1931 CIE standard observer were only related to the areas of fovea corresponding to 2° of visual angle. In the 1950s, Stiles and Burch conducted experiments using larger angles of 10° and high levels of illumination. Speranskaya made similar efforts using considerably lower illumination levels. In 1964, the CIE standardized a set of color matching functions for the 10° standard observer averaging the two ascertained data sets by Stiles and Burch and Speranskaya. These CMFs, referred to as $\bar{x}_{10}(\lambda)$, $\bar{y}_{10}(\lambda)$ and $\bar{z}_{10}(\lambda)$, are recommended for use whenever color matching conditions exceed 4° of visual angle. Since $\bar{y}_{10}(\lambda)$ does not directly relate to $V(\lambda)$ and hence does not represent the luminance, 10° CMFs should be used with caution (Berns, 2000). Due to transformation mathematics, both function

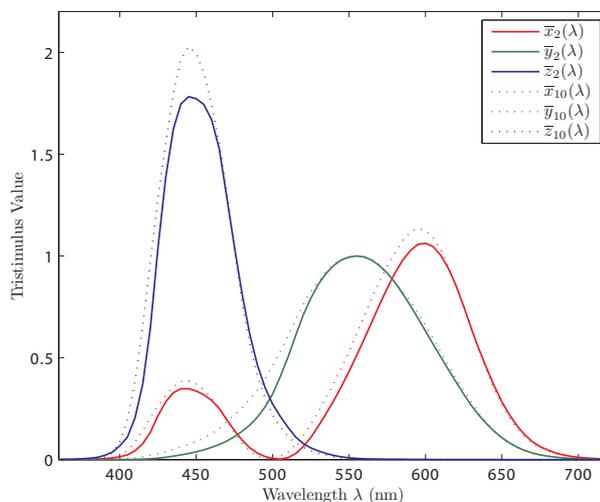


Figure 3.3.: The XYZ color matching functions for the 1931 CIE 2° and 1964 CIE 10° standard observers.

curves $\bar{y}(\lambda)$ and $\bar{y}_{10}(\lambda)$ are bimodal exhibiting two peaks as can be seen in Figure 3.3.

In case of continuous data, tristimulus values X , Y and Z for the CIE 2° and 10° standard observers can be calculated from the spectral distributions in the same way as using $\bar{r}(\lambda)$, $\bar{g}(\lambda)$ and $\bar{b}(\lambda)$ (Equation 3.1 on page 29). Discrete data can be obtained using quantization and linear matrix operations.

Sometimes it is necessary to separate hue and luminance. Using a one-point perspective projection of the data from the XYZ color space into an xyY domain allows for this separation. The hue is represented by the *chromaticity coordinates* x and y and the luminance axis Y completes the model. Since the pure chromaticity coordinates do not provide any luminance data, they are insufficient to fully specify a color stimulus (cf. Figure 2.7 on page 21) and should be used with care. An xyY triple should be used to completely describe a color. The transformation from XYZ color space to xy -chromaticity plane is described in Equation 3.5 (Fairchild, 1998).

$$\begin{aligned} x &= \frac{X}{X + Y + Z} \\ y &= \frac{Y}{X + Y + Z} \end{aligned} \tag{3.5}$$

Chromaticity values are abstract quantities and have no direct physical interpretation. A color is represented as a point within the horseshoe-shaped spectral locus in the CIE xy -chromaticity diagram as can be seen in the Figure 2.6 on page 20. It should be noted that the CIE xy -chromaticity diagram is not perceptually uniform, hence it can not be used when judging color differences.

3.1.3. Color Temperature

One of the most important artificial light sources is a tungsten filament lamp. When heated to a certain temperature, which depends on the resistance of the filament and the applied voltage, it emits light. The color of this light is closely related to the filament's temperature. Although a tungsten lamp is not a black-body or Planckian radiator, their emitted relative spectral power distributions are very identical. The SPDs of the light emitted by such sources are given by Planck's Radiation Law, as described in Equation 3.6. $P(\lambda)$ is the power in watts emitted per square centimeter of a surface per micrometer band of a wavelength λ . T is a physical temperature given in kelvins, e is the Euler's constant equals to approx. 2.718. The Planckian constants are $c_1 \approx 3,7418 \text{ W} \cdot \text{m}^2$, $c_2 = 1,4388 \cdot 10^{-2} \text{ m} \cdot \text{K}$ (Eichler et al., 1993).

$$P(\lambda, T) = \frac{c_1}{\pi \lambda^5} \cdot \frac{1}{e^{c_2/\lambda T} - 1} \tag{3.6}$$

For small values of λT , i.e. when $T < 5000 \text{ K}$ and λ is in the range of the visible spectrum, $e^{c_2/\lambda T}$ becomes significantly larger than one, thus the Plank's

3. Colorimetry in Video Applications

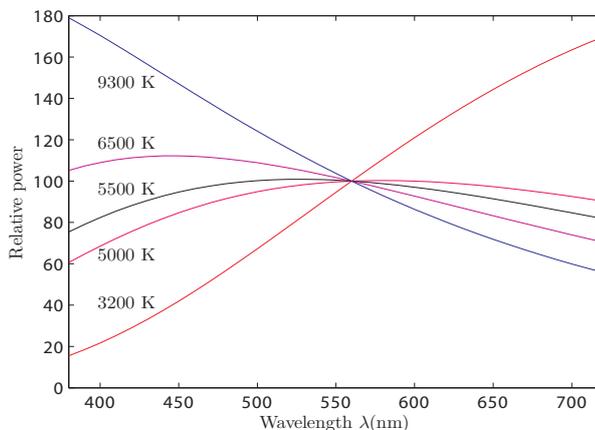


Figure 3.4.: Relative SPDs of Planckian radiators, normalized to equal power at 560 nm.

law can be approximated by Wien's Radiation Law as stated in Equation 3.7 (Eichler et al., 1993).

$$P(\lambda) = \frac{c_1}{\pi\lambda^5 e^{c_2/\lambda T}} \quad (3.7)$$

This approximation is adequate for tungsten filament lamps and the visible part of the light spectrum with accuracy of about 1%. With increasing temperature of the Planckian radiator, the emitted energy also increases and the peak shifts towards shorter wavelengths. Figure 3.4 illustrates this behavior on normalized SPDs.

If the absolute temperature of a black body, also referred to as its *color temperature*, is given, it is possible to reconstruct its spectral distribution. Since Planckian emitters are rare and can only be found in specialized laboratories, *correlated color temperature* (CCT) is a useful quantity to describe an illuminant. It refers to the temperature of a black-body radiator that has nearly the same color (Hunt, 2004). CCTs can be even assigned to light sources that emit light regardless of their physical temperature, such as fluorescent lights. The color of the sky, for instance, changes during the day. Correlated color temperature can be used to describe sky's hue, although the color changes are not related to a change of the temperature, but to the Rayleigh scattering in the atmosphere.

Generally, CCT is stated in kelvins (K) or alternatively in micro-reciprocal degrees (mireds). The latter unit is derived from kelvin and can be calculated as shown in Equation 3.8, where T is the color temperature in kelvins. It is a common unit when dealing with spectral power conversion filters, e.g. in a broadcasting studio.

$$M = \frac{1\,000\,000}{T} \quad (3.8)$$

Correlated color temperature gives an idea of a hue of an illuminant, but since it is an approximation related to a black-body radiator, it is more appropriate to describe the hue using CIE chromaticity coordinates, CIE XYZ tristimulus values or even to use relative spectral power distributions.

3.1.4. Standard Illuminants

In order to describe a color of a not self-luminous source it is important to have detailed knowledge of the illuminant used. The CIE have defined a number of spectral power distributions, referred to as CIE standard illuminants, to provide reference spectra for colorimetric issues. The illuminants are denoted by a letter or a letter-number combination. Their SPDs are normalized to a value of 100 at a wavelength of 560 nm in this work. The diagrams in this section were partly created using the *OptProp Toolbox* for MATLAB by Jerker Wågberg (More Research and DPC, Mid Sweden University).

CIE Illuminants A, B and C. The first three standard illuminants were introduced in 1931. Illuminant A represents an incandescent tungsten filament lamp.

“[CIE standard illuminant A] is intended to represent typical, domestic, tungsten-filament lighting. Its relative spectral power distribution is that of a Planckian radiator at a temperature of approximately 2856 K. CIE standard illuminant A should be used in all applications of colorimetry involving the use of incandescent lighting,

unless there are specific reasons for using a different illuminant.”
(CIE S005, 1999)

Illuminants B and C represent direct and shady daylight respectively. They can be derived from illuminant A using liquid conversion filters with high absorbance in the red part of the spectrum. More on these filters can be found in Hunt (2004). Due to their deficiency at wavelengths below 400 nm, that are important when e.g. working with fluorescent optical brighteners, illuminants B and C are considered deprecated in favor of the CIE D-series of illuminants. Practical realization of CIE illuminants A, B and C is possible since it is defined in the standard. Relative spectral power distributions of the illuminants A and C are depicted in Figure 3.5.

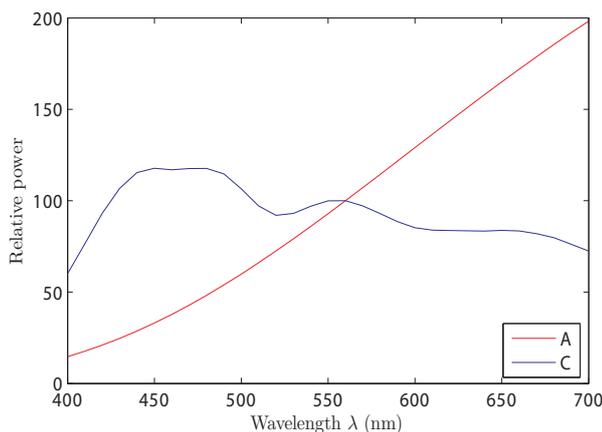


Figure 3.5.: SPD's of the CIE Standard Illuminants A and C.

CIE Illuminant Series D. These is a series of illuminants, that has been statistically defined in 1964 upon numerous measurements of real daylight. Although mathematically described, they can hardly be realized artificially. The correlated color temperatures of the commonly used illuminants D_{50} , D_{55} and D_{65} are slightly different to the values suggested by their names. Due to the revision of an estimate of one of the constant factors in Planck's law after the standards were defined, the correlated color temperature was shifted

a little. For example, the CCT of D_{50} is 5003 K and that of D_{65} is 6504 K. SPD's of some illuminants of the CIE Series D are illustrated in Figure 3.6.

“[CIE standard illuminant D_{65}] is intended to represent average daylight and has a correlated colour temperature of approximately 6500 K. CIE standard illuminant D_{65} should be used in all colorimetric calculations requiring representative daylight, unless there are specific reasons for using a different illuminant. Variations in the relative spectral power distribution of daylight are known to occur, particularly in the ultraviolet spectral region, as a function of season, time of day, and geographic location. However, CIE standard illuminant D_{65} should be used pending the availability of additional information on these variations.” (CIE S005, 1999)

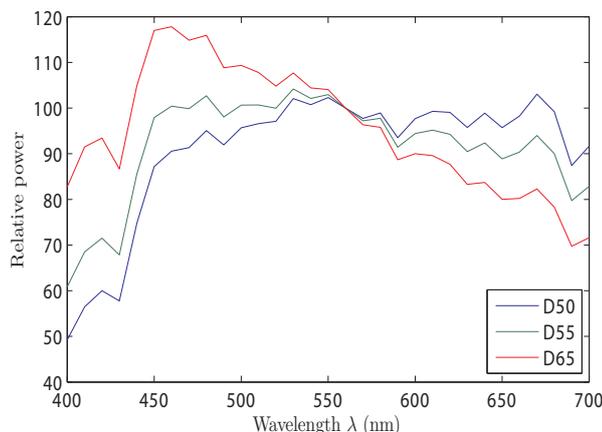


Figure 3.6.: SPD's of the CIE Standard Illuminants D_{50} , D_{55} and D_{65} .

CIE Illuminant E. This is a hypothetical reference radiator. All wavelengths in CIE illuminant E are weighted equally with a relative spectral power of 100.0. Since it is not a Planckian radiator, no color temperature is given, however it can be approximated by a CIE D illuminant with a CCT of 5455 K. Canonical standard illuminant D_{55} is the closest to match its color temperature. The SPD of this hypothetical radiator is shown in Figure 3.7.

3. Colorimetry in Video Applications

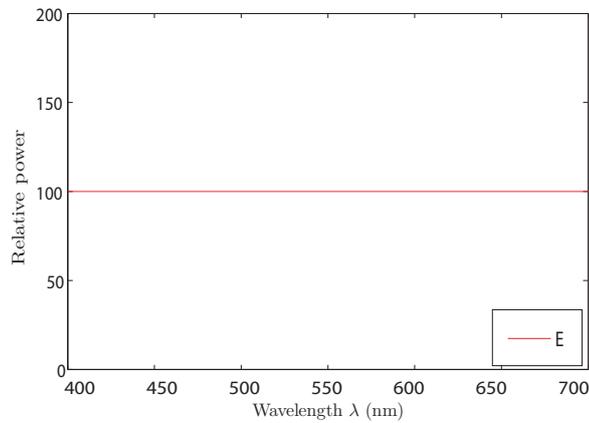


Figure 3.7.: SPD of the CIE Standard Illuminant E.

CIE Illuminant Series F. Twelve F illuminants represent typical relative SPDs for different types of fluorescent light sources. Illuminant F_2 , for instance, describes a cool-white light with the CCT of 4230 K, F_8 simulates daylight standard illuminant D_{50} at 5000 K and F_{11} stands for a triband source with 4000 K. Such triband sources are popular because of their color rendition properties and their light efficiency (Fairchild, 1998). Relative spectral powers of the F-series illuminants discussed here are illustrated in Figure 3.8.

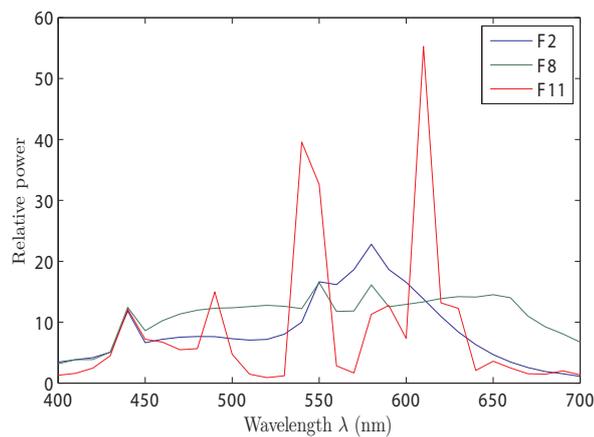


Figure 3.8.: SPD's of the CIE Standard Illuminants F_2 , F_8 and F_{11} .

Relative spectral distribution of an illuminant, weighted with the CIE color matching functions for a 2° standard observer and converted to the XYZ tristimulus values, is referred to as illuminant's white point.

3.2. Color Spaces

3.2.1. Definition

The terms *color model* and *color space* are often used carelessly. To avoid confusion, definitions of these terms have to be provided. In the following, *color model* refers to an abstract mathematical concept defining the axes of a coordinate system for the representation of colors. For instance, the *RGB* color model states that a color can be defined as a linear combination of the primary red, green and blue colors, subtractive *CMYK* model for color printing defines colors consisting of four primary components. Numerical representation of colors is meaningless referring to a color model without a well-defined scale. A *color space* refers to a color model replenished with definitions of how the numerical values have to be interpreted. In this work, the definition of a color space implies the information about the related color model, coordinates of its primaries and scaling information, white point data and description of any non-linear transfer characteristics.

3.2.2. Digital Image Color Workflow

Image data obtained by a color camera originate from the sampling by a sensor with applied color filters. Spectral attributes of these filters, mostly with passbands in red, green and blue parts of the light spectrum, are different to those of the phosphors or filters in computer displays or TV sets, although both are based on the RGB color model. Moreover, the color representation of a video signal is converted for transmission to the luma/chroma model for historical

and engineering reasons. For the purpose of predictable color rendition, well-defined transformations between different color spaces are necessary.

The generalization of the electronic image data flow considering the color is described by Süsstrunk et al. (1999) and illustrated in Figure 3.9. The RGB data obtained by a camera initially resides in the *sensor space*. This color space is device-dependent and scene-related, i.e. associated to the RGB primaries of the sensor and related to the original illumination. The spectral filter characteristics of capturing devices are chosen on the basis of engineering considerations and technical feasibility. Thus, there is no standard sensor color space now and there will unlikely be one (Süsstrunk et al., 1999).

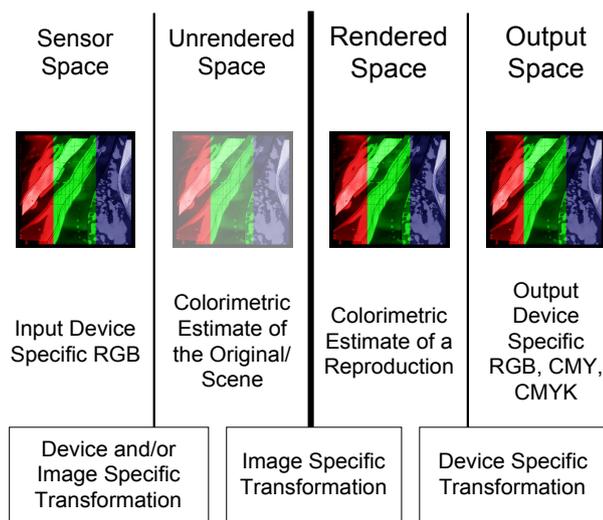


Figure 3.9.: Schematic representation of a digital image color flow (Süsstrunk et al., 1999).

An optional transformation to an *unrendered space* can be carried out. Such device-independent color space represents an estimate of the scene’s colorimetry. The transformation from a sensor space is device- and scene-specific. Its quality depends on the adequate scene-adapted white point estimation and the proper transformation matrix. Although an image in an unrendered space can not be viewed without further conversion, it allows for tone and color processing with any rendering intent for any reproduction device and appliance of

color appearance models. *CIE XYZ* and *CIE L*a*b** are device-independent color spaces that can be used as an unrendered space.

The colorimetric estimation of a color reproduction takes place in a *rendered space*, for which several standards have been established. Transformed from either sensor space or unrendered space, it is based on the characteristics of a real or virtual output device. In the computer graphics, *sRGB* is an example of a rendered color space. In video applications, RGB color spaces as defined in EBU Tech. 3213-E (PAL), SMPTE RP 145 (NTSC) or in ITU-R BT.709 (2002) (HDTV) as well as their luma/chroma representations, e.g. $Y'_{601}UV$ or $Y'_{709}C_B C_R$, can exemplify a rendered color space. The luma subscripts refer to the different ITU recommendations as explained later in detail. Usually, the chosen characteristics closely resemble that of generic output devices. Since some transformations into a rendered space require gamut and dynamic range compression to match the properties of an output device, this conversion is non-reversible without knowledge of the rendering mathematics used (Süsstrunk et al., 1999).

When gamut, dynamic range and viewing conditions of the color space of an output device significantly differ from the image representation in a rendered color space, an additional conversion into an *output space* is necessary. This kind of a color space is device-dependent. The transformation from a rendered space maps the color data to match the device's primaries and transfer function. It is unusual to exchange data in device-specific output color space.

