Perceptually-Inspired Gamut Mapping for Display and Projection Technologies

Syed Waqas Zamir

DOCTORAL THESIS UPF / YEAR 2017

Thesis Advisors: Marcelo Bertalmío and Javier Vazquez-Corral
Department of Information and Communication Technologies

To my parents, grandmother, and Khazina

Acknowledgments

First and foremost, I would like to express my utmost gratitude to my supervisors Marcelo Bertalmío and Javier Vazquez-Corral for their encouragement, continuous support and guidance throughout all these long years of research. Without their invaluable experience and mentoring this work would not have been possible.

I would like to thank all the participants of the psychophysical experiments. Special thanks to Stephan Cattan, from Deluxe (Spain), for providing us with test images and for all his help and suggestions. I would like to express my gratitude to Barco N.V. for their invaluable help, without which this work would not have been possible.

Many thanks and cheers go to my dear wife Khazina for her endless love, encouragement and unconditional support even when she had to endure practical and emotional discomfort several times during my PhD.

I am very grateful to Raquel Gil Rodriguez for being an awesome friend and always willing to help whether it was at work, or fighting against Bureaucracy in public offices. For all the favors I am greatly in your debt — forever. Thanks to my very kind (and annoyingly polite) friend and colleague Gabriela Ghimpeteanu for always hearing me out patiently. A very special thanks to my friend Praveen Cyriac for invaluable discussions, help and support throughout my PhD. I am also very thankful to my many other friends and colleagues: Zulqarnain Rashid, Thomas Batard, David Kane, Jihyun Kim, Antoine Grimaldi, Adnan Butt, Umar Iqbal, Hashim Shahid, Sadegh Mohammadi, Kamruddin Nur, Raul Parada, Aamir Saleem and Jose Maria Rojano.

Very special thanks to the most amazing and friendly secretaries: Jana Safrankova, Lydia Garcia, Joana Soria, Joana Clotet, Venessa Jimenez, Magda Castellnou, Ruth Temporal, Montserrat Serrano, and Lluis Bosch.

Words cannot express how thankful I am to my parents for all their unconditional support, love, help, encouragement and hard work in getting me where I am today. Thanks to my other family members: my brothers Sohail and Haroon, my sister Sahar, my grandparents especially my late grandmother Sughra Bibi (whom I miss dearly).

Abstract

The cinema and television industries are continuously working in the development of image features that can provide a better visual experience to viewers; these image attributes include large spatial resolution, high temporal resolution (frame rate), greater contrast, and recently, with emerging display technologies, much wider color gamut. The gamut of a device is the set of colors that this device is capable of reproducing. Gamut Mapping Algorithms (GMAs) transform colors of the original content to the native color palette of the display device with the simultaneous goals of (a) reproducing content accurately while preserving the artistic intent of the original content's creator and (b) exploiting the full color rendering potential of the target display device. There are two types of gamut mapping: Gamut Reduction (GR) and Gamut Extension (GE). GR involves the transformation of colors from a larger source gamut to a smaller destination gamut. Whereas in GE, colors are mapped from a smaller source gamut to a larger destination gamut.

In this thesis we propose three spatial Gamut Reduction Algorithms (GRAs) and four spatial Gamut Extension Algorithms (GEAs). These methods comply with some basic global and local perceptual properties of human vision, producing state of the art results that appear natural and are perceptually faithful to the original material. Moreover, we present a psychophysical evaluation of GEAs specifically for cinema using a digital cinema projector under cinematic (low ambient light) conditions; to the best of our knowledge this is the first evaluation of this kind reported in the literature. We also show how currently available image quality metrics, when applied to the gamut extension problem, provide results that do not correlate well with users' choices.

Resum

Les indústries de cinema i televisió estan treballant contínuament en el desenvolupament de diferents característiques de la imatge que puguin proporcionar una millor experiència visual per als espectadors; aquests atributs d'imatge inclouen la resolució espacial, la resolució temporal (fotogrames per segon), major contrast i, recentment, amb les noves tecnologies de visualització emergents, una gamma de colors (gamut) molt més ampli. El gamut d'un dispositiu és el conjunt de colors que aquest dispositiu és capa de reproduir. Els algoritmes de modificació de gamut (GMA, de les seves sigles en anglés) transformen els colors del contingut original a la paleta de color del dispositiu de visualització amb els objectius de (a) reproduir el contingut amb precisió preservant al mateix temps la intenció artística del creador del contingut original i (b) utilitzar tot el gamut de color del dispositiu de visualització. Hi ha dos tipus d'algoritmes de modificació de gamut: Reducció de Gamut (GR) i Extensió de Gamut (GE). GR implica la transformació dels colors d'un gamut d'origen més gran a un gamut de destinació més petit. Mentre que a GE, els colors s'assignen d'un gamut d'origen petit a un gamut de destinació més gran.

En aquesta tesi es proposen tres algoritmes de Reducció de Gamut (GRAs) i quatre algoritmes d'extensió de Gamut (GEAs). Aquests mètodes compleixen amb algunes propietats perceptives globals i locals bàsiques de la visió humana, produint resultats que són estat de l'art, i que són naturals i perceptualment fidels al material original. D'altra banda, es presenta una avaluació psicofísica del problema d'extensió de Gamut específicament dissenyada per a cinema utilitzant un projector de cinema digital en condicions cinemàtiques (baixa llum ambiental); aquest estudi creiem que és el primer del seu tipus a la literatura. També mostrem com les mètriques de qualitat d'imatge disponibles actualment proporcionen resultats que no es correlacionen bé amb l'elecció dels usuaris quan s'apliquen al problema d'Extensió de Gamut.

Contents

Nomenclature

2D Two Dimensions

3D Three Dimensions

BT.2020 Abbreviation for the next-generation TV standard ITU-R BT.2020 [ITU-R, 2012]

BT.709 Abbreviation for the broadcast standard ITU-R BT.709 [ITU-R, 2002]

CAT Chromatic Adaptation Transform

CCD Charge Coupled Device

CCFL Cold Cathode Fluorescent Lamp

cd/m^2 Candela per square meter

CFA Color Filter Array

CIE International Commission on Illumination

CMOS Complementary Metal Oxide Semiconductor

CRT Cathode Ray Tube

DCI Digital Cinema Initiatives

DCI-P3 Abbreviation for the cinema standard [SMPTE, 2011]

DLP Digital Light Processing

DMD Digital Micromirror Device

EOTF Electro-Optical Transfer Function

GE Gamut Extension

GEA Gamut Extension Algorithm

GM Gamut Mapping

GMA Gamut Mapping Algorithm

GR Gamut Reduction

GRA Gamut Reduction Algorithm

HCFL Hot Cathode Fluorescent Lamp

HDR High Dynamic Range

HDTV High-definition Television

HLG Hybrid Log Gamma

HPMINDE Hue Preserving Minimum ΔE

HSV Hue Saturation Value

K Kelvin

LCD Liquid Crystal Display

LCoS Liquid Crystal on Silicon

LED Light Emitting Diode

LGN Lateral Geniculate Nucleus

LUT Look Up Table

MSE mean square error

nm Nano meters

OETF Opto-Electrical Transfer Function

OOG Out-of-gamut

OOTF Opto-Optical Transfer Function

PQ Perceptual Quantizer

PSNR Peak Signal to Noise Ratio

QD Quantum Dot

RGB Red Green Blue

SGCK Sigmoidal Gaussian Cusp Knee

SMPTE Society of Motion Picture and Television Engineers

SPD Spectral Power Distribution

TV Television

UHDTV Ultra High-definition Television

WCG Wide Color Gamut

List of Figures

List of Tables

CHAPTER 1

Introduction

To enhance the overall viewing experience both for cinema and television (TV), the media industry is continuously striving to improve image quality with higher frame rates, larger resolution, vivid colors, and greater contrast. The normal human eye contains three types of cone cells that respond to incident light to produce the sensation of color. Two lights that produce the same cone response triplet are perceived as having the same color, even if they have a different physical power distribution. This allows to generate any perceivable color by a proper mixture of any three given lights (as long as they are colorimetrically independent, i.e. that the mixing of two of them does not produce the same color as the remaining one), in what is known as the trichromacy property. Therefore, given three light 'primaries', any color is characterized by the triplet of weights with which it can be generated as a mixture of the primaries. So colors can be represented as points in a three-dimensional space, although it is common to ignore light intensity and just represent the chromatic content of light as two-dimensional points on a plane: Fig. 1.1 shows the standard CIE (International Commission on Illumination) xy chromaticity diagram, where the horseshoe-shaped region corresponds to the chromaticities of all the colors a standard observer can physiologically see. While the trichromacy property states that any color can be expressed as a linear combination of a given set of primaries, it is important to note that the weights for the linear combination can be negative. Since most display systems are based on mixing three primary lights by regulating the power contribution of each, the color set that a display can generate (its *color*

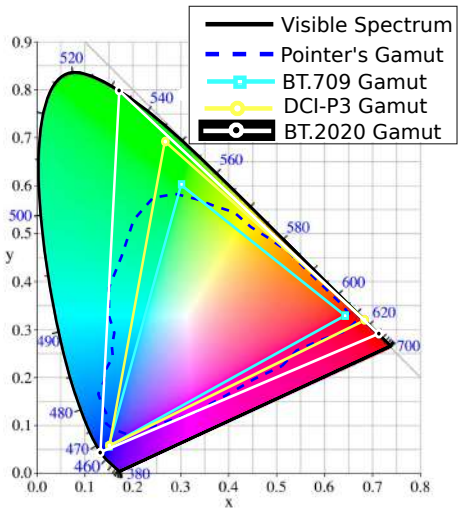


Figure 1.1: Gamuts on CIE xy chromaticity diagram.