3.3. RGB Encoding

3.3.1. Computer Graphics

The idea behind the color spaces based on the RGB model is to simulate the detection of colors by the human eye. This trichromatic theory was discussed in Section 3.1.1 Spectrum and Tristimulus Values. RGB encoding is a powerful

data compression method compared to spectral representation. Nevertheless, high correlation between red, green and blue channels leads to high data redundancy. Some representatives of the rendered RGB color space family — sRGB, AdobeRGB and ROMM RGB — are discussed below.

sRGB. This color space is specified in the standard IEC 61966-2-1 by the International Electrotechnical Commission. It was designed to match typical non-professional viewing conditions as available in home and office environments. For this reason, the RGB primary values have been selected according to ITU-R BT.709, representing generic monitor primaries. sRGB is intended for data exchange in multimedia environments. Since sRGB primaries and the white point closely match characteristics of typical monitors and TV sets, the transformation into a device-dependent output space can often be omitted. The transformation from unrendered CIE XYZ tristimulus color space into linear RGB domain is given in Equation 3.9 (Stokes et al., 1996).

$$\begin{bmatrix} R_{sRGB} \\ G_{sRGB} \\ B_{sRGB} \end{bmatrix} = \begin{bmatrix} 3.2410 & -1.5374 & -0.4986 \\ -0.9692 & 1.8760 & 0.0416 \\ 0.0556 & -0.2040 & 1.0570 \end{bmatrix} \cdot \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \quad (3.9)$$

Let C represent a single linear color channel R_{sRGB} , G_{sRGB} or B_{sRGB} . The non-linear transfer function into a gamma corrected channel C'_{sRGB} is then described in Equation 3.10 (Stokes et al., 1996).

$$C'_{sRGB} = \begin{cases} 12.92C_{sRGB} & 0 \leq C_{sRGB} < 0.0031 \\ 1.055C_{sRGB}^{1/\gamma} - 0.055 & 0.0031 \leq C_{sRGB} < 1 \end{cases} \quad (3.10)$$

$$\gamma = 2.4$$

The non-linear color data is encoded in sRGB color space with 8-bit precision.

AdobeRGB. This color space was introduced in 1998 by Adobe Systems Inc. with the intent to provide a suitable working color space for print production (Adobe Systems Inc., 2005). Its gamut is larger than that of sRGB, especially

in cyan and green areas, allowing for the representation of the colors reproducible with the most CMYK color printers at that time. Similar to sRGB, AdobeRGB is encoded in 8-bit non-linear representation. Equations 3.11 and 3.12 show the conversion matrix and the non-linear gamma function, respectively. The gamma transfer function of AdobeRGB is not defined piecewise, in contrast to other RGB color spaces discussed in this work.

$$\begin{bmatrix} R_{AdobeRGB} \\ G_{AdobeRGB} \\ B_{AdobeRGB} \end{bmatrix} = \begin{bmatrix} 2.04159 & -0.56501 & -0.34473 \\ -0.96924 & 1.87597 & 0.04156 \\ 0.01344 & -0.11836 & 1.01517 \end{bmatrix} \cdot \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \quad (3.11)$$

$$\begin{aligned} C'_{AdobeRGB} &= C_{AdobeRGB}^{1/\gamma} \\ \gamma &= 2.2 \end{aligned} \quad (3.12)$$

ROMM RGB. Reference Output Medium Metric (ROMM) RGB, sometimes referred to as ProPhoto RGB, is a wide-gamut color space, designed to represent image data without typical device-specific gamut limitations. It is intended to be used for storing, exchanging and manipulating of images. For proper encoding of the large gamut, ROMM RGB allows optional 12-bit and 16-bit quantization. The conversion matrix and the transfer function of ROMM RGB are given in Equations 3.13 and 3.14.

$$\begin{bmatrix} R_{ROMM} \\ G_{ROMM} \\ B_{ROMM} \end{bmatrix} = \begin{bmatrix} 1.3460 & 0.2556 & 0.0511 \\ -0.5446 & 1.5082 & 0.0205 \\ 0.0000 & 0.0000 & 1.2123 \end{bmatrix} \cdot \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \quad (3.13)$$

$$C'_{ROMM} = \begin{cases} 0 & C_{ROMM} < 0.0 \\ 16 & 0.0 \leq C_{ROMM} < E_t \\ C_{ROMM}^{1/\gamma} & E_t \leq C_{ROMM} < 1.0 \\ 1 & C_{ROMM} \geq 1.0 \end{cases}$$

$$\gamma = 1.8$$

$$E_t = 16^{\gamma/(1-\gamma)} \approx 0.001953 \tag{3.14}$$

Chromaticity coordinates of the described color spaces are summarized in the Table A.2 in the Appendix. For further information about sRGB, AdobeRGB and ROMM RGB color spaces, e.g. concerning viewing conditions and encoding to digital code values, please refer to Stokes et al. (1996), Adobe Systems Inc. (2005) and Spaulding et al. (2000), respectively.

3.3.2. Video Applications

RGB primaries for video applications are defined in different standards. European Broadcasting Union introduced the chromaticities for studio monitors in EBU Tech. 3213-E (1975). For standard definition television, ITU-R BT.601 solely defines encoding parameters of digital television. It does not refer to any color space used to form component signals. Thus, it is implicit allowed to use both EBU Tech. 3213-E and ITU-R BT.709 primaries to calculate non-linear $R'G'B'$ data. The conversion matrix between CIE XYZ color space and linear EBU RGB signals is given in Equation 3.15 (Ford and Roberts, 1998; Pascale, 2003).

$$\begin{bmatrix} R_{EBU} \\ G_{EBU} \\ B_{EBU} \end{bmatrix} = \begin{bmatrix} 3.0629 & -1.3932 & -0.4758 \\ -0.9693 & 1.8760 & 0.0416 \\ 0.0679 & 0.2289 & 1.0694 \end{bmatrix} \cdot \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \tag{3.15}$$

For high definition environment, ITU-R BT.709 defines both encoding and color transformation coefficients. The chromaticity coordinates of the primaries are identical to those defined in sRGB standard. Thus, the same color

space conversion function is used when transforming from CIE XYZ (see Equation 3.9 on page 41). However, the transfer function in ITU-R BT.709 is not the same as defined for sRGB (see Equation 3.10 on page 41). The transfer function from a linear color signal C to a non-linear quantity C' according to ITU-R BT.709 is shown in Equation 3.16 (Ford and Roberts, 1998; Poynton, 2007).

$$C'_{709} = \begin{cases} 4.5C_{709} & 0 \leq C_{709} < 0.018 \\ 1.099C_{709}^{1/\gamma} - 0.099 & 0.018 \leq C_{709} < 1 \end{cases}$$

$$\gamma = 2.2 \tag{3.16}$$

Transfer functions for standard definition television have been poorly specified. According to Poynton (2007), it is appropriate to use the transfer function defined in ITU-R BT.709 as described above.

3.4. Luma and Chroma Encoding

3.4.1. Luma and Chroma

The term *relative luminance*, denoted Y , has been introduced by the CIE in 1931. It can be calculated by integrating the relative SPD filtered with the color matching function $\bar{y}(\lambda)$, which represents the CIE standard luminosity function $V(\lambda)$, or from the given linear tristimulus RGB components using a linear matrix operation. Video and broadcasting applications often transform the linear RGB signals obtained by the camera into one signal representing lightness and two signals representing color. For engineering reasons, video systems first apply a non-linear transfer function, referred to as *gamma correction*, to the red, green and blue channels (Poynton, 2007). The weighted sum of these non-linear R' , G' and B' signals is then computed and forms *luma*, denoted Y' . The prime symbol signalizes the non-linear origin of a quantity. However, it is usually omitted denoting color difference signals in-

troduced next, since no practical image coding system employs non-linear color differences (Poynton, 2007). It is important to clearly distinguish between luminance and luma to avoid ambiguity. For more information on these terms and definitions please refer to the appendix “YUV and luminance considered harmful” in Poynton (2007).

After the luma signal is computed from the gamma corrected $R'G'B'$ components, two color difference signals $B' - Y'$ and $R' - Y'$ are calculated. These two color difference signals, also referred to as *chroma*, are subsequently scaled and often subsampled. In component signal applications, scaled chroma signals are referred to as P_B and P_R for analog and C_B and C_R for digital video. Using composite signals, chroma is referred to as U and V signals forming the modulated chroma part of the video signal. The analog composite color coding system $Y'IQ$, where the chroma basically consists of the rotated and axis-exchanged U and V components, introduced in 1953 for the NTSC standard, is now considered obsolete. Although the image coding systems $Y'P_BP_R$, $Y'C_BC_R$, $Y'UV$ and $Y'IQ$ are all derived from the color difference signals $B' - Y'$ and $R' - Y'$, they all differ in scaling factors and offsets.

3.4.2. Standard Definition TV

ITU Recommendation ITU-R BT.601 defines weighting factors to derive luma and chroma components from a non-linear $R'G'B'$ triple. These coefficients can be applied in the standard definition domain. They can be easily implemented as a linear transformation as shown in equation 3.17. Since ITU introduced a different set of luma transformation coefficients for HDTV, the related ITU recommendation is labeled by a subscript to maintain clarity. The matrices in this section are given according to normalized $R'G'B'$ signals taking values between 0 and 1.

$$\begin{bmatrix} Y'_{601} \\ B' - Y'_{601} \\ R' - Y'_{601} \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ -0.299 & -0.587 & 0.886 \\ 0.701 & -0.587 & -0.114 \end{bmatrix} \cdot \begin{bmatrix} R' \\ G' \\ B' \end{bmatrix} \quad (3.17)$$

While the luma component can take values between 0 and 1, the color difference components have inconvenient limits of ± 0.886 for $B' - Y'_{601}$ and ± 0.701 for $R' - Y'_{601}$. These component signals have to be scaled to be in accordance with the standards for analog and digital video.

Color difference components for analog standard definition television nominally have the same excursions as the luma component. An offset forces each of the chroma signals to range between -0.5 and 0.5 of the luma excursion. These components, referred to as P_B and P_R , can be directly calculated from non-linear $R'G'B'$ data as shown in equation 3.18. The transformation matrix already includes scaling factors as well as the offset.

$$\begin{bmatrix} Y'_{601} \\ P_B \\ P_R \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ -0.168736 & -0.331264 & 0.5 \\ 0.5 & -0.418688 & -0.081312 \end{bmatrix} \cdot \begin{bmatrix} R' \\ G' \\ B' \end{bmatrix} \quad (3.18)$$

As discussed in section 2.2 Units and Levels, digital component signals for SDTV exhibit foot- and headroom. Furthermore, digital representation does not allow for the employment of negative values. For example, in 8-bit domain luma component has a value range of 219 digital codes between 16 and 235, chroma signals C_B and C_R have a positive offset of 128 and reference values of 16 and 240. Equation 3.19 introduces the transformation matrix to obtain 8-bit $Y'C_BC_R$ signal from gamma corrected $R'G'B'$ channels. To obtain this matrix, the scaling factors from equation 3.18 are multiplied by the excursion ranges of 219 for luma and 224 for each chroma component. The additional offset vector ensures proper foot- and headrooms and avoids negative values. For more transformation matrices, concerning 10-bit domain, $R'G'B'$ data with different value ranges or full-range $Y'C_BC_R$ coding, please refer to Poynton (2007).

$$\begin{bmatrix} Y'_{601} \\ C_B \\ C_R \end{bmatrix} = \begin{bmatrix} 16 \\ 128 \\ 128 \end{bmatrix} + \begin{bmatrix} 65.481 & 128.553 & 24.966 \\ -37.797 & -74.203 & 112 \\ 112 & -93.786 & -18.214 \end{bmatrix} \cdot \begin{bmatrix} R' \\ G' \\ B' \end{bmatrix} \quad (3.19)$$

So far, the transformation matrices discussed in this section are used to form component video signals. One luma and two chroma channels are employed to deal with the video data. However, broadcasting standards PAL and NTSC both employ chroma modulation techniques to form one composite video signal, containing luma as well as chroma information. Both standards refer to $Y'UV$ image coding system derived from color difference components $B' - Y'_{601}$ and $R' - Y'_{601}$. The components U and V are used in an intermediate step to modulate chroma signal C , that, combined with the luma channel, finally forms one composite signal. The chroma components in the analog composite video domain are scaled, so that the resulting composite signal has an excursion from $-1/3$ to $+4/3$ of the luma signal excursion (Poynton, 2007). To ensure this, color difference signals $B' - Y'_{601}$ and $R' - Y'_{601}$ are multiplied by the factors of 0.492111 and 0.877283, respectively. Incorporating these scaling factors from the SMPTE 170M standard into Equation 3.17 on page 45, the transformation matrix shown in Equation 3.20 can be derived (Poynton, 2007).

$$\begin{bmatrix} Y'_{601} \\ U \\ V \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ -0.147141 & -0.288869 & 0.436010 \\ 0.614975 & -0.514965 & -0.100010 \end{bmatrix} \cdot \begin{bmatrix} R' \\ G' \\ B' \end{bmatrix} \quad (3.20)$$

For the sake of completeness, the deprecated composite NTSC coding system $Y'IQ$ is described below. It was considered too expensive to implement in decoders and in 1990 SMPTE accepted U and V components encoding. The chroma channels I (in-phase) and Q (quadrature) were filtered to have different bandwidths with the intent to improve color rendition of the transmitted signal. Narrow-band component Q was limited to 600 kHz while wide-band I had an upper limit of 1.3 MHz. These components were computed from the $Y'UV$ model rotating the U and V components by 33° and subsequently exchanging the horizontal and vertical axes (Poynton, 2007). These rotation and

exchange, related to Equation 3.20 are incorporated to form the transformation matrix as stated in Equation 3.21 (Poynton, 2007).

$$\begin{bmatrix} Y'_{601} \\ I \\ Q \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ 0.595901 & -0.274557 & -0.321344 \\ 0.211537 & -0.522736 & 0.311200 \end{bmatrix} \cdot \begin{bmatrix} R' \\ G' \\ B' \end{bmatrix} \quad (3.21)$$

3.4.3. High Definition TV

Luma coding coefficients for high definition television based on the ITU-R BT.709 differ from the factors defined in ITU-R BT.601 for standard definition video. The new set of weighting factors was introduced to match the chosen HDTV primaries theoretically. Except for this fact, the methodology to form analog and digital signals has been retained.

Equation 3.22 gives the transformation matrix used to obtain color difference signals $B - Y'_{709}$ and $R - Y'_{709}$ from non-linear $R'G'B'$ values ranging from zero to unity.

$$\begin{bmatrix} Y'_{709} \\ B' - Y'_{709} \\ R' - Y'_{709} \end{bmatrix} = \begin{bmatrix} 0.2126 & 0.7152 & 0.0722 \\ -0.2126 & -0.7152 & 0.9278 \\ 0.7874 & -0.7152 & -0.0722 \end{bmatrix} \cdot \begin{bmatrix} R' \\ G' \\ B' \end{bmatrix} \quad (3.22)$$

The analog color representation in high-definition video, that is conform to ITU-R BT.709, is employed using P_B and P_R chroma signals. Equation 3.23 depicts the formation of the analog components.

$$\begin{bmatrix} Y'_{709} \\ P_B \\ P_R \end{bmatrix} = \begin{bmatrix} 0.2126 & 0.7152 & 0.0722 \\ -0.114572 & -0.385428 & 0.5 \\ 0.5 & -0.454153 & -0.045847 \end{bmatrix} \cdot \begin{bmatrix} R' \\ G' \\ B' \end{bmatrix} \quad (3.23)$$

For 8-bit digital applications, HDTV image data can be coded using the offset vector and the transformation matrix as shown in Equation 3.24.

$$\begin{bmatrix} Y'_{709} \\ C_B \\ C_R \end{bmatrix} = \begin{bmatrix} 16 \\ 128 \\ 128 \end{bmatrix} + \begin{bmatrix} 46.559 & 156.629 & 15.812 \\ -25.664 & -86.336 & 112.000 \\ 112.000 & -101.730 & -10.270 \end{bmatrix} \cdot \begin{bmatrix} R' \\ G' \\ B' \end{bmatrix} \quad (3.24)$$

The luma transformation coefficients can be found in the recommendation ITU-R BT.709 (2002) as well as in Steinberg (1997). Additional transformation matrices are listed in Poynton (2007).

3.4.4. Conversion between SD- and HDTV

Due to different luma weighting factors defined in recommendations ITU-R BT.601 and ITU-R BT.709 (see Equation 3.25), image data has to be re-coded when converting from standard-definition to HDTV and vice versa.

$$\begin{aligned} Y'_{601} &= 0.299R' + 0.587G' + 0.114B' \\ Y'_{709} &= 0.2126R' + 0.7152G' + 0.0722B' \end{aligned} \quad (3.25)$$

The transformation can be performed using an intermediate step, first converting back to $R'G'B'$ and then calculating the new luma and chroma signal values. Combined, these two steps can be expressed in form of a multiplication with a single 3×3 matrix. Equations 3.26 and 3.27 outline the resulting conversion (Poynton, 2007).

$$\begin{bmatrix} Y'_{709} \\ C_B \\ C_R \end{bmatrix} = \begin{bmatrix} 1 & -0.115550 & -0.207938 \\ 0 & 1.018640 & 0.114618 \\ 0 & 0.075049 & 1.025327 \end{bmatrix} \cdot \begin{bmatrix} Y'_{601} \\ C_B \\ C_R \end{bmatrix} \quad (3.26)$$

$$\begin{bmatrix} Y'_{601} \\ C_B \\ C_R \end{bmatrix} = \begin{bmatrix} 1 & 0.099312 & 0.191700 \\ 0 & 0.989854 & -0.110653 \\ 0 & -0.072453 & 0.983398 \end{bmatrix} \cdot \begin{bmatrix} Y'_{709} \\ C_B \\ C_R \end{bmatrix} \quad (3.27)$$

In case of subsampled chroma channels, up- and down-conversion both require interpolation before, and subsampling after the color transformation. These computationally intensive calculations are necessary to avoid color reproduction errors due to different luma encoding coefficients.

3.5. Legal and Valid Signals

Every encoding system has certain limits for signal excursion. For example, analog $R'G'B'$ systems have a lower voltage limit of 0 V and an upper limit of 700 mV. Ambiguously, such coding range is often referred to as *gamut* in technical video literature, although in color science this term describes a subset of reproducible colors. If a signal remains within the allowed coding range for the particular encoding system, it is considered to be *legal*. If a conversion of a legal signal into another system, e.g. from $Y'P_BP_R$ to $R'G'B'$, still produces a legal signal, the original signal is referred to as *valid* (Schmidt, 2005). However, if the resulting signal exceeds the permitted limits for its coding system, the source signal is invalid. Valid signals are always legal, but not every legal signal is necessarily valid. Figure 3.10a shows the transformation of a legal and valid color difference signal $Y'P_BP_R$ into a legal $R'G'B'$ signal. When the legal $Y'P_BP_R$ signal is distorted, as shown in Figure 3.10b, it still remains within its coding range. However, the $R'G'B'$ signal resulting from the conversion exceeds its lower limit. In this case, the $Y'P_BP_R$ signal is legal, but invalid, and the $R'G'B'$ is illegal.

Valid $Y'P_BP_R$ signal values only occupy about 25 % of a legal $Y'P_BP_R$ space, as shown in Figure 3.11a. All the signals within the $Y'P_BP_R$ cube are $Y'P_BP_R$ legal. However, about 75 % of all $Y'P_BP_R$ signals correspond to $R'G'B'$ com-

3. Colorimetry in Video Applications

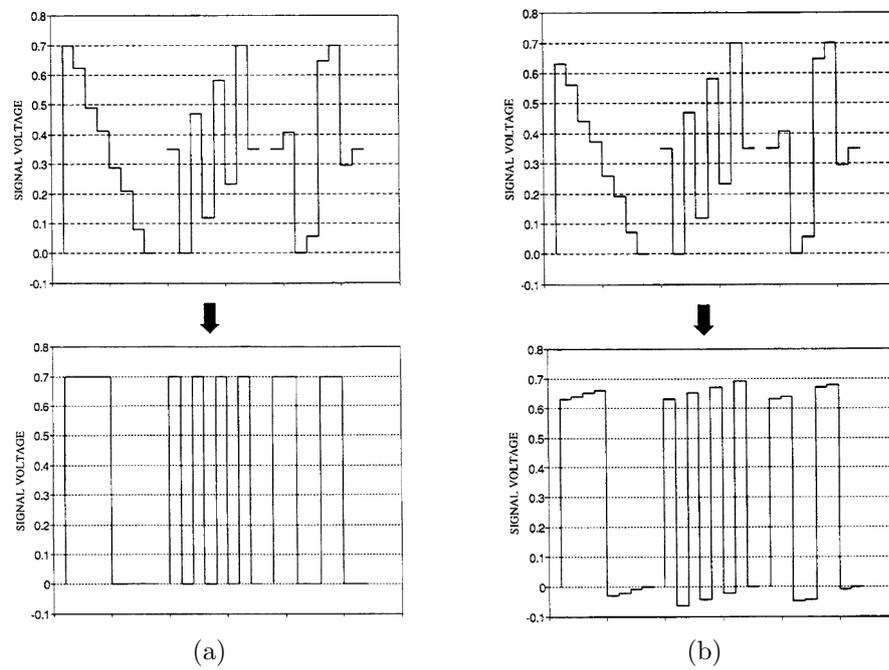
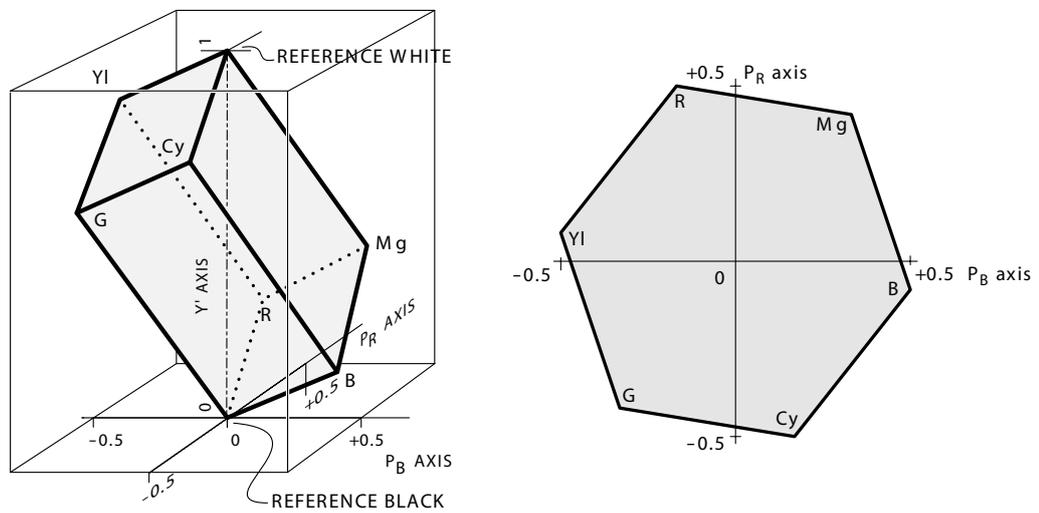


Figure 3.10.: Valid and invalid color difference component signals (Tektronix Inc., 2002).

binations outside the legal $R'G'B'$ cuboid (Poynton, 2007). Although legal in $Y'P_BP_R$ space, they are invalid and would produce illegal $R'G'B'$ signals.

3. Colorimetry in Video Applications



(a) $Y'P_BP_R$ cube and $R'G'B'$ cuboid (b) Projection of $R'G'B'$ cuboid onto P_BP_R plane

Figure 3.11.: $Y'P_BP_R$ cube. The luma excursion is normalized to 1. P_B and P_R chroma signals must remain within the range of ± 0.5 . (Poynton, 2007).

4. Software Design

4.1. Programming Language and Tools

The main challenge of this work was to develop and to implement a software interface for the objective analysis of the image quality of video capturing devices. In order to achieve this goal, an appropriate hardware interface had to be selected and its software counterpart had to be implemented. The software part had to be integrated into the IE-Analyzer by Image Engineering — a modular software solution for objective testing of imaging devices discussed in detail in Section 4.2 IE-Analyzer.

Since IE-Analyzer is written in MATLAB, it was a consequent decision to utilize this programming language for the development of the video interface, referred to as the *Video Module* hereafter. MATLAB is a platform-independent commercial environment by The MathWorks, Inc. intended for numerical computing. It provides a high-level technical programming language, also referred to as MATLAB. MATLAB is often used for rapid prototyping, algorithm development and data analysis. The range of its functions can be extended by collections of special-purpose functions, called *toolboxes*. For this work, the Image Acquisition Toolbox (IMAQ) and the Image Processing Toolbox are used in particular. IMAQ provides direct acquisition of images and video from hardware devices into MATLAB. It has native support for a wide range of capturing devices, from simple webcams and industrial frame-grabbers to high-end scientific equipment. Image Processing Toolbox provides a set of algorithms and tools for image processing and analysis. In order to run a com-

piled MATLAB program, the interpreter — MATLAB Component Runtime (MCR) — has to be installed first. Since MATLAB is an interpreted language, its code execution is sometimes slower compared to the direct machine code execution.

An appropriate hardware interface is necessary to attain video data for further analysis. Multibrige Pro by Blackmagic Design Ltd. became the device of choice, because it supports various video connections including HDMI, standard- and high-definition SDI 4:2:2, SDI 4:4:4, analog Y'P_BP_R, S-Video and NTSC/PAL video (see Figure 4.1). The connection between the Multibrige Pro and the host PC is established via a 4-lane PCI Express card allowing data transmission rates up to 10 Gbit/s.

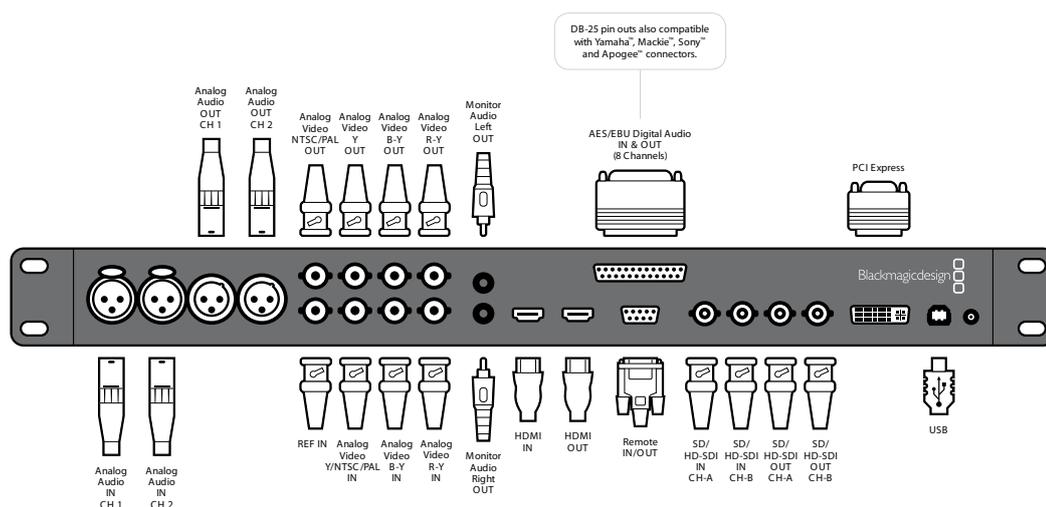


Figure 4.1.: Connection diagram for Blackmagic Multibrige Pro (Blackmagic Design Ltd., 2009).

Multibrige Pro is not natively supported by the MATLAB's IMAQ toolbox. To ensure interoperability between hard- and software, Blackmagic Design provides a Software Development Kit (SDK) for a programmatic control of its hardware products. The SDK includes a set of Application Programming Interfaces (APIs) for C++ containing both low-level control of hardware and high-level interfaces for common tasks. On the MATLAB's side, IMAQ provides an Adaptor Kit for C++, that allows software developers to write custom

adaptors for the Image Acquisition Toolbox. An adaptor is a dynamic-link library (DLL) that provides the connection between the IMAQ engine and a device driver using a hardware vendor's programming interface. The adaptor design for the Multibridge Pro is discussed in Section 4.3 DeckLink Adaptor.

4.2. IE-Analyzer

The IE-Analyzer is a stand-alone software by Image Engineering for testing digital imaging devices, such as digital still cameras or scanners. The Video Module, developed in the context of this work and discussed in detail in Section 4.4 Video Module, provides an additional interface for the video capturing devices. For the analysis purpose, dedicated test charts are used as a reference. The images of a test chart, taken under defined conditions with the device to be analyzed, are processed by the IE-Analyzer in order to quantify the characteristics of the device. Since the access to the different functional units of a capturing device is often a privilege of the manufacturer, the imaging system is generally considered as a black box in IE-Analyzer.

IE-Analyzer consists of different modules, each serving a specific function. Various ISO standards provide a basis for the analysis. The modules are briefly introduced in this section.

OECF Module. Opto-Electronic Conversion Function (OECF) describes transfer characteristic of a digital imaging device pointing out the relationship between the scene's luminance and corresponding captured digital values. Various properties of an imaging device, such as characteristic OECF curve (ISO 14524, 2009), sensitivity (ISO 12232, 2006), noise performance (ISO 15739, 2003), white balance and dynamic range (ISO 15739, 2003), can be measured in the OECF Module by means of a single grayscale test chart.

Color Module. This module is used for the measurement of the color rendition of a capturing system. A perceptual uniform color space CIE $L^*a^*b^*$, which is a transformation of CIE XYZ , allows the measurement of color distance, that is closely related to human perception of color differences. Based on one of the the CIE color distance formulas (CIE76, CIE94 or CIEDE2000), the colors from the digital image of a test target, for example X-Rite ColorChecker or ColorChecker SG, are compared to the chart-specific reference values within the CIE $L^*a^*b^*$ color space. The Color Module is able to handle different white points, chromatic adaptation and distinction between luminance, saturation and hue in the results.

Resolution Module. When the measurement of detail reproduction and sharpness is required, the Resolution Module provides information about Spatial Frequency Response (SFR) of an imaging system. Different test targets can provide a reference, since the calculation can be performed on both the images of slanted edges and Siemens stars (ISO 12233, 2009). Furthermore, the impact of some image enhancement algorithms, such as sharpening, can be determined by using a test target exhibiting a known noise distribution. Image processing algorithms often affect this distribution in the captured image.

Shading Module. The loss of intensity from the center of an image to its periphery is referred to as *shading* or *vignetting*. The natural illumination falloff is inherent to every lens, but it can be amplified by the lens geometry and chosen aperture. This effect can be evaluated on the images of uniformly illuminated homogeneous test targets, such as a diffuser plate. Irregularities in the color reproduction as well as the noise, potentially increased by shading correction algorithms, can also be measured using the Shading Module of the IE-Analyzer.

Distortion Module. Distortion is a form of optical aberration and describes a deviation from rectilinear projection. The radial distortion (due to the symmetry of the lens) typically exhibits either barrel or pincushion shape. Both

types — TV distortion (EBU Tech. 3249-E, 1995) and lens geometric distortion (I3A CPIQ, 2009) — can be calculated using a test chart with a rectilinear pattern. Expressing distortion as a function of the image height, a polynomial approximation can be performed in order to assist the development of compensating techniques. Operating on different color channels, the Distortion Module provides sub-pixel accuracy. Furthermore, lateral chromatic aberration can be determined by measuring the offset between the channels.

Histogram Module. Histograms were discussed in Section 2.3.3 Histogram in detail. In IE-Analyzer, they are used to locate defective pixels, that are too dark (dead pixels) or too bright (hot pixels). For this purpose, a uniform black, gray or white image has to be analyzed. An adjustable threshold allows users to define, which pixels should be considered defective. Their positions are then stored in a pixel map for the sake of correction and further processing.

4.3. DeckLink Adaptor

4.3.1. C++ Framework

The Image Acquisition Toolbox for MATLAB provides a C++ framework for creating own adaptors. An adaptor is a dynamic-linked library (DLL) that establishes the connection between IMAQ and a device driver. It can represent either a specific device, multiple devices from one specific manufacturer or a generic class of devices. In this work, the implemented DeckLink Adaptor represents multiple devices by Blackmagic Design, that are supported by the DeckLink SDK, particularly Multibrige Pro. The C++ framework is a set of classes following a predefined design. The development of an own adaptor is performed by implementing framework classes. The resulting layered architecture, consisting of MATLAB, DeckLink Adaptor and Multibrige Pro, is illustrated in Figure 4.2.

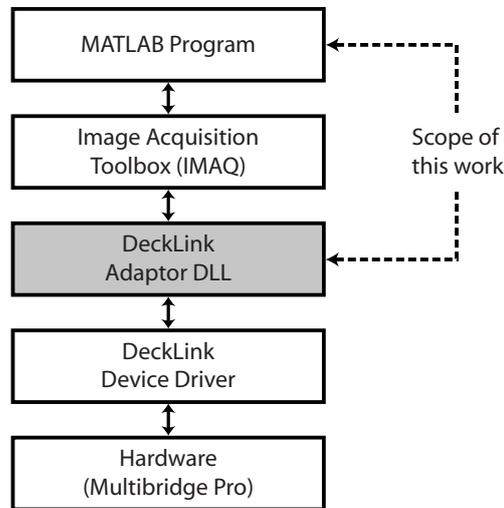


Figure 4.2.: Relationship of adaptor to toolbox components and hardware.

According to IMAQ’s adaptor class design, every adaptor must provide five functions in order to communicate with the toolbox engine. Some of them, however, can be empty, if there is no need for their implementation. These functions are `initializeAdaptor()`, `getAvailHW()`, `getDeviceAttributes()`, `createInstance()` and `uninitializeAdaptor()`.

initializeAdaptor(). This function can perform initialization required by the adaptor or the device driver, for example loading required DLLs. No hardware initialization is performed here. Since this functionality is not necessary for the DeckLink driver, the implementation of `initializeAdaptor()` only includes several null pointer initializations for safety reasons.

getAvailHW(). The implementation of this function defines the device name, its ID and video format names in order to provide the IMAQ toolbox with the information about the connected device. This information is stored in the objects of three classes as shown in Figure 4.3.

Information on each particular device, that is available through DeckLink Adaptor, is stored in a separate `IDeviceInfo` container. These containers are

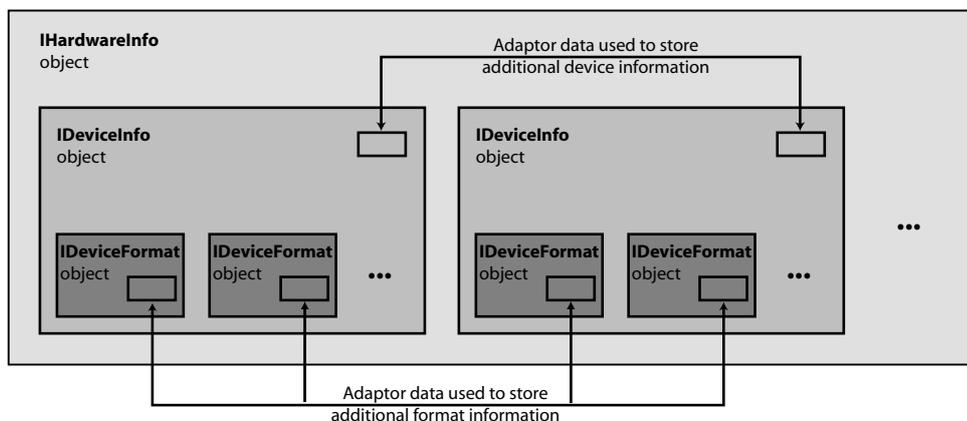


Figure 4.3.: Classes used to store device and format information.

stored in the `IHardwareInfo` object — a global container for hardware data. For each format supported by a device, an `IDeviceFormat` object holds format-specific data. It is, in turn, stored in the corresponding `IDeviceInfo` object. The information about the connected devices and their formats is attained dynamically using the functions provided by the DeckLink API. The following steps outline the algorithm for the `getAvailHW()` function as implemented in this work. Figure 4.4 illustrates this algorithm in flowchart form. Table A.3 on page 94 lists the formats, supported by the DeckLink API.

1. Determine which devices are available using DeckLink API. Exit loop if no device found.
2. For each device, create an `IDeviceInfo` object.
 - For each format supported by the device, create an `IDeviceFormat` object.
 - Add each created device format object to the `IDeviceInfo` object.
3. Add the `IDeviceInfo` object to the `IHardwareInfo` object, that is passed to the `getAvailHW()` function by the toolbox engine as a parameter.
4. Return to step 1 and repeat this procedure for each device available on the user's system.

Figure 4.4.: Flowchart of `getAvailHW()` function.

getDeviceAttributes(). This function defines video sources that are available to the user. A video source refers to one or more hardware inputs that are treated as a single entity. Each video source has a name, visible to the user, and an ID number. DeckLink API provides five hardware inputs — SDI, HDMI, Component, Composite and S-Video connections. A source listener class is associated with the `SelectedSourceName` property of a video input object in the MATLAB domain. When this property is changed in the MATLAB program, the listener in the adaptor notifies the device driver to change the input source accordingly.

createInstance(). This function is called by the toolbox in order to instantiate an object of the adaptor class. It is a subclass of the `IAdaptor` abstract class from the IMAQ C++ framework. This abstract class provides several pure virtual functions that have to be implemented in the particular adaptor

class. Some of them provide the information about the adaptor itself (e.g. `getDriverDescription()` and `getDriverVersion()`), the others refer to the properties of the current video format (e.g. `getMaxWidth()`, `getMaxHeight()`, `getNumberOfBands()` or `getFrameType()`). The functions `closeDevice()` and `openDevice()` handle the adaptor's connection to the IMAQ toolbox. `startCapture()` and `stopCapture()` begin and terminate the acquisition of the frames.

`uninitializeAdaptor()`. This function is called when the IMAQ toolbox is reset or the MATLAB program terminates. In DeckLink Adaptor, it is used to free allocated memory safely.

4.3.2. Image Acquisition

The IMAQ toolbox provides some class design rules to implement a threaded image acquisition. It should be used whenever the device driver operates in the pull-mode, so that every frame to be captured has to be requested from the driver. Blackmagic's DeckLink driver works in the push-mode, though. A listener object derived from the virtual `IDeckLinkInputCallback` class is notified every time a new frame arrives. For this reason, the design of the acquisition loop makes use of both frameworks — DeckLink API to capture frames and IMAQ's functions to send them to the MATLAB program.