gamut) is limited to the colors that can be obtained with linear combinations that use *positive* weights, because negative weights imply a physically-unrealizable light of negative power. Given that each primary can be represented as a point in the CIE xy diagram, the color gamut of a display is therefore the triangle that has the primaries as its vertices, i.e. the region covered by linear combinations with positive weights, and as a consequence trichromatic displays are unable to reproduce many colors that we can see. With the goal of making a display device that can reproduce more of the colors that our visual system can perceive, several multiple-primary displays have been proposed that make use of four [Chino et al., 2006], five [Ueki et al., 2009], [Cheng et al., 2010], and even six color primaries [Yang et al., 2005], [Sugiura et al., 2006b], [Sugiura et al., 2006a], [Ajito et al., 2000], [Roth et al., 2003], however, the quest to make an optimal display still continues.

State of the art digital movie cameras (ARRI Alexa, RED Epic, Canon EOS C700, etc.) are capable of shooting content with a large range of colors. However, before a movie is released, its colors have to be reproduced according to the standard gamuts: DCI-P3 [SMPTE, 2011] is used for digital cinema projection and it is based on the gamut of Xenon lamps, and BT.709 [ITU-R, 2002], with

primaries close to the phosphor primaries of cathode ray tubes (CRTs), is used for cable and broadcast TV, DVD, Blu-Ray and streaming. These standard color gamuts DCI-P3 and BT.709 (shown in Fig. 1.1) exist to ensure a consistent movie presentation across different digital cinema projectors and TVs, respectively. This adaptation to standard gamut implies altering the range of colors (and contrast) of the original captured content. To reach the objective of color reproduction, in the cinema industry, colorists (expert technicians) at the post-production stage perform gamut modifications using three-dimensional look-up tables (LUTs). These LUTs contain millions of entries and colorists only specify a few colors manually, while the rest are interpolated without taking care of their spatial or temporal context [Bertalmío, 2014]. Subsequently, the resulting video may have false colors that were not present in the original material and intensive manual correction is usually necessary, commonly performed in a shot-by-shot, object-by-object basis. This process is difficult, time consuming and expensive, and therefore it makes an automated procedure called Gamut Mapping (GM) very desirable: GM transforms colors of an input image to the colors of a target device.

There are two types of GM: Gamut Reduction (GR) and Gamut Extension (GE). Gamut reduction involves the mapping of colors from a larger source gamut to a smaller destination gamut. For example, cinema footage needs to be passed through a gamut reduction method in order to be displayed on a television screen [Bankston, 2005] [Kennel, 2007]. On the other hand, gamut extension refers to the transformation of colors from a smaller source gamut to a larger destination gamut. Gamut extension is needed by state of the art digital cinema projection systems, which often receive a cinema signal that is encoded with a limited gamut as precaution against commonly poor projectors; therefore, in order to realize the full potential of these projectors in terms of colors, a gamut extension procedure is required [Bertalmío, 2014].

Informally, any color gamut that is larger than the BT.709 gamut is referred as wide color gamut (WCG). Gamut mapping usually implies the use of a GR procedure due to the fact that the professional movie cameras capture content that has a wider color gamut than that of the standard gamuts. However, the process of GE is gaining importance. [Pointer, 1980] analyzed frequently occurring real surface colors and derived what is commonly known as "Pointer's gamut", shown in Fig. 1.1. Although both DCI-P3 and BT.709 cover a reasonable amount of Pointer's gamut, many interesting real world colors fall outside these standard gamuts. In 2012, the International Telecommunication Union-Radiocommunication (ITU-

R) recommended a new standard gamut BT.2020 [ITU-R, 2012] for the next generation ultra-high definition TV that encompasses DCI-P3 and BT.709 and covers 99.9% of Pointer's gamut. New laser projectors have pure (very saturated) primaries [Beck, 2014] and therefore they are able to cover the very wide BT.2020 gamut [Silverstein et al., 2011], [Kusakabe et al., 2013], reproducing nearly every color found in nature and providing the audience with a more compelling color experience. But if the inputs are movies with DCI-P3 gamut, as virtually all professional movies currently are, the full color rendering potential of these new projectors cannot be realized. The same issue happens presently when DCI-P3 projectors are used to display pictures that come in BT.709 (usually because the movie distributor wants to prevent issues with lower quality or older projectors, or the movie was originally prepared for broadcast only). In both cases there is a pressing need to develop gamut extension (GE) techniques that automatically extend the gamut of the movie content, with the very challenging constraint that this has to be done in a way that the appearance of the gamut extended result preserves as much as possible the artistic intent of the content's creator.