As explained in the previous section, IMAQ's function `openDevice()` is called by `createInstance()` in order to establish the connection between adaptor and the toolbox engine. In this work, `openDevice()` is used to configure the video input, according to selected video format, and to instantiate and configure a listener object for arriving frames. When acquisition is requested from the MATLAB program, IMAQ's function `startCapture()` triggers the acquisition loop calling the `StartStreams()` function of the DeckLink API. Every time the frame is available to the driver interface, the DeckLink's listener object is notified and performs image data transfer to the IMAQ engine.

For this purpose, the data is packed into the `IAdapterFrame` class that is standardized by the IMAQ toolkit. In this manner it is guaranteed that the data frame can be read by the Image Acquisition toolbox, independent of the driver's native data format. Figure 4.5 illustrates the acquisition architecture.

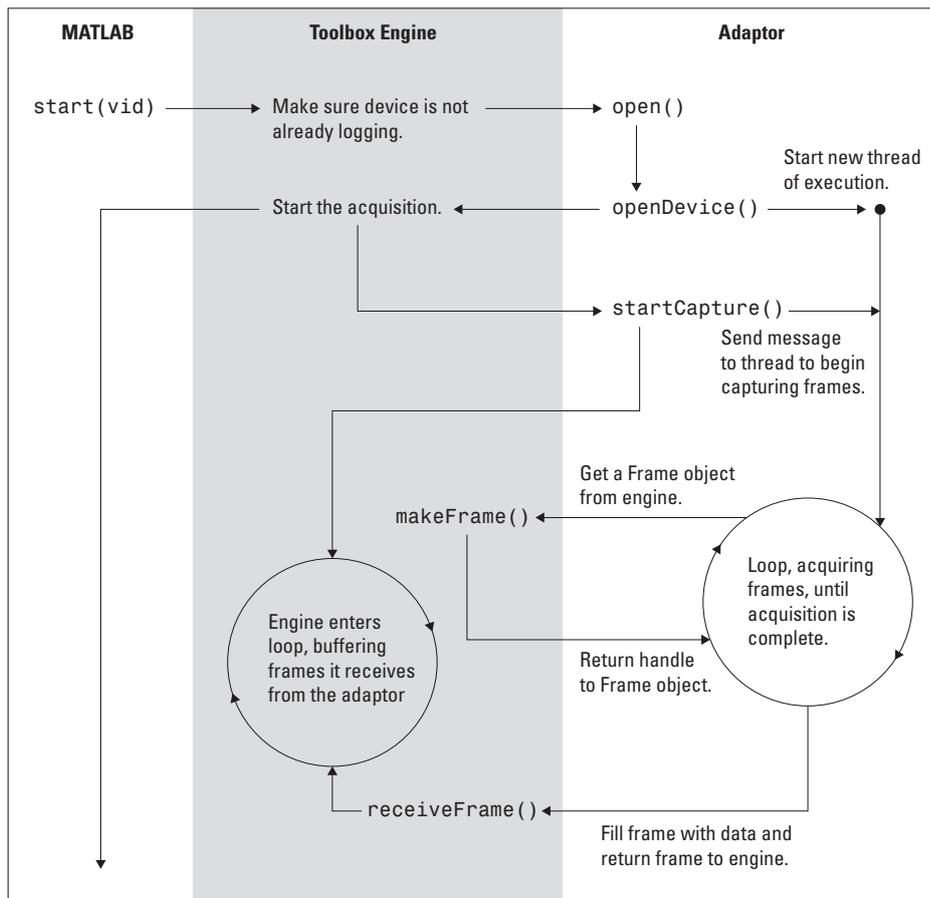


Figure 4.5.: Flow of control among the adaptor acquisition functions (The MathWorks Inc., 2009).

The frame acquisition loop grabs frames from the device driver and sends them to the IMAQ engine. This process, depicted in a simplified schematic manner in Figure 4.6, consists of the following steps:

1. Check whether the specified number of frames has been acquired using IMAQ's `isAcquisitionNotComplete()` method. Exit loop if the acquisition is complete.

2. Grab a frame from the device using DeckLink's `GetBytes()` function and put the image data into a previously allocated buffer.
3. Check whether it's needed to send the acquired frame to the IMAQ engine using IMAQ's `isSendFrame()`. This allows to reduce the frame rate. Put the frame data from the buffer into an `IAdaptorFrame` object and pass it to the IMAQ toolbox engine if the frame should be sent. Otherwise, skip to step 4.
 - Create a frame object of specified image dimensions using the `makeFrame()` method. Image data beyond the region of interest (ROI) will be discarded in the next step.
 - Put the acquired image data from the buffer into the frame object using the `setImage()` function.
 - Send the packaged frame to the IMAQ calling the engine's member function `receiveFrame()`.
4. Increment the frame counter for every frame (no matter whether sent or not) using `incrementFrameCount()`.
5. Return to step 1 and repeat.

4.4. Video Module

4.4.1. Principle

The Video Module is designed to provide an interface for the analysis of video capturing devices with IE-Analyzer. Taking advantage of MATLAB's Image Acquisition Toolbox, it supports numerous hardware devices, such as miscellaneous USB cameras, frame grabbers and FireWire (IEEE 1394) devices conform to the DCAM/IIDC standard for uncompressed video (1394 Trade

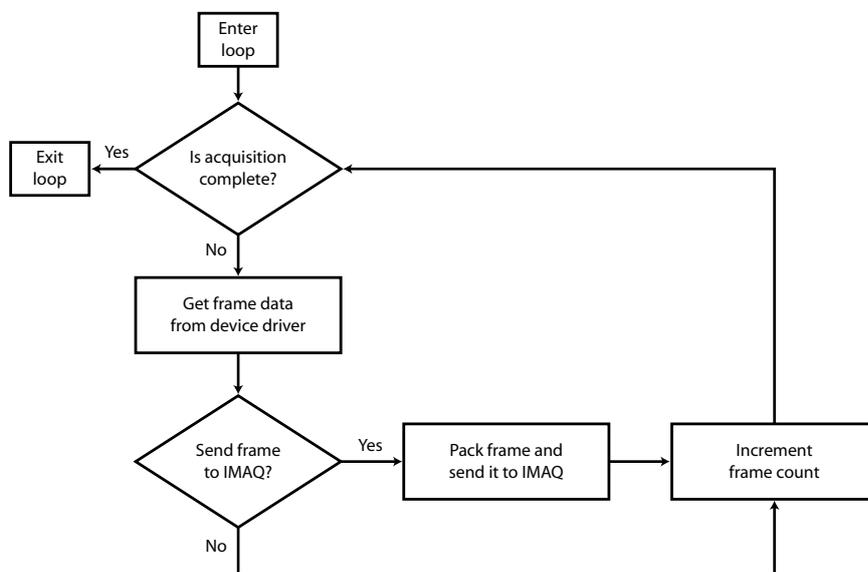


Figure 4.6.: Frame acquisition loop.

Association, 2004). Furthermore, the DeckLink Adaptor implemented in the context of this work provides support for the DeckLink product series by Blackmagic Design, in particular Multibrige Pro, as discussed in Section 4.3 Deck-Link Adaptor.

The Video Module is divided into three submodules. The *Preview* submodule allows a first visual evaluation and selection of a region of interest (ROI), a part of the image that is relevant for the measurements. The image data outside the ROI is discarded in the adaptor domain, offering processing speed-up. Signal evaluation using traditional instruments is done in the *Measurement* submodule consisting of four monitors — video preview, histogram, waveform monitor and vectorscope. Measurement of color reproduction related to dedicated test charts and corresponding reference data can be performed in the *Comparison* submodule. Moreover, a selectable number of frames can be saved. The image files can then be easily passed to any other IE-Analyzer module for further analysis. Figure B.1 on page 96 in the Appendix provides screenshots of the submodules.

4.4.2. Graphical User Interface (GUI)

In general, a module GUI of the IE-Analyzer is divided into panels with settings and panels for the image or results display. The general settings should be easily accessible, hence they are always visible. An advanced settings panel provides rarely changed options. It can be expanded on demand. Image and result panels occupy the most space in the GUI. The Video Module consistently follows this layout.

Settings. The general settings panel of the Video Module offers a selection of available hardware, including input interface and video format settings. Device specific properties, such as gain, brightness, saturation or similar, are only available if the corresponding adaptor DLL provides these options. If available, the properties dialog window can be accessed from the general settings panel. DeckLink adaptor, for instance, does not have any device related settings. An example of device specific settings dialog is illustrated in Figure B.3 on page 98 in Appendix. Trigger control for the frame grabbing, defining how many frames should be captured and at which frame rate, is situated in the panel for the general settings as well as the list containing captured frames and the button starting the acquisition. Advanced settings can be expanded for access using a toggle button in the general settings panel. It features options concerning saving acquired pictures, selection of a custom graticule layout for the vectorscope, color model selection and warning levels for the results of the Comparison submodule. Figure B.2 on page 97 in Appendix provides screenshots of the settings panels.

Preview Submodule. The Preview submodule shows the live video assisting in environmental setup and adjustment of the camera position. A region of interest can be defined by an adjustable rectangle as shown in Figure B.4 on page 99 in Appendix. The buttons “ROI/Full” and “Confirm” invoke the ROI selection and display of a full frame image.

Measurement Submodule. The Measurement submodule consists of a preview panel showing live video, a histogram monitor allowing the evaluation of the pixel value distribution as well as a waveform monitor and a vectorscope for common signal monitoring. Each of these monitors can be turned on and off separately and provides specific settings, such as signal selection or graticule options. The preview panel allows switching between the color, luma and blue channel display. The latter is helpful when adjusting monitor's contrast and brightness using a color bars test signal. A mouse click on the video selects the video line to be analyzed within the WFM and the vectorscope. Histogram and waveform monitors offer selection of the displayed signal source (Y', C_B, C_R, R', G', B') as well as buttons for parade display of either component or $R'G'B'$ signal. Vectorscope panel provides the graticule customization for color bars signals with 75% and 100% saturation as well as custom layouts that can be selected in the advanced settings panel. In addition, WFM and vectorscope panels can either visualize the data from a single video line selected in the preview panel or the whole video frame. Screenshots in Figure B.5 on page 100 in Appendix depict the measurement monitors.

Comparison Submodule. The Comparison submodule allows an evaluation of color reproduction on live video streams. For this purpose, a test chart exhibiting color patches must be used as reference, for example EBU or NTSC color bars, X-Rite ColorChecker or ColorChecker SG. The latter two provide more color samples than the color bars, featuring skin tones for instance. A file describing the chart layout as well as a file containing reference color values must be selected in order to compare captured colors and the reference. After the positions of the chart and its patches are detected in the image, a visual comparison panel allows subjective estimation displaying the mean patch color data and its reference side by side. Human visual system is a very sensitive instrument concerning color differences. However, a calibrated and profiled monitor is highly recommended for this task. An objective measurement is also possible, computing color differences using the $CIE76 \Delta E_{ab}$ formula as described in detail in Section 4.4.3 Live Color Comparison.

Moreover, the Comparison submodule allows saving the color data from the video in a reference file. It can assist in matching color reproduction of different cameras in a broadcasting studio for instance. For this purpose, one particular camera provides reference data. A user can choose how many frames from the live video stream should be averaged in order to create an own reference dataset. The other cameras are then adjusted to match the color reproduction of the reference. Screenshots in Figure B.6 on page 101 in Appendix illustrate the color comparison task.

4.4.3. Live Color Comparison

Chart Detection. The implementation of the detection of a test chart within a video image for the purpose of live color comparison was beyond the scope of this work. An already existing MATLAB function performing pattern recognition by template matching was utilized instead. Its mathematics is briefly discussed here for the sake of completeness. Locating the chart position in the image is performed by means of *normalized cross-correlation*. In signal processing, cross-correlation is a measure of similarity between two signals. In the Video Module, one of these signals is the image acquired from the video stream, the other one is a template image of a test target that is stored in a chart layout file. Figure 4.7 shows an example of an image and a template.

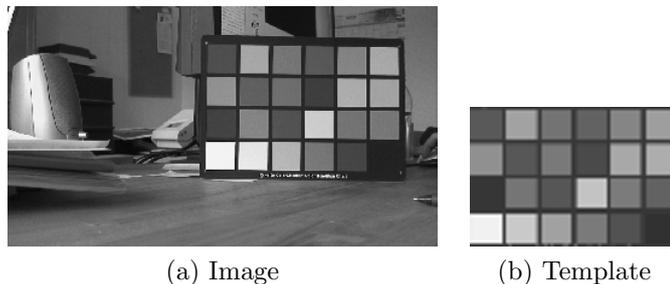


Figure 4.7.: Example of the image and template used for the chart detection. In order to reduce computing time, the processing is performed on the luma channel of a downscaled image.

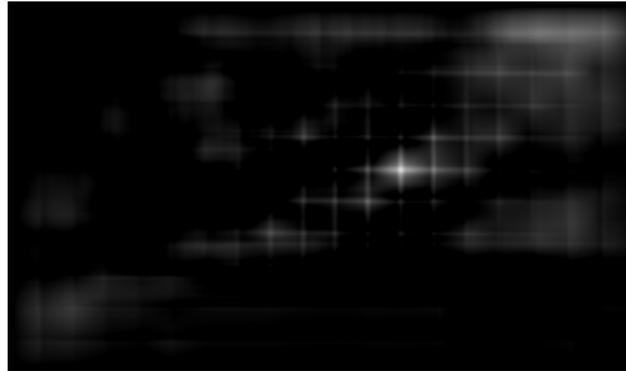
For the image $f(x, y)$ and a template $t(x, y)$ located in the image at the position (u, v) , the cross-correlation function $c(u, v)$ is given in Equation 4.1.

$$c(u, v) = \sum_{x,y} f(x, y)t(x - u, y - v) \quad (4.1)$$

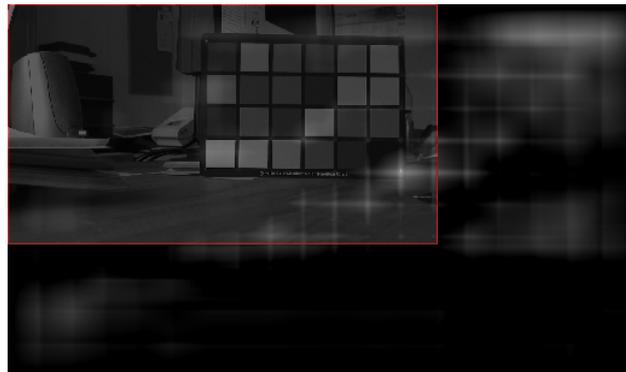
To compensate for the possible changes in image brightness which can occur due to lighting and exposure conditions, the image data can be normalized. The normalization is performed by subtracting the mean value at each pixel and dividing by the standard deviation. The normalized cross-correlation is defined as stated in Equation 4.2, where the mean value of a signal $s(x, y)$ is denoted \bar{s} and the standard deviation σ_s .

$$c_n(u, v) = \frac{\sum_{x,y} [f(x, y) - \bar{f}] [t(x - u, y - v) - \bar{t}]}{\sigma_f \sigma_t} \quad (4.2)$$

MATLAB provides the function `normxcorr2` in order to calculate the cross-correlation coefficients. More information on the normalized cross-correlation is provided in the work of Lewis (1995). The results of the calculation are illustrated in Figure 4.8. The maximum of the coefficients marks the lower right corner of the test target looked for. The dimensions of the coefficients plane are greater than those of the original image to allow for the sliding the template over the image in terms of the calculation. After the position of the test chart is found, the patch locations are distributed relative to the chart coordinates and according to the layout defined in the chart layout file. The calculation of the color differences discussed next is performed by comparing the mean data from the patch regions in the video and the reference data provided in a chart reference file. Thus, two auxiliary files are required in order to perform color comparison — one defining the chart layout and one containing the reference color data.



(a)



(b)

Figure 4.8.: Results of a normalized cross-correlation function for the image and template in Figure 4.7. Picture (a) visualizes the absolute values of the coefficients of the normalized cross-correlation function. The superimposed original image in (b) is used to emphasize the effect.

Color Differences. The colors in the *CIE XYZ* color space are not distributed uniformly referring to human visual perception. In Figure 4.9, each ellipse, referred to as *MacAdam ellipse*, marks the region on a chromaticity diagram which contains all colors that are perceived visually equal to the color at the center of the ellipse. To overcome this issue, other color spaces — *CIE L*a*b** and *CIE L*u*v** — were introduced in terms of transformations from the *CIE XYZ* space. In this work, *CIE L*a*b** color space is used for the computation of color differences.

*“These spaces [*CIE L*a*b** and *CIE L*u*v**] are intended to apply to comparisons of differences between object colours of the same size and shape, viewed in identical white to middle-grey surroundings, by an observer photopically adapted to a field of chromaticity not too different from that of average daylight.”* (CIE Publ. 15.2, 1986)

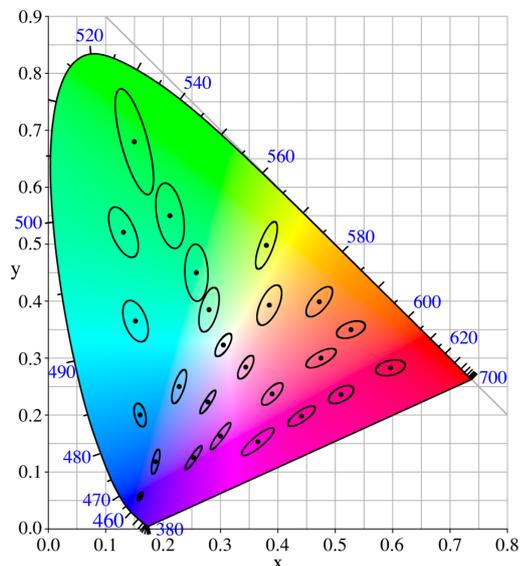


Figure 4.9.: MacAdam ellipses. The ellipses are ten times their actual size in order to emphasize the proportions (Wikipedia, 2010).

The *CIE L*a*b** introduces an opponent-color encoding. The L^* axis represents the lightness. The a^* axis corresponds to the red-green opponent hues, while the b^* axis corresponds to the yellow-blue ones. Positions along the pos-

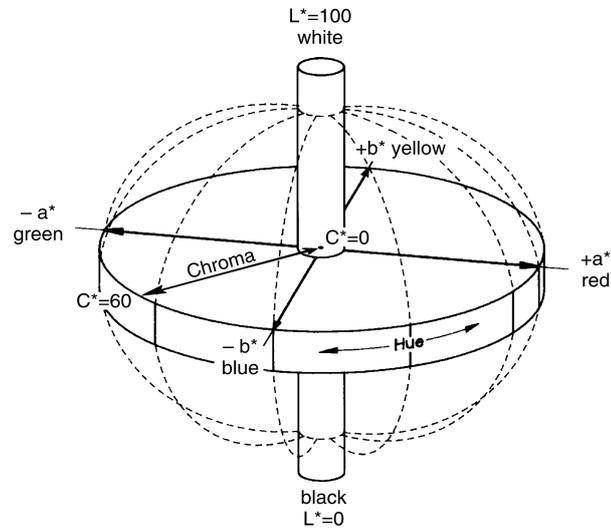


Figure 4.10.: Schematic representation of the *CIE L*a*b** color space (X-Rite Inc., 2007).

itive a^* axis represent the measure of redness and positions along the negative a^* axis correspond to the measure of greenness. Likewise, the positive b^* axis represents the measure of yellowness and the negative b^* axis stands for blueness (cf. Figure 4.10). The transformation from *CIE XYZ* to *CIE L*a*b** is described in Equation 4.3 (CIE Publ. 15.2, 1986).

$$\begin{aligned}
 L^* &= 116f(Y/Y_n) - 16 \\
 a^* &= 500[f(X/X_n) - f(Y/Y_n)] \\
 b^* &= 200[f(Y/Y_n) - f(Z/Z_n)]
 \end{aligned}
 \tag{4.3}$$

The values X_n, Y_n and Z_n are the tristimulus values of the reference white. The function $f(\cdot)$ of the quotients is piecewise defined as shown in Equation 4.4 (CIE Publ. 15.2, 1986).

$$f(Q) = \begin{cases} Q^{1/3} & Q > 0.008856 \\ 7.787Q + 16/116 & \text{otherwise} \end{cases}$$

$$Q \in \{X/X_n, Y/Y_n, Z/Z_n\} \quad (4.4)$$

In the approximately perceptually uniform color space $CIE L^*a^*b^*$, the color difference $CIE76 \Delta E_{ab}$ is defined as the Euclidean distance between the reference color (L_r^*, a_r^*, b_r^*) and the sample color (L_s^*, a_s^*, b_s^*) as shown in Equation 4.5.

$$CIE76 \Delta E_{ab} = \sqrt{(L_r^* - L_s^*)^2 + (a_r^* - a_s^*)^2 + (b_r^* - b_s^*)^2} \quad (4.5)$$

Transforming $CIE L^*a^*b^*$ color space from Cartesian into cylindrical coordinates as described in Equation 4.6, the color is given by its lightness L^* , chrominance (saturation) C^* and hue-angle h° . This representation is sometimes more convenient when discussing colors since it resembles the way humans percept colors.

$$C^* = \sqrt{a^{*2} + b^{*2}}$$

$$h^\circ = \arctan\left(\frac{b^*}{a^*}\right) \quad (4.6)$$

The Comparison submodule features the calculation of both ΔE_{ab} and color differences in $L^*C^*h^\circ$ color space representation. Since ΔE_{ab} only indicates the overall color differences, differences in lightness, saturation and hue provide information that is more distinct. The difference in the hue-angle h° can be large for desaturated colors lying close to the achromatic axis, although they appear similar. In order to exhibit the same difference in the hue-angle, two saturated colors must be distant to each other. To overcome this problem,

CIE defines the hue-difference ΔH_{ab}^* eliminating the impact of saturation. The hue-difference is introduced so that the overall color difference ΔE_{ab} can be split into components ΔL^* , ΔC^* and ΔH^* with $\Delta E_{ab}^2 = \Delta L^{*2} + \Delta C^{*2} + \Delta H^{*2}$, while hue-angle h° does not have this property (CIE Publ. 15.2, 1986). The computation of lightness, saturation and hue differences is shown in Equation 4.7.

$$\begin{aligned}\Delta L^* &= L_r^* - L_s^* \\ \Delta C_{ab}^* &= C_r^* - C_s^* \\ \Delta H_{ab}^* &= \sqrt{(\Delta E_{ab})^2 - (\Delta L^*)^2 - (\Delta C_{ab})^2}\end{aligned}\tag{4.7}$$

In the Video Module, video data is encoded in the *sRGB* color space due to the design of the Image Acquisition Toolbox. In order to calculate color differences, the conversion is performed first scaling *sRGB* data to the range [0..1], transforming it to the *CIE XYZ* space using inverse matrix to the one stated in Equation 3.9 on page 41 and then calculating *CIE L*a*b** values. A pseudo-color representation, encoding difference values from green (low difference) over yellow to red (high difference), assists in evaluating the results at a glance (cf. Figure B.6 on page 101). The warning levels for the maximum difference value can be defined individually in the advanced settings as shown in Figure B.2 in Appendix.

In order to compensate for the rest of perceptual nonuniformities in the *CIE L*a*b** color space, the CIE has refined the definition of the color difference formula over the years. After the previously covered CIE76 formula introduced in 1976, two other formulas — CIE94 and CIEDE2000 — were defined in 1994, respectively 2000. While still computed in the *CIE L*a*b** color space, they feature application-specific weights (CIE94) and various compensation factors as well as a hue rotation term (CIEDE2000). It is possible to calculate color differences using these formulas in IE-Analyzer's Color Module. In order to process video data, it has to be captured first using the Video Module and subsequently passed for calculation to the Color Module. More information on CIE94 formula can be found in the paper by McDonald and

Smith (1995). The work of Sharma et al. (2005) provides further details on the CIEDE2000 color difference calculation.

4.4.4. Integration into IE-Analyzer

Class Design. In order to implement the functions of the Video Module, the object-oriented programming approach was used. Each monitor in the GUI is represented by an object of a class that describes monitor's characteristics as *class properties* and its behavior by means of *class methods*. Since multiple classes can exhibit the same properties, such as a reference to the axes the monitors are painted on, and have the same methods, such as functions for selecting ROIs in the preview monitors, the concept of inheritance was utilized. For example, `videostreamMonitor` and `comparisonMonitor` classes both inherit the method `updateImageSize()` from the class `previewMonitor`. This function is implemented once in `previewMonitor` and can be used by any derived class. The parent class `previewMonitor`, in turn, inherits from the class `monitor` that declares the reference (*handle*) to the GUI area (*axes*), containing the monitors. Figure 4.11 illustrates the class inheritance hierarchy by means of a UML diagram. Appendix C lists the implemented classes with their properties and methods.

Abstract methods are only declared in the parent classes and must be implemented in the inheriting classes. For example, all classes derived from `previewMonitor` implement the functions `play()` and `stop()` in their own manner. The implementation is hidden from the software developer using these classes, the only thing he/she needs to know, is how to call these methods (e.g. `videostreamMonitor.play`) and what kind of behavior to expect. Moreover, properties and methods only used for internal purpose are declared as *private* and hence are invisible for the class user. This concept is known as *encapsulation*.

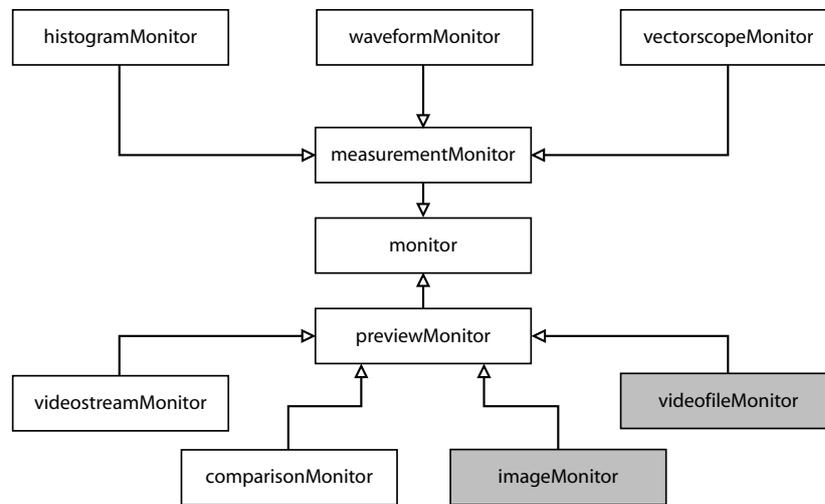


Figure 4.11.: UML class hierarchy of the Video Module. Grey shaded classes are not implemented in this work.

Module Initialization. Before the Video Module can start, the initialization of the hardware and GUI elements as well as instantiation of the monitor classes have to be accomplished. This is necessary, because the hardware is detected automatically at every start of the Video Module. Following steps describe this procedure:

1. Initialize hardware:
 - Register adaptor DLLs in order to communicate with the hardware.
 - Find connected devices supported by the registered adaptors.
 - Retrieve device information and supported video formats for each found device.
2. Populate GUI elements, particularly popup menus, with the attained hardware data (names, interfaces, video formats).
3. Create `videoinput` object representing the currently selected combination of the hardware, interface and video format.

4. Restore the state of the remaining GUI elements from a file if necessary. Otherwise set default values.
5. Create video monitor instances according to the `videoinput` object by calling the constructor methods of monitor classes.
6. Associate callback functions with the monitors. Callback routines are invoked automatically whenever the specified event occurs.

Callback Functions. As mentioned in the previous paragraph, the callback functions are invoked due to a fulfilled condition. The callback function associated with the object of the `videostreamMonitor`, for example, is called every time the new video frame is available. It checks whether the preview and measurement monitors (histogram, vectorscope and waveform monitor) are turned on. If so, it performs data preprocessing and calls corresponding member functions of the objects of currently active monitors passing on the required data.

The callback routine associated with the `comparisonMonitor` is also invoked on every incoming frame. Depending on the current selection, it displays the video data or calls the methods of `comparisonMonitor` in order to create and display the side-by-side color comparison image or to calculate color differences using reference data. The results returned by the object's functions are postprocessed in the callback method and displayed in the GUI.

When the user triggers the beginning of the frame acquisition by pressing the "Capture" button, a callback function associated with the `videoinput` object is invoked. This routine is responsible for the proper naming and saving of the image files onto the hard disk according to the current settings.

5. Results

5.1. Test conditions

This chapter illustrates the measurements that have become possible as a result of this work. The algorithms and their implementations for the most measurements discussed here are not a part of this thesis. Nevertheless, using IE-Analyzer and the developed Video Module allows to perform the measurements in a convenient way. Since the Video Module only monitors the video signals, numerical calculations are done within other modules and the results are replenished by screenshots of the Video Module when possible. This test does not include the evaluation of all possible characteristics, it merely illustrates some of the applicable software features.

The test object was a Sony HDR-HC5E camcorder. It has a single 1/3 inch CMOS sensor and a Carl Zeiss Vario-Sonnar T* 1.8/5.1 – 51 mm lens. The images have been captured in HD format (8 bit, 1920×1080 pixels, sRGB) via HDMI interface. To avoid color casts and to evaluate the auto balancing performance, the white balance was set to the automatic mode. The sharpening level was at the default value selected by the manufacturer. In order to acquire sharp images, the autofocus was activated during the tests. The used IE-Analyzer version was 4.5 (Build 072).

The measurement conditions for the transparent and reflective test charts were set up as illustrated in Figure 5.1. Ten frames with the rate of one frame per second were saved as uncompressed TIFF files and analyzed in dedicated IE-Analyzer modules for each measurement. The results presented here con-

sist of the average data of the ten frames. The output files containing the numerical results can be found on the attached CD.

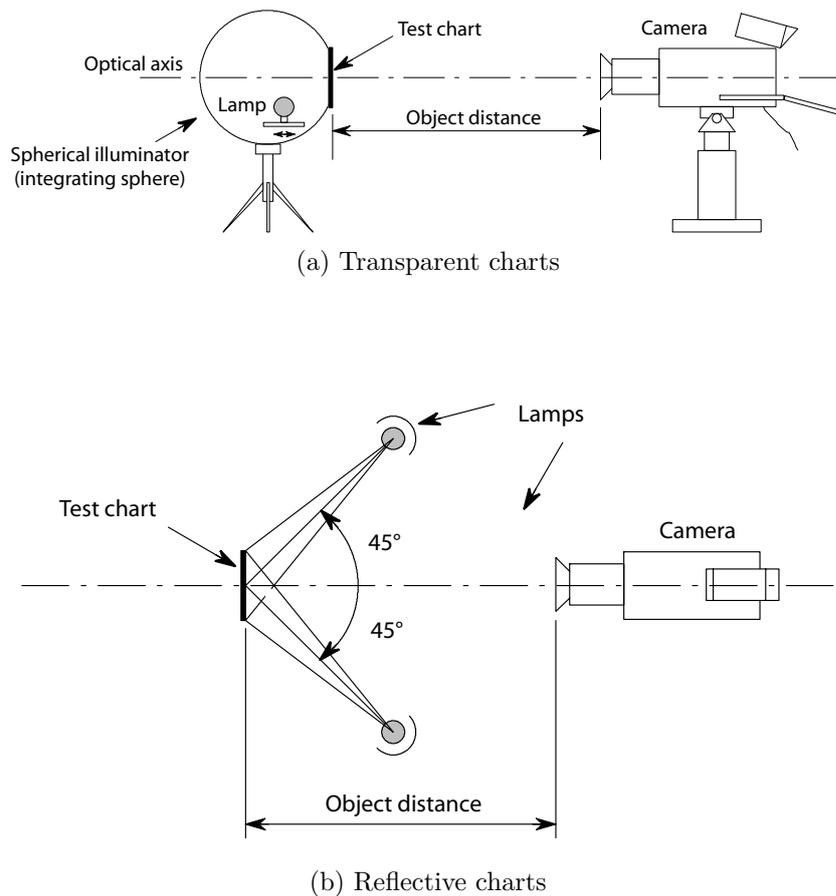


Figure 5.1.: Arrangement and illumination of test charts according to EBU Tech. 3281-E (1995).

5.2. OECF and Noise

Opto-electronic conversion function (OECF) describes the capability of a camera to transfer luminance in a scene into digital values in the image. OECF can be calculated with the aid of an arranged gray scale with known characteristics. The utilized transparent test chart exhibits contrast range of 10,000:1. The

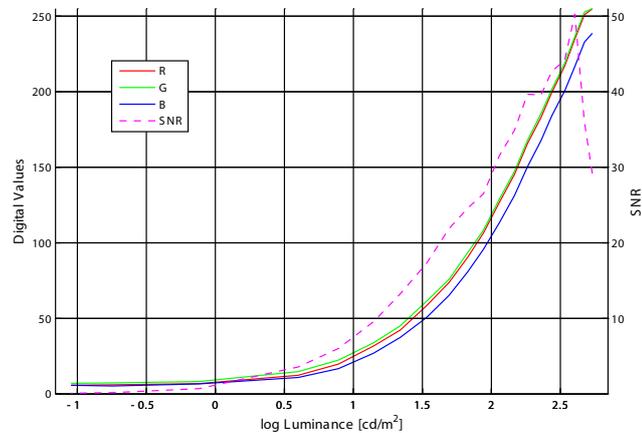
5. Results

OECF measurement procedure and the chart layout are described in ISO 14524 (2009) and ISO 15739 (2003). Signal-to-noise ratio (SNR) is also calculated as the ratio of the signal value to its standard deviation for each patch and is plotted in the results diagram. The luminance values of the patches used for the calculation have been obtained using Gossen Mavo-Monitor luminance meter. Figure 5.2 shows the used test target and the results.



(a) Example of a captured frame

(b) WFM display of the central part



(c) Average OECF and SNR

Figure 5.2.: OECF and SNR measurement results for Sony HDR-HC5E.

The blue channel's curve exhibits an undershoot compared to the red and green ones. It indicates the insufficient performance of the white balancing. The slight yellowish color cast can be seen when examining the captured images.

More characteristics of an imaging device, such as dynamic range and visual noise, can be evaluated in the OECF Module. In contrast to SNR, the visual noise is evaluated as an output referred noise. Its calculation takes into account that the spatial distribution of noise could be irregular and considers that human observers react differently to color intensity noise. More information on visual noise can be found in the paper of Kleinmann and Wueller (2007).

5.3. Resolution

The measurement of the resolution is performed on images of a reflective test target with nine sine modulated Siemens stars distributed over the image (Figure 5.3). The measurement procedure is described in ISO 12233 (2009) (Committee Draft Rev.1). The resulting modulation transfer function (MTF) describes the response of a camera as a function of spatial frequency. MTF values represent the relative decrease of modulation. The measurement mathematics are described in the work of Loebich et al. (2007).

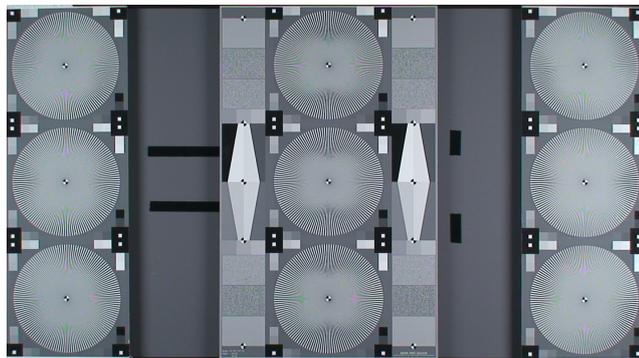


Figure 5.3.: Example of a captured frame.

The overview diagram (Figure 5.4) illustrates the resolution performance at a glance. The black circles represent the Nyquist frequency, the white areas the limiting resolution of each of the eight star segments at 10% of the modulation. The percentage value gives the ratio between Nyquist frequency and the average limiting resolution of all segments. The resolution on the left

5. Results

and right is lower than on the top and bottom, because the stars 3 and 7 are closer to the optical axis.

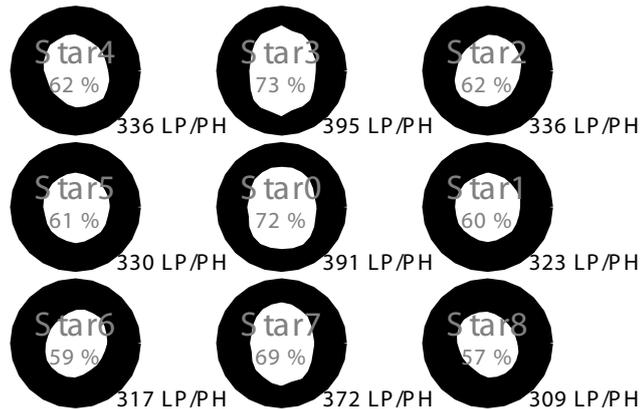


Figure 5.4.: MTF overview.

Figure 5.5 shows the MTF curves for the nine Siemens stars. The spatial resolution in the diagram is given in line pairs per picture height (LP/PH). The curves particularly exhibit values greater than one. The reason is most likely the default sharpening within the camera's signal processing.

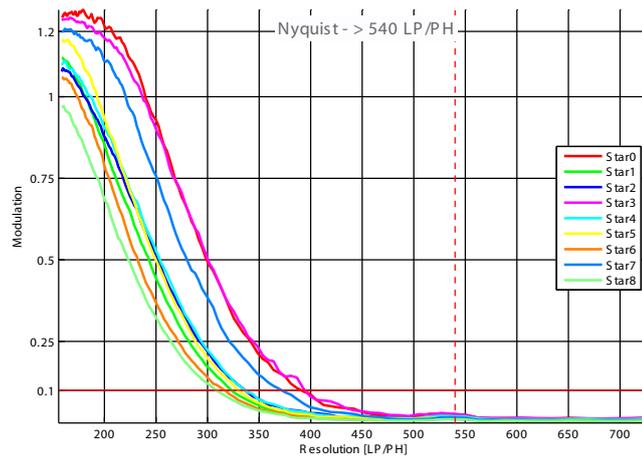


Figure 5.5.: Average MTF results.

5.4. Color

The calculation of the color differences between the captured image and the reference data is performed as discussed in Section 4.4.3 Live Color Comparison. The used reflective test chart, X-Rite ColorChecker SG, is illustrated in Figure 5.6. It exhibits 140 color patches, such as skin tones or highly saturated colors. The color reference data is provided by the chart manufacturer.