Gamut mapping is not only important in the film or the broadcast industry but it is also an essential module in the image reproduction pipeline of printing technologies, where the end goal is to minimize the perceptual difference of the same image when it is viewed on a display device and when it is printed. Other application domains of gamut mapping include handheld devices (mobile, tablet computers), websites, photo-sharing online platforms, computer graphics, animation and video games.

To summarize the aforementioned discussion, we would like to mention a key difference between the application of gamut reduction and gamut extension. Gamut reduction is a necessary procedure to perform when the colors of the input image fall outside the display's gamut in order to reproduce colors accurately; if not, the display will reproduce the image with artifacts and loss of spatial detail. On the contrary, gamut extension is not a must procedure to apply, rather it is considered as an enhancement operation [Morovič, 2008]. For example, displaying BT.709 footage (represented in BT.2020 container) as it is on a wide-gamut BT.2020 supported display device will not cause any visual color distortion, but if we do not extend colors of the input footage to the gamut of the display we will be missing the color rendering potential of the wide-gamut display device. Thus the objective of this thesis is the following: to formulate gamut mapping algorithms (GMAs) that can reproduce colors of a given footage (captured using

any mobile camera, low-end consumer camera, or professional cinema camera) according to the gamut of the display on which it is going to be displayed, while staying perceptually faithful to the original material.

1.1 Main Contributions

A large number of GMAs have been proposed in the literature (as we shall see in chapter 4). GMAs are generally classified into two broad categories. The first category consists of global (also called non-local or non-adaptive) GMAs that modify each color of an image independently, meaning that these methods completely ignore the spatial color distribution in the image. However, this is not how the visual system works. The human visual system is very effective in processing the visual information in highly complex scenes. The interpretation of stimulus occurs at different levels in the visual system, from retina to cortex. And our perception of the color of an object is not completely determined by the light coming from it, instead it is heavily influenced by the light coming from other objects in its neighborhood [Land and McCann, 1971]. For example, the two inner squares in Fig. 1.2 have physically the same color but appear to us differently because of the surrounding colors. To incorporate such functionality, the second category of gamut mapping methods involve local GMAs that modify pixel values taking into account their local neighborhoods in order to better preserve colors and details in the reproduction. This property certainly makes GMAs adaptive and flexible but at the same time far more complex and computationally expensive than global GMAs. As of today, in order to reproduce movie content, the time consuming manual procedures by skilled technicians are preferred over

Figure 1.2: Effect of local neighborhood on color sensation. The small squares in both patches have same color value but different color sensation.

automatic gamut mapping algorithms, because colorists always work following and respecting the vision of the movie creator [Van Hurkman, 2013], while automatic methods are prone to issues for which audiences are very sensitive and that therefore severely affect the quality of the results, like introducing color artifacts, hue shifts, and not preserving the integrity of memory colors (e.g. the blue of the sky, the green of grass) or skin tones [Heckaman and Sullivan, 2011], [Pan and Daly, 2008], [Morovič, 2008]. Aiming at the accurate reproduction but low computational cost, in this study we present several local gamut reduction algorithms (GRAs) and gamut extension algorithms (GEAs).

1. We present two spatial GMAs (one GRA and one GEA) in RGB color space that rely on a perceptually-based variational framework. Our algorithms adapt the image energy functional of [Bertalmío et al., 2007] whose minimization leads to image enhancement and contrast modification. We show how by varying the importance of the contrast term in the image functional we are able to perform both gamut reduction and gamut extension.

2. We propose a mathematical framework that allows us to modify our first GRA in order to better respect the saliency of the original image in the reproduced image.

3. We propose a spatial GEA, implemented as a partial differential equation-based optimization procedure related to visual perception models, that performs gamut extension in CIELAB color space by taking into account the analysis of distortions in hue, chroma and saturation. Moreover, in order to evaluate GEAs, we present user studies performed using a digital cinema projector under cinematic (low ambient light, large screen) conditions. We also show how currently available image quality metrics, when applied to the gamut extension problem, provide results that do not correlate with users' choices.

4. We present two spatial GMAs (one GRA and one GEA) in HSV color space that are based on the work of [Bertalmío et al., 2009]. The proposed GEA adapts itself according to the color distribution of the input image. On one side, the proposed GEA modifies in a controlled manner those colors that require a special treatment such as skin tones, less saturated natural objects, neutral colors and some particular memory colors. On the other side, it extends normally colors of high chromatic (natural and artificial)

objects. The proposed GRA is developed by introducing simple but key modifications to the framework of our GEA (in HSV color space).