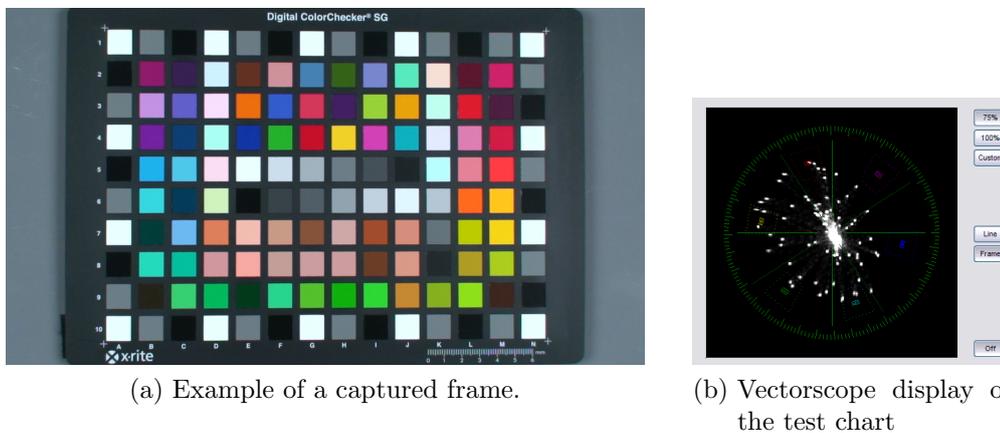


Figure 5.6.: X-Rite ColorChecker SG.

A first visual evaluation of camera's capability of color reproduction can be done comparing image and reference data side by side. A calibrated monitor has to be used in order to judge the differences visually. The comparison image is shown in Figure 5.7.

More sophisticated results can be obtained calculating numerical color difference values ΔE , ΔL , ΔC and ΔH . The comparison of the live measurements in the Video Module and the results from the IE-Analyzer's Color Module is shown in Figure 5.8. The numerical results are displayed in pseudo-colors, low color differences are painted green and high differences ($\Delta E_{ab} > 20$) dark red. The tested camera has poor color rendition performance in greens and blues as can be seen in Figure 5.8. The visual comparison (Figure 5.7) verifies this conclusion.

5. Results

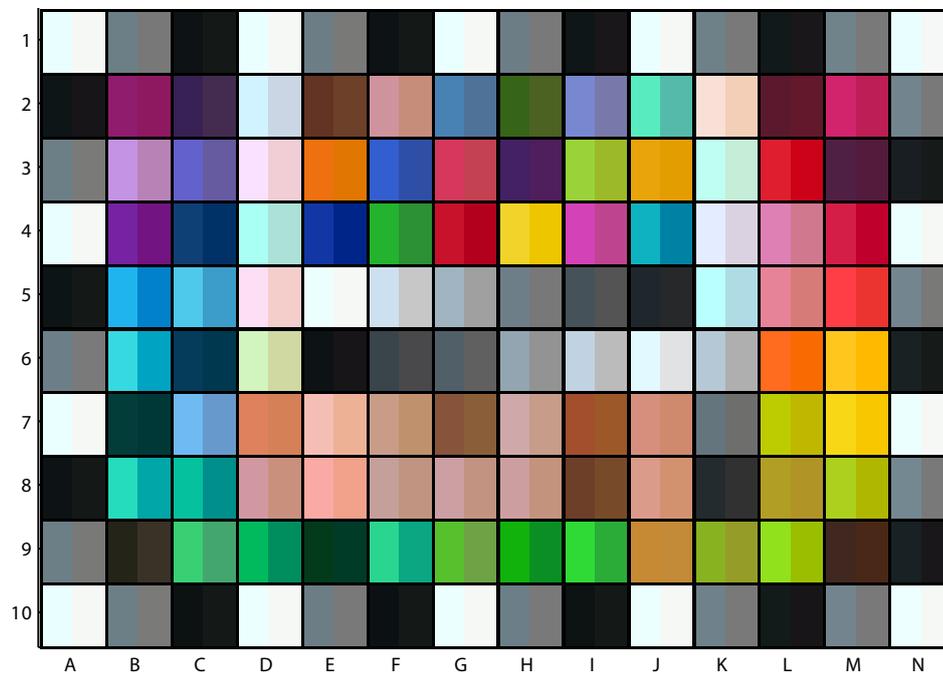
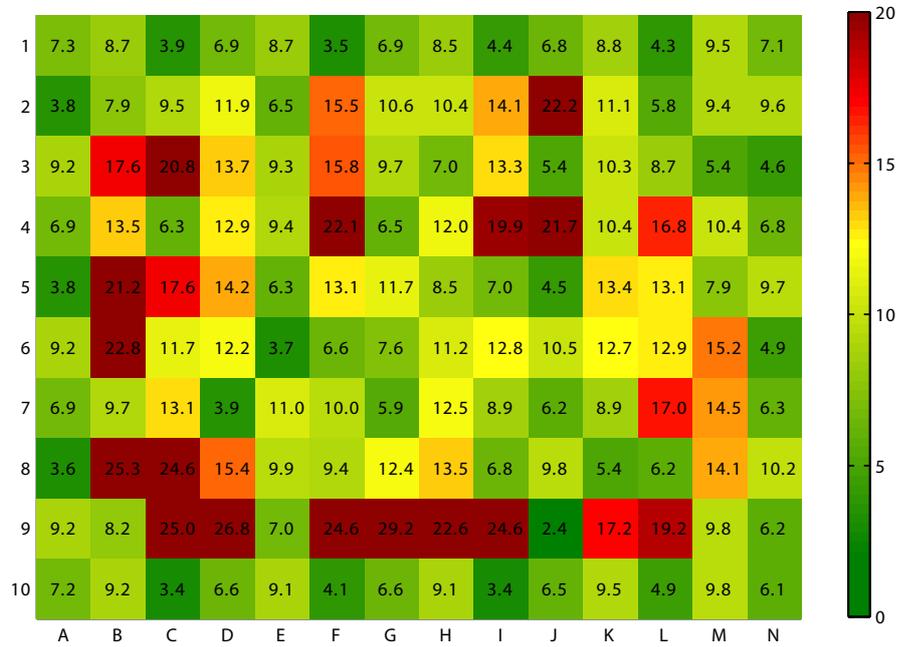
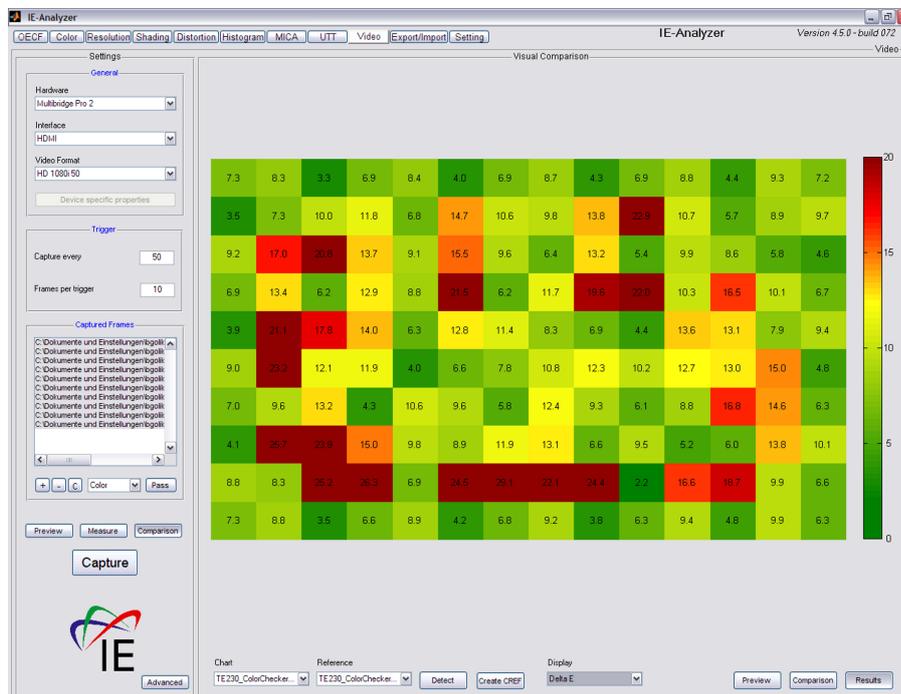


Figure 5.7.: Visual side by side comparison of the image data (left) and the reference values.

5. Results



(a) Average ΔE_{ab} results calculated in the Color Module



(b) Screenshot of the live ΔE_{ab} results from the Video Module

Figure 5.8.: Color difference measurement results for Sony HDR-HC5E.

6. Conclusion and Outlook

The aim of this thesis was to develop and implement a software interface between a video input hardware and the IE-Analyzer — the software made by Image Engineering for the purpose of standard-compliant image quality analysis of digital imaging devices.

It has been shown that the software-based objective analysis of the digital video image quality is possible and provides high potential of giving complex and reliable information about a video camera.

This work provides a foundation for further implementation of objective video quality assessment methods. Further steps may include the support of video files as signal source, evaluation of compression artifacts or implementation of standard-compliant measurement techniques, such as OECF or resolution measurements, performing on live video streams.

Bibliography

- [1394 Trade Association 2004] 1394 TRADE ASSOCIATION: *TA Document 2003017: IIDC 1394-based Digital Camera Specification (Ver.1.31)*. 2004
- [Adobe Systems Inc. 2005] ADOBE SYSTEMS INC.: *Adobe RGB (1998) Color Image Encoding*. [Online]: <http://www.adobe.com/digitalimag/pdfs/AdobeRGB1998.pdf> (visited on January 22, 2010). 2005
- [Berns 2000] BERNS, Roy S.: *Billmeyer and Saltzman's Principles of Color Technology*. 3rd Edition. John Wiley & Sons Inc., 2000
- [Blackmagic Design Ltd. 2009] BLACKMAGIC DESIGN LTD.: *Multibridge Pro & Multibridge Eclipse Operation Manual*. 2009
- [Bovik 2000] BOVIK, Alan C. (Editor): *Handbook of Image and Video Processing*. Academic Press, 2000
- [Burger and Burge 2006] BURGER, Wilhelm ; BURGE, Mark J.: *Digitale Bildverarbeitung: Eine Einführung mit Java und ImageJ*. 2nd Edition. Springer Verlag, 2006
- [CIE Publ. 15.2 1986] COMMISSION INTERNATIONALE DE L'ECLAIRAGE (CIE): *Colorimetry. Second Edition. Publication No. 15.2*. 1986
- [CIE S005 1999] COMMISSION INTERNATIONALE DE L'ECLAIRAGE (CIE): *CIE Standard Illuminants for Colorimetry*. [Online]: <http://www.cie.co.at/publ/abst/s005.html> (visited on January 21, 2010). 1999

- [EBU Tech. 3213-E 1975] EUROPEAN BROADCASTING UNION (EBU):
Tech. 3213-E: Standard for Chromaticity Tolerances for Studio Monitors.
1975
- [EBU Tech. 3249-E 1995] EUROPEAN BROADCASTING UNION (EBU):
*Tech. 3249-E: Measurement and Analysis of the Performance of Film and
Television Camera Lenses.* 1995
- [EBU Tech. 3281-E 1995] EUROPEAN BROADCASTING UNION (EBU):
*Tech. 3281-E: Methods for the Measurement of the Characteristics of CCD
Cameras.* 1995
- [Eichler et al. 1993] EICHLER, Hans-Joachim ; FLEISCHER, Axel ; KROSS,
Jürgen ; KRISTEK, Michael ; LANG, Heinwig ; NIEDRIG, Heinz ; RAUCH,
Helmut ; SCHMAHL, Günter ; SCHOENEBECK, Heinz ; SEDLMAYR, Erwin ;
WEBER, Horst ; WEBER, Kurt ; NIEDRIG, Heinz (Editor): *Bergmann
Schaefer: Lehrbuch der Experimentalphysik.* Volume 3 Optik. 9th Edition.
Walter de Gruyter, 1993
- [EN 61146-1 1997] COMITÉ EUROPÉEN DE NORMALISATION ÉLEC-
TROTÉCHNIQUE (CENELEC): *Methods of measurement for video cameras
(PAL/SECAM/NTSC) - Part 1: Non-broadcast single-sensor cameras.* 1997
- [EN 61146-2 1998] COMITÉ EUROPÉEN DE NORMALISATION ÉLEC-
TROTÉCHNIQUE (CENELEC): *Video cameras (PAL/SECAM/NTSC) –
Methods of measurement – Part 2: Two- and three-sensor professional cam-
eras.* 1998
- [EN 61146-3 1998] COMITÉ EUROPÉEN DE NORMALISATION ÉLEC-
TROTÉCHNIQUE (CENELEC): *Video cameras (PAL/SECAM/NTSC) –
Methods of measurement – Part 3: Non-broadcast camera-recorders.* 1998
- [EN 61146-4 1999] COMITÉ EUROPÉEN DE NORMALISATION ÉLEC-
TROTÉCHNIQUE (CENELEC): *Video cameras (PAL/SECAM/NTSC) –
Methods of measurement – Part 4: Automatic functions of video cameras
and camera-recorders.* 1999

Bibliography

- [Fairchild 1998] FAIRCHILD, Mark D.: *Color Appearance Models*. Addison Wesley Longman Inc., 1998
- [Ford and Roberts 1998] FORD, Adrian ; ROBERTS, Alan: *Colour Space Conversions*. [Online]: <http://www.poynton.com/PDFs/coloureq.pdf> (visited on January 22, 2010). 1998
- [Harris Corporation 2005] HARRIS CORPORATION: *Application Note 52: Digital Gamut Iris Display*. [Online]: <http://www.broadcast.harris.com/applicationnotes/testmeasurement/appnotes/app52.pdf> (visited on January 10, 2010). 2005
- [Hunt 2004] HUNT, Robert W. G.: *The Reproduction of Colour*. 6th Edition. John Wiley & Sons Ltd., 2004
- [I3A CPIQ 2009] INTERNATIONAL IMAGING INDUSTRY ASSOCIATION (I3A): *Camera Phone Image Quality (CPIQ) – Phase 2: Lens Geometric Distortion (LGD)*. 2009
- [ISO 12232 2006] INTERNATIONAL ORGANIZATION FOR STANDARDIZATION (ISO): *Photography – Digital still cameras – Determination of exposure index, ISO speed ratings, standard output sensitivity, and recommended exposure index*. 2006
- [ISO 12233 2009] INTERNATIONAL ORGANIZATION FOR STANDARDIZATION (ISO): *Photography – Electronic still-picture cameras – Resolution measurements (Committee Draft Rev.1)*. 2009
- [ISO 14524 2009] INTERNATIONAL ORGANIZATION FOR STANDARDIZATION (ISO): *Photography – Electronic still-picture cameras – Methods for measuring opto-electronic conversion functions (OECFs)*. 2009
- [ISO 15739 2003] INTERNATIONAL ORGANIZATION FOR STANDARDIZATION (ISO): *Photography – Electronic still-picture imaging – Noise measurements*. 2003

- [ITU-R BT.500 2002] INTERNATIONAL TELECOMMUNICATION UNION (ITU): *Methodology for the subjective assessment of the quality of television pictures*. 2002
- [ITU-R BT.601 2007] INTERNATIONAL TELECOMMUNICATION UNION (ITU): *Studio encoding parameters of digital television for standard 4:3 and wide-screen 16:9 aspect ratios*. 2007
- [ITU-R BT.709 2002] INTERNATIONAL TELECOMMUNICATION UNION (ITU): *Parameter values for the HDTV standards for production and international programme exchange*. 2002
- [Jack 2005] JACK, Keith: *Video Demystified: A Handbook for the Digital Engineer*. 4th Edition. Elsevier Inc., 2005
- [Jack and Tsatsulin 2002] JACK, Keith ; TSATSULIN, Vladimir: *Dictionary of Video and Television Technology*. Elsevier Science, 2002
- [Kleinmann and Wueller 2007] KLEINMANN, Johanna ; WUELLER, Dietmar: Investigation of Two Methods to Quantify Noise in Digital Images Based on the Perception of the Human Eye. In: *Proceedings SPIE – IS&T Electronic Imaging: Image Quality and System Performance IV* Volume 6494, 2007
- [Lewis 1995] LEWIS, J. P.: Fast Normalized Cross-Correlation. In: *Vision Interface*, Canadian Image Processing and Pattern Recognition Society, 1995, p. 120–123
- [Lindbloom 2007] LINDBLOOM, Bruce J.: *Useful Color Information: Using the Chromaticity Diagram for Color Gamut Evaluation*. [Online]: <http://www.brucelindbloom.com/index.html?ChromaticityGamuts.html> (visited on January 11, 2010). 2007
- [Loebich et al. 2007] LOEBICH, Christian ; WUELLER, Dietmar ; KLINGEN, Bruno ; JAEGER, Anke: Digital Camera Resolution Measurement Using Sinusoidal Siemens Stars. In: *Proceedings SPIE – IS&T Electronic Imaging: Digital Photography III* Volume 6502, 2007

Bibliography

- [McDonald and Smith 1995] MCDONALD, R. ; SMITH, K. J.: CIE94 – A New Colour-Difference Formula. In: *Journal of the Society of Dyers and Colourists* 111 (1995), No. 12, p. 376–379
- [Pascale 2003] PASCALE, Danny: *A Review of RGB Color Spaces: From xyY to $R'G'B'$* . [Online]: <http://www.babelcolor.com/download/A%20review%20of%20RGB%20color%20spaces.pdf> (visited on January 22, 2010). 2003
- [Poynton 2007] POYNTON, Charles: *Digital Video and HDTV: Algorithms and Interfaces*. Morgan Kaufmann Publishers, 2007
- [Schmidt 2005] SCHMIDT, Ulrich: *Professionelle Videotechnik*. 4th Edition. Springer Verlag, 2005
- [Sharma et al. 2005] SHARMA, Gaurav ; WU, Wencheng ; DALAL, Edul N.: The CIEDE2000 Color-Difference Formula: Implementation Notes, Supplementary Test Data, and Mathematical Observations. In: *Color Research & Application* 30 (2005), No. 1, p. 21–30
- [Spaulding et al. 2000] SPAULDING, Kevin E. ; GIORGIANNI, Edward ; WOOLFE, Geoffrey J.: Reference Input/Output Medium Metric RGB Color Encoding (RIMM/ROMM RGB). In: *Proceedings IS&T/PICS, 2000*
- [Süsstrunk et al. 1999] SÜSTRUNK, Sabine ; BUCKLEY, Robert ; SWEN, Steve: Standard RGB Color Spaces. In: *Proceedings IS&T/SID 7th Color Imaging Conference* Volume 7, 1999
- [Steinberg 1997] STEINBERG, Victor: *Video Standards: Signals, Formats and Interfaces*. Snell & Wilcox Ltd., 1997
- [Stokes et al. 1996] STOKES, Michael ; ANDERSON, Matthew ; CHANDRASEKAR, Srinivasan ; MOTTA, Ricardo: *A Standard Default Color Space for the Internet — sRGB*. [Online]: <http://www.w3.org/Graphics/Color/sRGB.html> (visited on January 22, 2010). 1996

Bibliography

- [Tektronix Inc. 1998] TEKTRONIX INC.: *WFM90 & WFM91 User Manual*. 1998
- [Tektronix Inc. 2002] TEKTRONIX INC.: *Application Note: Preventing Illegal Colors*. [Online]: http://www.tek.com/Measurement/App_Notes/25_15609/25W_15609_0.pdf (visited on January 10, 2010). 2002
- [The MathWorks Inc. 2009] THE MATHWORKS INC.: *Image Acquisition Toolbox Adaptor Kit: User's Guide*. 2009
- [University College London 2010] UNIVERSITY COLLEGE LONDON: *Colour & Vision Research Laboratory and Database*. [Online]: <http://www-cvrl1.ucsd.edu> (visited on January 20, 2010). 2010
- [Wang and Bovik 2006] WANG, Zhou ; BOVIK, Alan C.: *Modern Image Quality Assessment*. Morgan & Claypool, 2006
- [Wikipedia 2010] WIKIPEDIA: *MacAdam Ellipse*. [Online]: http://en.wikipedia.org/wiki/MacAdam_ellipse (visited on February 09, 2010). 2010
- [Wueller 2006] WUELLER, Dietmar: *White Paper: Image Engineering Digital Camera Tests*. [Online]: http://www.image-engineering.de/images/pdf/whitepaper/cameratests_whitepaper_1.0.pdf (visited on January 12, 2010). 2006
- [Wyszecki and Stiles 2000] WYSZECKI, Guenther ; STILES, W.S.: *Color Science: Concepts and Methods, Quantitative Data and Formulae*. 2nd Edition. John Wiley & Sons Ltd., 2000
- [X-Rite Inc. 2007] X-RITE INC.: *A Guide to Understanding Color Communication*. [Online]: http://www.xrite.com/documents/literature/en/L10-001_Understand_Color_en.pdf (visited on February 10, 2010). 2007

Appendices

A. Tables

Standard		Red	Green	Blue	White	White Point
NTSC (1953)	x	0.670	0.210	0.140	0.3100	CIE C
	y	0.330	0.710	0.080	0.3160	
	z	0.000	0.080	0.780	0.3740	
EBU Tech. 3213	x	0.640	0.290	0.150	0.3127	CIE D65
	y	0.330	0.600	0.060	0.3290	
	z	0.030	0.110	0.790	0.3582	
SMPTE RP 145	x	0.630	0.310	0.155	0.3127	CIE D65
	y	0.340	0.595	0.070	0.3290	
	z	0.030	0.095	0.775	0.3582	
ITU-R BT.709	x	0.640	0.300	0.150	0.3127	CIE D65
	y	0.330	0.600	0.060	0.3290	
	z	0.030	0.100	0.790	0.3582	

Table A.1.: CIE x, y and z coordinate values of primary chromaticities for today's video standards (according to Poynton (2007)).