The goal of this research is to develop GMAs that can reproduce content respecting as much as possible the artist's vision. One important aspect that we would like to discuss here is the role of the ground truth in the evaluation of our GMAs. In general, and not only in the case of legacy material, the ground truth data is never available. The original material is all there is, and under the scenario of gamut mismatch (between the source material and the destination device) this material has to have its gamut modified while preserving as much as possible its original appearance. In our tests we include a ground truth for evaluation purposes. To assess the performance of GRAs, we start from a wide-gamut material, reduce its gamut using a GRA and then ask users to compare the result with the wide-gamut ground truth. And in the gamut extension case, we start from a wide-gamut material, reduce its gamut, extend it with a GEA and then ask users to compare the result with the original wide-gamut ground truth. Nonetheless, by showing that a GMA performs well as evaluated using ground truth data, we expect that it also performs well when ground truth data is not available. This way of evaluation is a common approach in image processing (e.g. denoising algorithms are evaluated in terms of PSNR (peak signal to noise ratio) by comparing with a ground truth "clean" image that is never available in real scenarios, segmentation algorithms are evaluated by comparing their results to those of manual segmentations, etc.) In this study, we ask users to choose the most accurate result instead of the one they find most pleasant. The reason is that, for a gamut mapping technique to be adopted by the movie industry, it must yield gamut mapped results that preserve as much as possible the artistic intent of the content's creator. Therefore, user tests based on accuracy provide results that are less subjective (as users are simply asked to estimate differences with respect to the provided reference) and the preservation of the artistic intent is in a sense "built-in" in the procedure, since the ground truth acts as a stand-in for the content's creator intent. Apart from the psychophysical testing, in this study we also evaluate the performance of GMAs using several image quality metrics.

1.1.1 Publications

This study yielded the following papers in international conferences and journals.

Journals

- S.W. Zamir, J. Vazquez-Corral, and M. Bertalmío. Gamut Extension for Cinema. IEEE Transactions on Image Processing, 26(4):1595-1606, 2017.
- S.W. Zamir, J. Vazquez-Corral, and M. Bertalmío. Gamut mapping in cinematography through perceptually-based contrast modification. IEEE Journal of Selected Topics in Signal Processing, 8(3):490-503, 2014.

Conferences

- S.W. Zamir, J. Vazquez-Corral, and M. Bertalmío. Perceptually-based gamut extension algorithm for emerging wide color gamut display and projection technologies. In proc. of Society of Motion Picture and Television Engineers (SMPTE) Annual Technical Conference and Exhibition, 2016.
- J. Vazquez-Corral, S.W. Zamir, A. Galdran, D. Pardo, and M. Bertalmío. Image processing applications through a variational perceptually-based color correction related to Retinex. In proc. of IS&T Electronic Imaging Conference, 2016.
- S.W. Zamir, J. Vazquez-Corral, and M. Bertalmío. Gamut extension for cinema: psychophysical evaluation of the state of the art, and a new algorithm. In proc. of IS&T Electronic Imaging Conference, 2015.
- J. Vazquez-Corral, S.W. Zamir, and M. Bertalmío. Considering saliency in a perception inspired gamut reduction algorithm. In proc. of IS&T/SID 22nd Color and Imaging Conference, 2014.
- S.W. Zamir, J. Vazquez-Corral, and M. Bertalmío. Gamut Mapping through Perceptually-based Contrast Reduction. Pacific-Rim Symposium on Image and Video Technology (PSIVT), 2013.

1.2 Thesis Outline

The rest of this document is organized as follows. In chapter 2, we cover some basic concepts of color science including colorimetry and standard color spaces.

We also present details about the interaction of light with the human visual system. Moreover, we discuss some global and local characteristics of the human vision, and the effects of ambient lighting on perception. In chapter 3, we discuss concepts of the image processing pipeline of digital cameras. Later we review the standard color gamuts and present methods to perform transformations among them. There are a variety of display devices available in the marketplace; each of which has its own color gamut. We discuss briefly some of the most common types of television and projector technologies, specifically mentioning those factors that contribute in defining the color gamut of a display device. Chapter 4 presents a survey on gamut mapping algorithms published in the literature. Then, it mentions several ways of evaluating the performance of GMAs. Chapter 5 covers details of our GMAs that we apply in RGB color space. These methods are listed in the previous section as contribution 1 and 2. In chapter 6 we explain our GEA (mentioned as contribution 3) that uses CIELAB color space. Chapter 7 presents details of our proposed GMAs (contribution 4) that work in HSV color space. Finally, we conclude this study in chapter 8 and suggest future work.

CHAPTER 2

Fundamentals of Vision and Basic Color Science

2.1 Light

We live in a world full of electromagnetic radiations emitting either from natural origin or from sources made by humans. The full spectrum of electromagnetic radiation that the sun emits is shown in Fig. 2.1. However, we are only able to see a very narrow band of wavelengths ranging from $380nm$ to $740nm$ to which we commonly refer as visible light.