A. Tables

Standard		Red	Green	Blue	White	White Point
sRGB	x	0.640	0.300	0.150	0.3127	CIE D65
	y	0.330	0.600	0.060	0.3290	
	z	0.030	0.100	0.790	0.3582	
AdobeRGB	x	0.640	0.210	0.150	0.3127	CIE D65
	y	0.340	0.710	0.060	0.3290	
	z	0.020	0.080	0.790	0.3582	
ROMM RGB	x	0.7347	0.1596	0.0366	0.3457	CIE D50
	y	0.2653	0.8404	0.0001	0.3585	
	z	0.0000	0.0000	0.9633	0.2958	

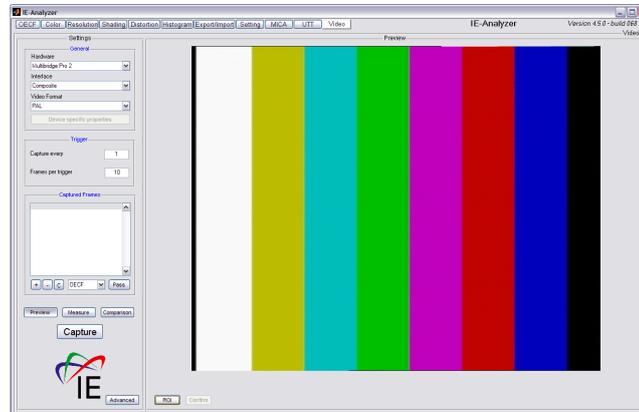
Table A.2.: CIE x, y and z coordinate values of primary chromaticities for some RGB color spaces for computer graphics (according to Süsstrunk et al. (1999)).

Mode	Width	Height	FPS	Fields per frame
bmdModeNTSC	720	486	29.97	2
bmdModeNTSC2398	720	486	29.97	2
bmdModePAL	720	576	25	2
bmdModeHD1080psf2398	1920	1080	23.97	2
bmdModeHD1080psf24	1920	1080	24	2
bmdModeHD1080i50	1920	1080	25	2
bmdModeHD1080i5994	1920	1080	29.97	2
bmdModeHD720p50	1280	720	50	1
bmdModeHD720p5994	1280	720	59.94	1
bmdModeHD720p60	1280	720	60	1
bmdMode2k2398	2048	1556	23.97	2
bmdMode2k24	2048	1556	24	2

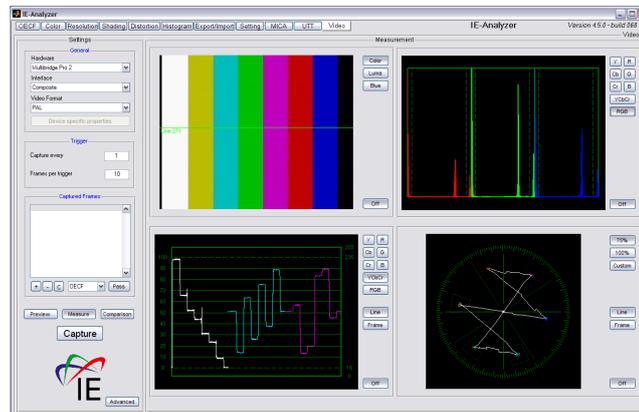
Table A.3.: Video formats supported by the DeckLink SDK.

B. Screenshots

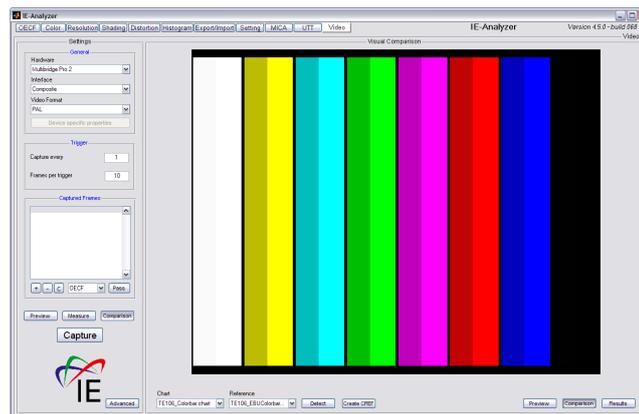
B. Screenshots



(a) Preview



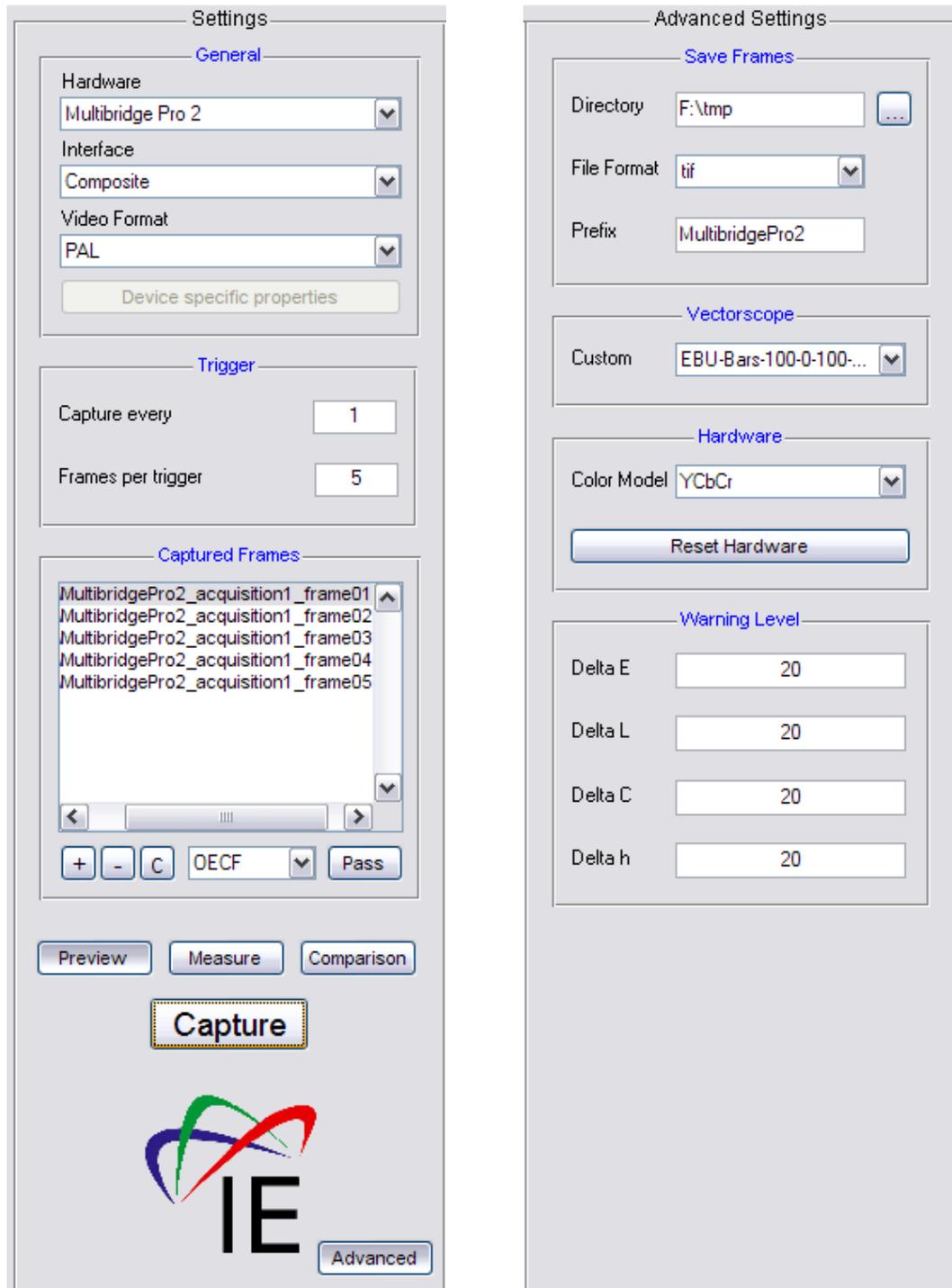
(b) Measurement



(c) Comparison

Figure B.1.: Screenshots of the Video Module.

B. Screenshots



(a) General settings

(b) Advanced settings

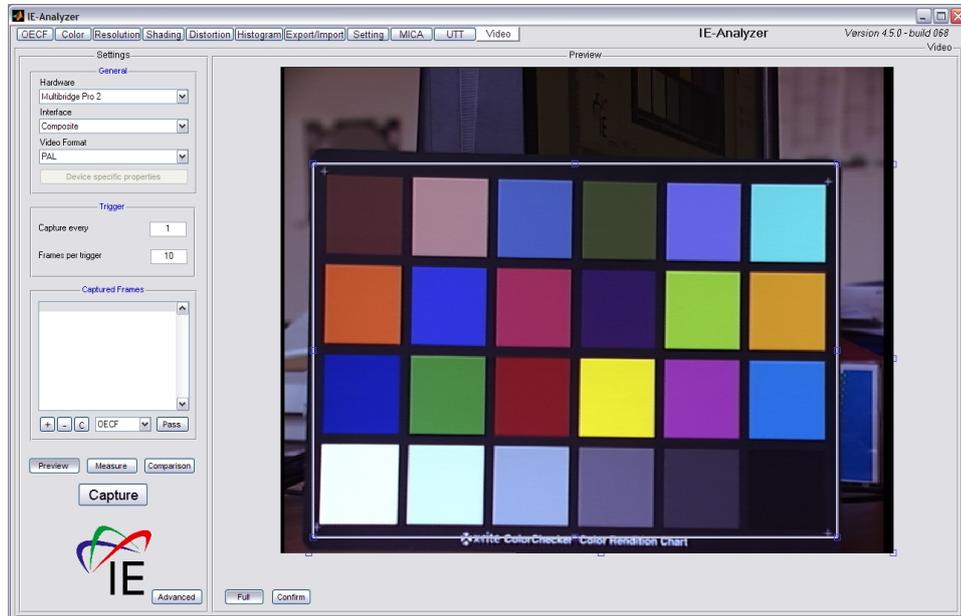
Figure B.2.: Screenshots of the settings panels of Video Module.

B. Screenshots

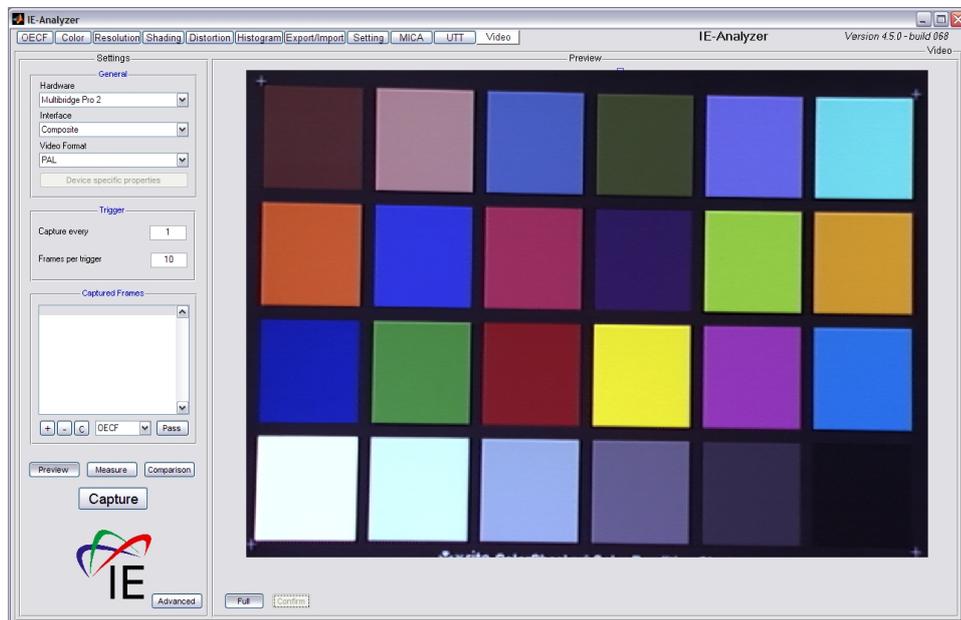


Figure B.3.: Device specific settings for Labtec WebCam Pro. The settings are dynamically collected from the adaptor DLL and summarized in the dialog window.

B. Screenshots



(a) Selected ROI



(b) Cropped image

Figure B.4.: Screenshots outlining the selection of the region of interest (ROI).

B. Screenshots

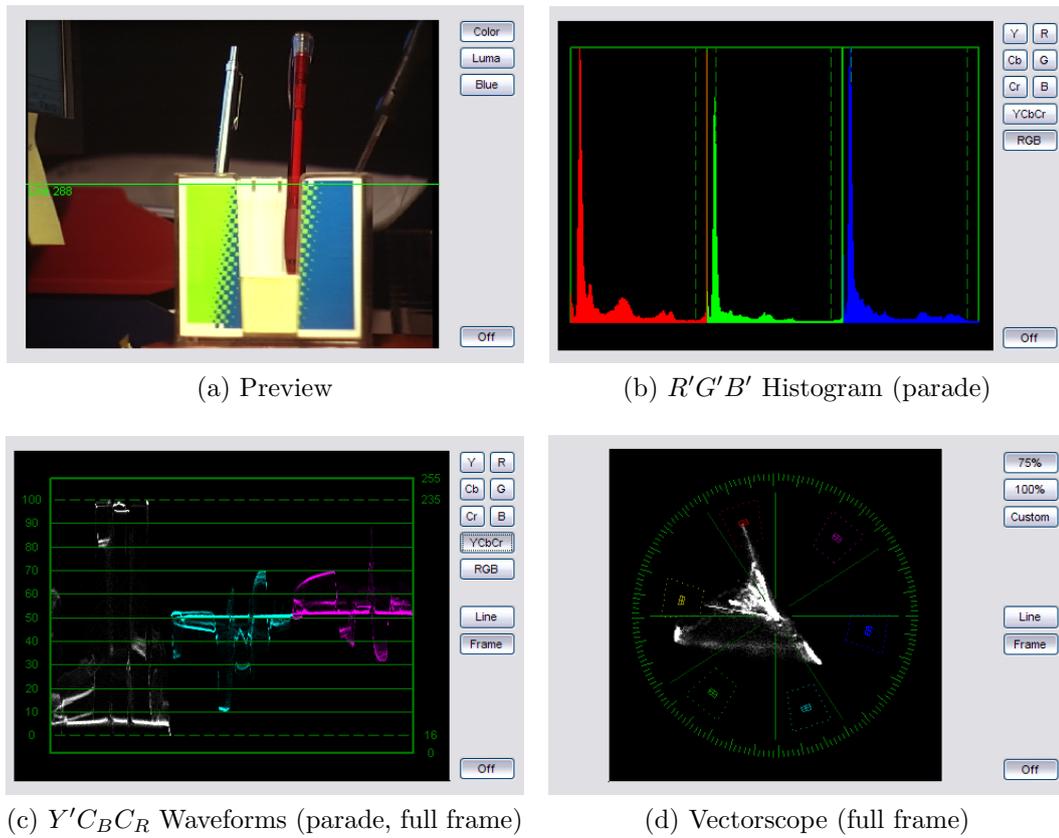
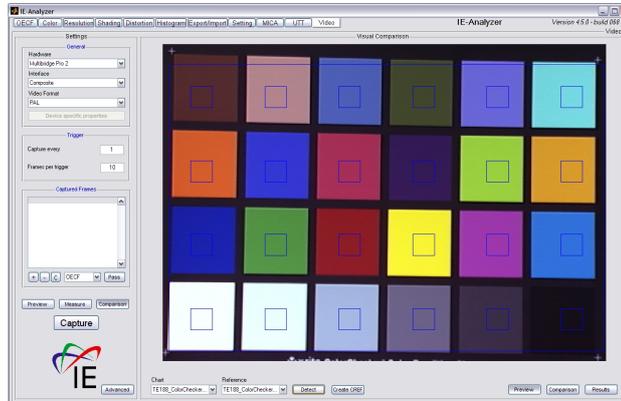
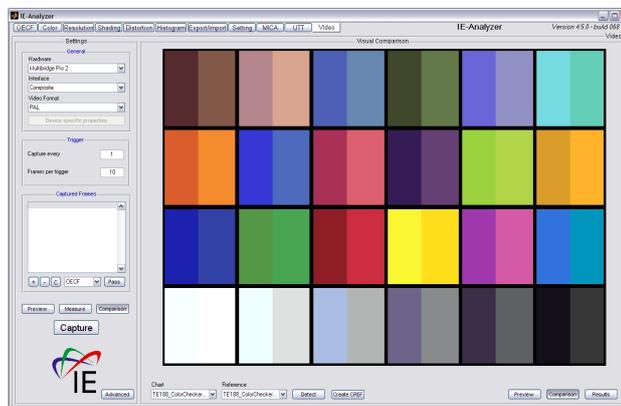


Figure B.5.: Measurement monitors of the Video Module.