Light comes from several different sources such as the sun, lasers, candles, tungsten bulbs, fluorescent tubes, light emitting diode (LED), just to name a

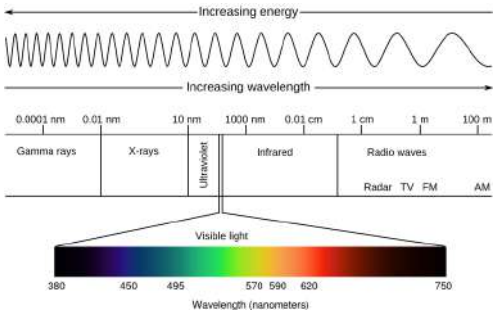


Figure 2.1: Electromagnetic spectrum and visible light. Image from [Bertalmío, 2014].

few. Radio-metrically, a light source can be characterized by its **spectral power distribution (SPD)** which is the physical composition of light as a function of wavelength λ. Light can reach our eye either directly from a light emitting source, or by first hitting a non-luminous object and then reaching to us. In the latter, more common, situation the light is partly absorbed and the rest is reflected depending on the absorption properties of the object. The reflected light that reaches our eye is called *radiance* $E(\lambda)$ which is the product of the SPD of the light source $I(\lambda)$ (*irradiance*) and the spectral *reflectance* $R(\lambda)$ of the object:

$$E(\lambda) = I(\lambda) \times R(\lambda) \tag{2.1}$$

The pictorial representations in Fig. 2.2 and Fig. 2.3 show the irradiance function of some light emitting sources and reflectances of different colored patches, respectively.

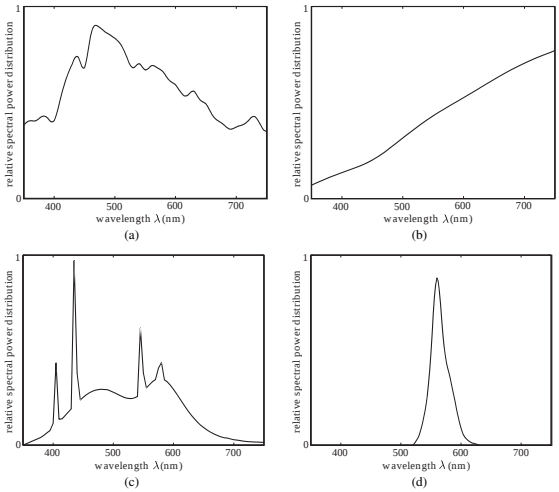


Figure 2.2: Spectral power distributions of some common light emitting sources: (a) Sun, (b) tungsten bulb, (c) fluorescent tube, and (d) LED. Figure from [Lam and Fung, 2008].

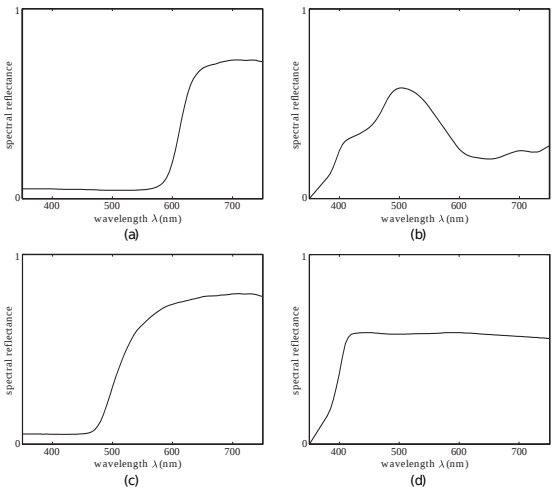


Figure 2.3: Spectral reflectance of various colored patches: (a) red patch, (b) blue patch, (c) yellow patch, and (d) gray patch. Figure from [Lam and Fung, 2008].

2.2 Human Eye Anatomy, Biology, and Vision Science

Once radiance reaches our eyes, it passes through the optics of the eye to the retina and gets translated into nerve signals for the brain to process and interpret the scene. The conversion from real world light to imagery completes by passing through three main processing units of the human visual system: 1) the eye optics, 2) the retina and 3) the thalamus and visual cortex. The schematic diagram of an human eye is illustrated in Fig. 2.4 and the explanation of the functionality of its core components is presented in the following sections.