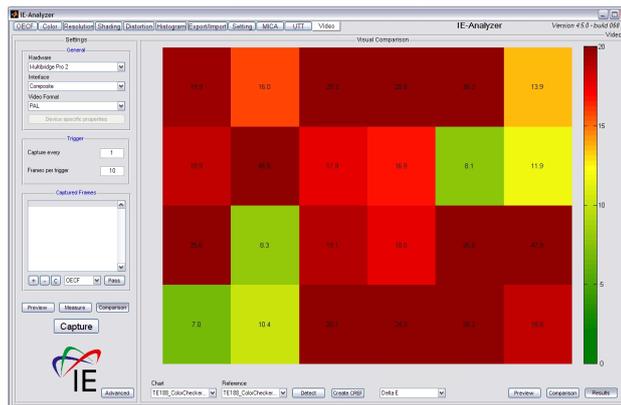
B. Screenshots



(a) Detected test chart



(b) Visual comparison. Video data (left) vs. color reference (right)



(c) Numerical $CIE76 \Delta E_{ab}$ evaluation

Figure B.6.: Screenshots of the Video Module's Comparison submodule. The pseudo-color coding of numerical ΔE_{ab} values can be adjusted in the advanced settings.

C. Classes (UML)

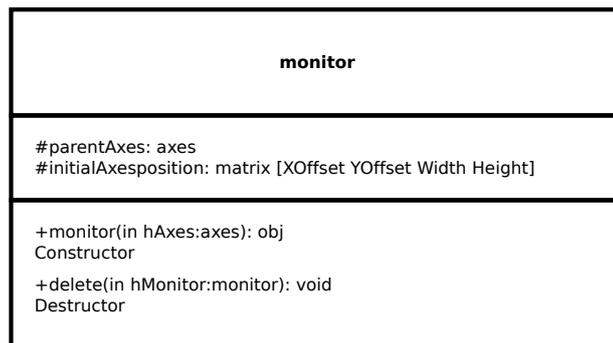


Figure C.1.: Class `monitor`.

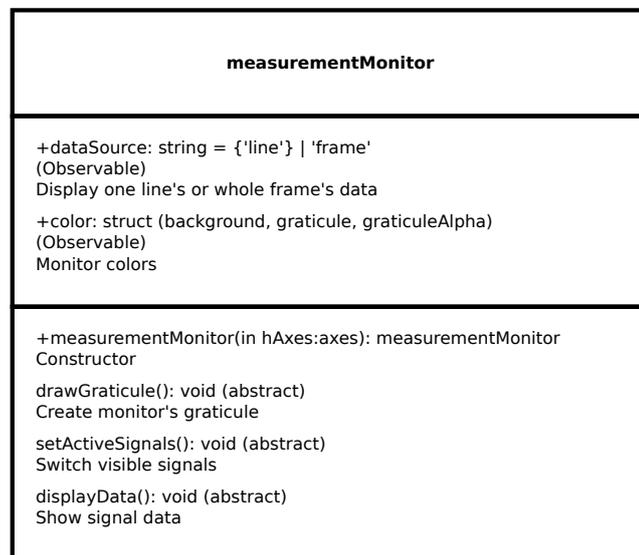


Figure C.2.: Class `measurementMonitor`.

C. Classes (UML)

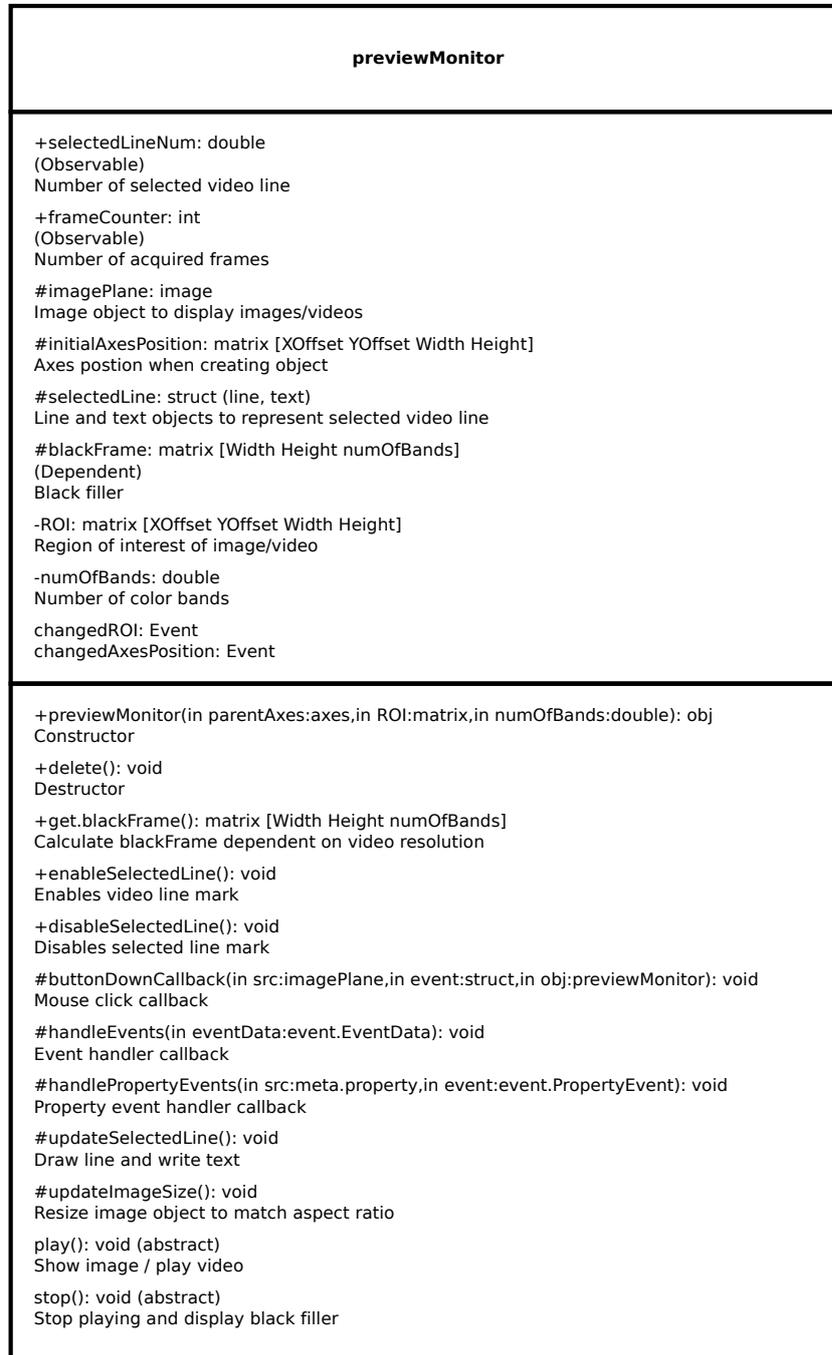


Figure C.3.: Class previewMonitor.

C. Classes (UML)

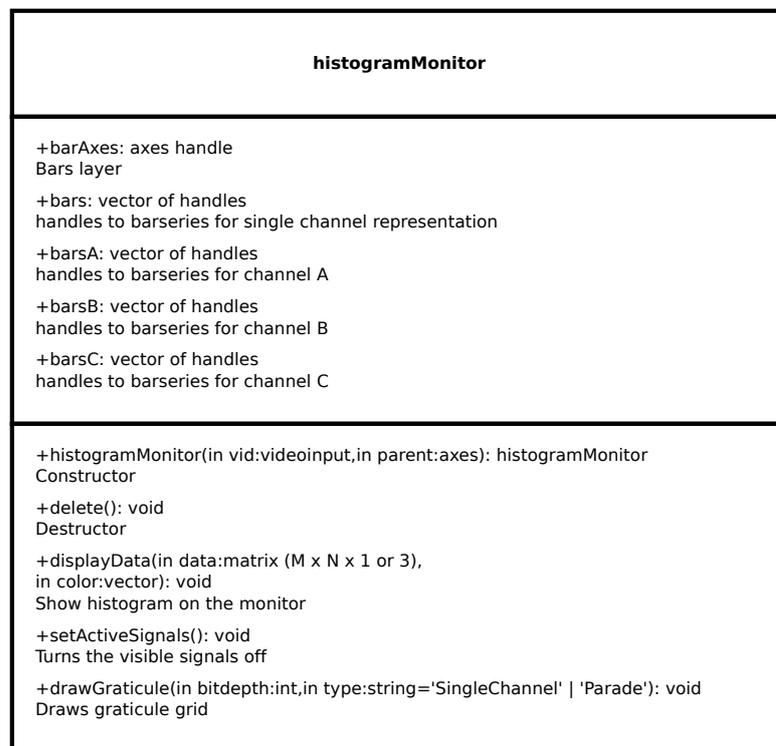


Figure C.4.: Class `histogramMonitor`.

C. Classes (UML)

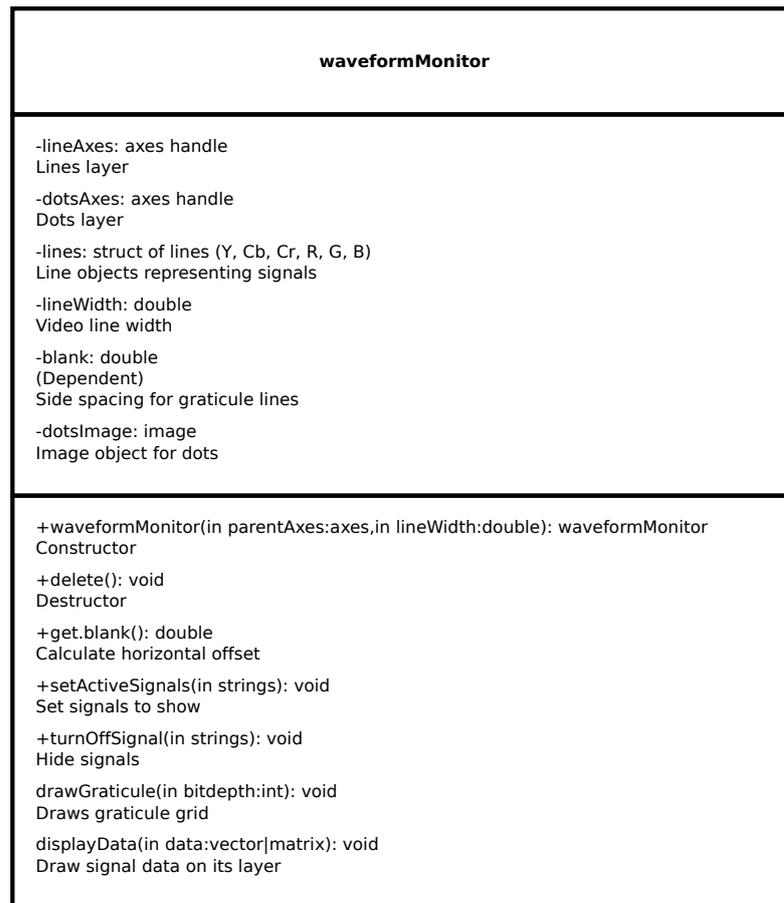


Figure C.5.: Class waveformMonitor.

C. Classes (UML)

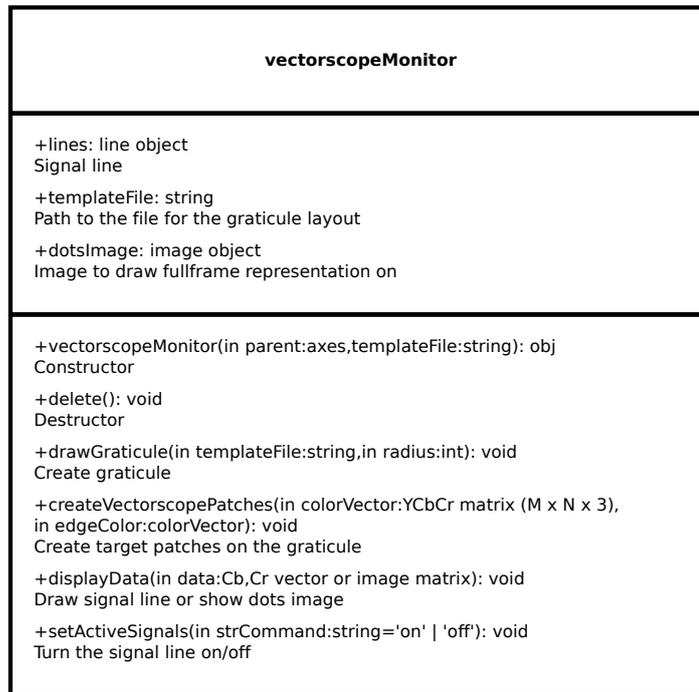


Figure C.6.: Class vectorscopeMonitor.

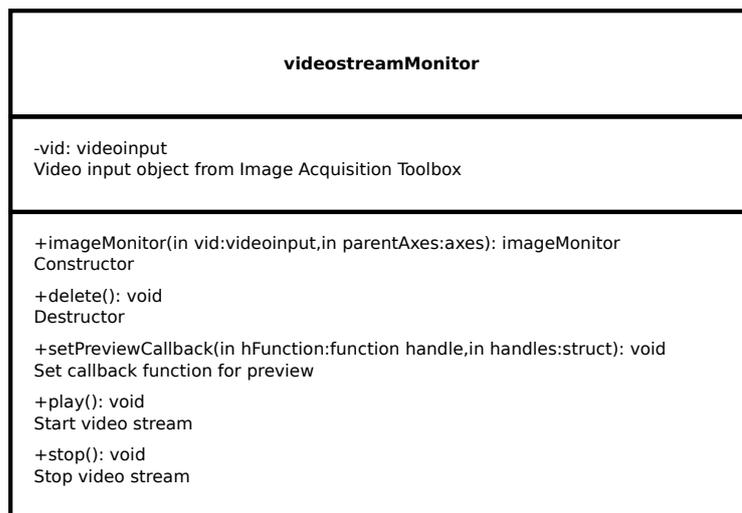


Figure C.7.: Class videostreamMonitor.

C. Classes (UML)

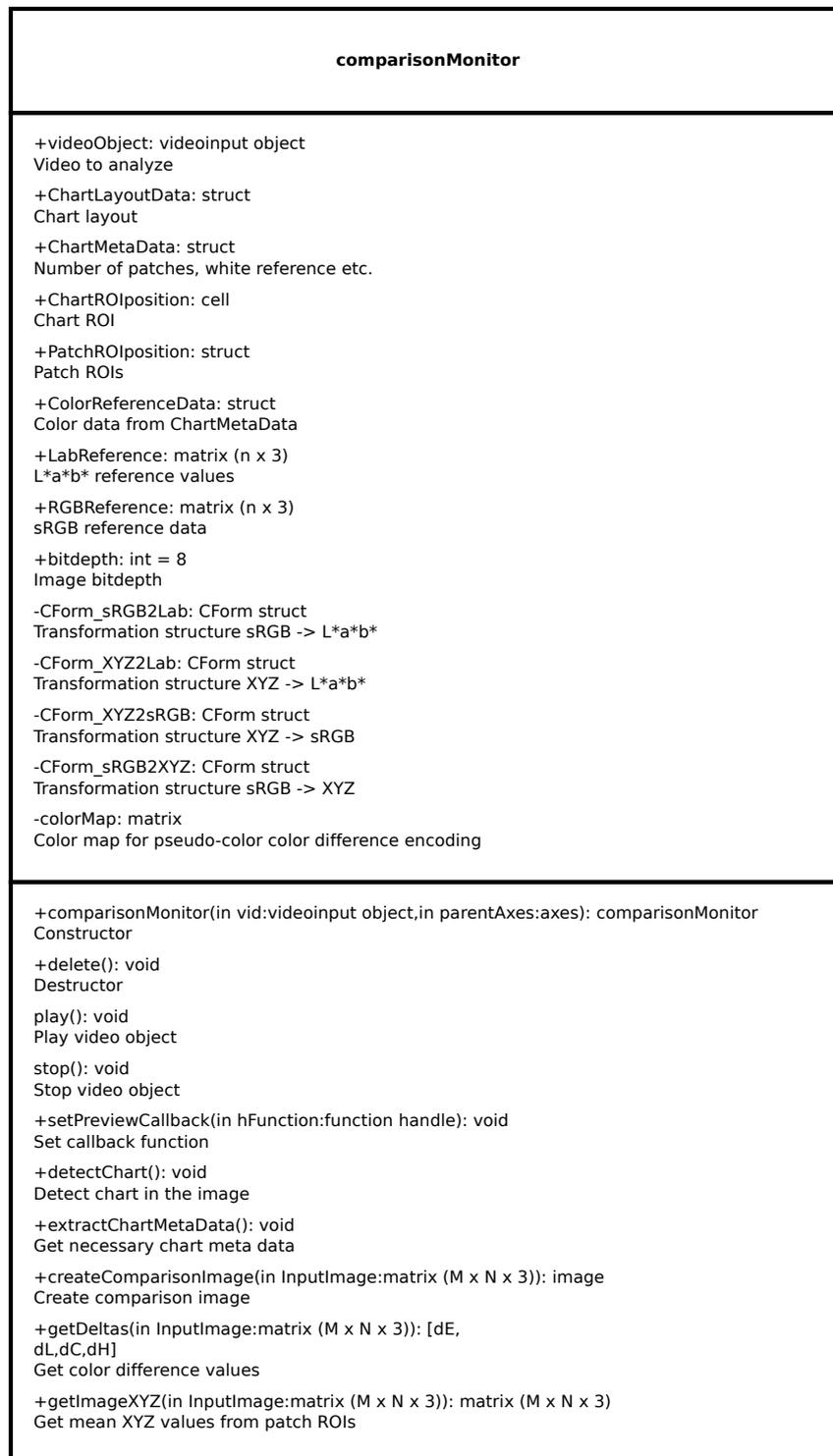


Figure C.8.: Class comparisonMonitor.

Remarks

Eidesstattliche Erklärung

Ich versichere, dass ich die vorliegende Arbeit in dem gemeldeten Zeitraum selbstständig verfasst und keine anderen als die angegebenen Hilfsmittel verwendet habe. Die Stellen, die anderen Werken wörtlich oder sinngemäß entnommen sind, sind als solche kenntlich gemacht. Ich versichere weiterhin, dass die Arbeit in gleicher oder ähnlicher Form noch keiner anderen Prüfungsbehörde vorgelegen hat.

Affirmation

I certify that the work presented here is, to the best of my knowledge and belief, original and the result of my own investigations, except as acknowledged, and has not been submitted, either in part or whole, for a degree at this or any other audit authority.

Köln, den 19. Februar 2010
Cologne, February 19, 2010

Borys Golik

Weitergabeerklärung

Ich erkläre hiermit mein Einverständnis, dass das vorliegende Exemplar meiner Masterarbeit oder eine Kopie hiervon für wissenschaftliche Zwecke verwendet werden darf.

Declaration of Publication

I declare that this thesis and/or a copy of it may be used for scientific purposes.

Köln, den 19. Februar 2010

Cologne, February 19, 2010

Borys Golik

Sperrvermerk

Die vorgelegte Arbeit unterliegt keinem Sperrvermerk.

Remark of Closure

This thesis is not closed.