2.2.1 The Eye Optics

The optical part of the human eye includes cornea, iris, pupil and lens; they all perform non-neural tasks. The cornea is a nearly transparent part of the eye that lets light to enter into the visual system, and two-thirds of the light refraction occurs at the boundary of cornea [Ferwerda, 2001]. At the cornea, 99% of ultra violet rays at wavelength less than $300nm$ are absorbed and therefore it protects the eye periphery from the excessive exposure of short wavelength rays [Packer

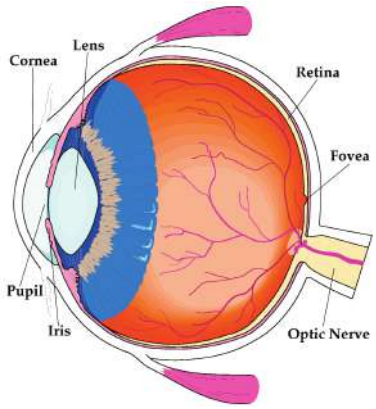


Figure 2.4: Schematic diagram of the human eye. Figure is from [Fairchild, 2013].

and Williams, 2003]. The iris is a distinctive (unique for everyone) colored part of the eye and in the center of the iris there is an opening called pupil. The iris can contract and expand to adjust the size of the pupil in order to let through a controlled amount of light to the other parts of the eye. In optical terms, the pupil serves as the aperture of the eye and the iris acts as the aperture stop. Light comes in through the cornea, gets regulated by the iris and then hits a biconvex transparent lens that performs a minor refraction operation in order to focus the incoming light to the back of the eye onto the retinal surface. Continual adjustment of the lens enables us to change focus and therefore see objects at various distances.

2.2.2 The Retina

The inner layer of neural tissues at the back of the eye is called retina. It serves as an interface between non-neural and neural parts of the eye. The retina has millions of photoreceptors that are responsible to transduce electromagnetic waves into nerve signals. As shown in Fig. 2.5, the retina has several interconnected cell layers, and it may seem somewhat unexpected, the light actually enters from the top layer, passes through the intermediate layers, gets slightly scattered and absorbed, before reaching to the light-sensitive layer of photoreceptors. The human visual system operates over a large range of light levels. It has two types

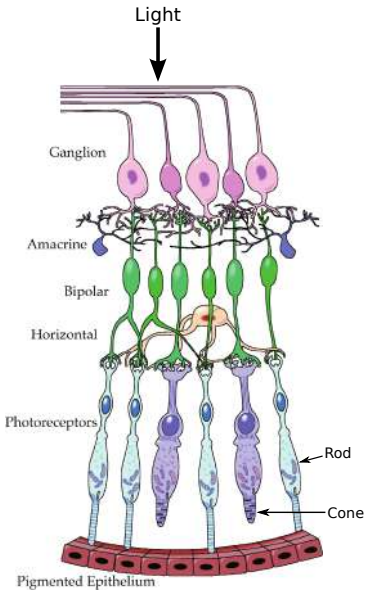


Figure 2.5: Simplified diagram of the human retina. This figure is adapted from [Fairchild, 2013].

of photoreceptors that respond depending on the power of the incoming light:

- **Rods** initiate reliable vision at low light levels in the range 10^{-6} to 10 cd/m^2 (candela per square meter) and register grayscale information, and although remain activated at high luminances, they become saturated. Rods are extremely sensitive to light; they get triggered and produce detectable photocurrent response even in the presence of one photon [Wandell, 1995]. Vision at low illumination levels is called *scotopic* and it is colorless.

- **Cones** have pigments that are typically 100 times less sensitive to light than rods [Kefalov et al., 2003]. Therefore, they only operate in bright conditions (0.01 to 10^8 cd/m^2) and detect fine details and colors. Vision at high light levels is referred as *photopic*. The range of luminance levels (0.01 to 10 cd/m^2) where both rods and cones contribute to the vision are

called *mesopic* levels, for example, in a typical movie theater we experiecne mesopic vision [Bankston, 2005].

Cones are further divided into three varieties: S-cones, M-cones and L-cones, where the capital letters denote *short, medium* and *long* wavelengths, respectively. Historically, Thomas Young in the early 1800s and Hermann Von Helmholtz in the 1850s realized that our eyes have three types of photoreceptors and each one is sensitive to different wavelengths of light. But their trichromatic theory, though mostly correct, was not based on experimental evidence, rather it was based on scientific reasoning only. In the late $19th$ century, Arthur König was the first to measure experimentally the spectral absorption function of each type of photoreceptor [König and Dieterich, 1892]. The curves in Fig. 2.6 show the region on the visible spectrum to which each of the three cone types are sensitive to. S-cones are sensitive to short wavelengths, with a peak at $420nm$. M-cones operate in the middle wavelengths and their peak sensitivity is $533nm$. And L-cones provide response to long wavelengths, with a peak sensitivity of $584nm$. Although these bell-shaped sensitivity curves are quite broad, each type of photoreceptor peaks at a distinct wavelength and these three distinct wavelength values, respectively, correspond to monochromatic violet, blue-green and yellow-green light (or *loosely* associate to blue, green and red light [Hubel, 1995]).

Under fixed viewing conditions, the cones response to the incident light can be modeled as a vector with three components, called **tristimulus values**. It is

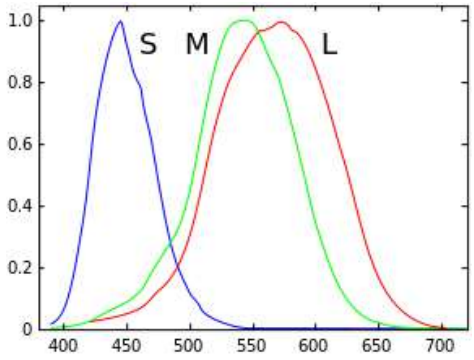


Figure 2.6: Cones sensitivity functions. Figure from [Wikipedia, 2007].

computed by taking the integral of the radiance $E(\lambda)$ times the cone's sensitivity, both as a function of wavelength, over the visible spectrum.

$$\begin{aligned} L &= \int_{380}^{740} l(\lambda)E(\lambda)d\lambda \\ M &= \int_{380}^{740} m(\lambda)E(\lambda)d\lambda \\ S &= \int_{380}^{740} s(\lambda)E(\lambda)d\lambda \end{aligned} \tag{2.2}$$

where $l(\lambda), m(\lambda)$ and $s(\lambda)$ denote sensitivity functions of long, medium and short cones, respectively.

Compared to 120 million rods, there are only 6 million cones. Cones are mostly concentrated in the central region of the retina named as *fovea*, that serves to our finest spatial and color vision. The fovea covers an area that subtends about two degrees of visual angle in the central field of vision. In order to see an object in sharp focus, we move our eyes to make the light rays coming from it to fall onto the fovea. The distribution of photoreceptors in Fig. 2.7 shows that the density of cones at the center of the fovea is really high (and there are virtually no rods in this part) and as we move away from the fovea, the concentration of cones decreases rapidly and therefore the visual acuity becomes inferior. In the same figure, it can also be seen that there is a region in the retina called blind spot where photoreceptors are not present. However, we do not notice this because the

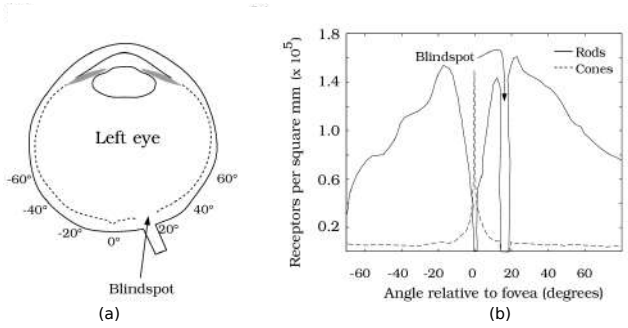


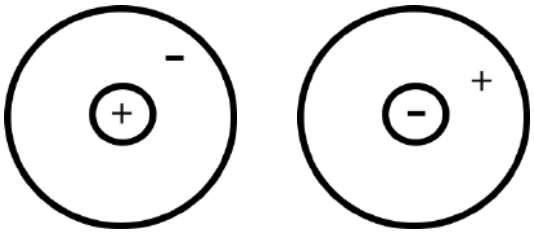
Figure 2.7: Distribution of photoreceptors across the human retina: (a) visual angles relative to the fovea, (b) concentration of cones and rods in the fovea. Figure is taken from [Wandell, 1995].

human visual system fills this region with information of the observed scene.

In the illustration of retina (Fig. 2.5), the bipolar cells on one end are connected with the photoreceptors and on the other end they have connection to the ganglion cells which form the optic nerve. What is important to note is that many photoreceptors provide input to multiple bipolar cells, and many bipolar cells provide input to multiple ganglion cells. This straight line of retinal processing ignores the horizontal cells and the amacrine cells, both of which provide lateral connections: the former among photoreceptors and bipolar cells, and the latter among bipolar and ganglion cells. The human retina is a legitimate part of the brain and although researchers have spent enormous amount of time to understand the specific processing that occurs at each type of cell, it is still not fully understood.

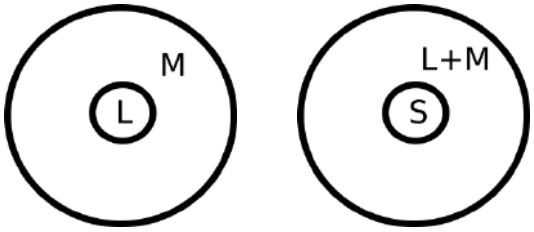
The processing of visual information within the retina is significantly complex. The output signal of retina that ganglion cells provide to the optic nerve is the result of a series of complex operations performed by the layers of neural cells within the retina. The retina has roughly 126 million photoreceptors, but only 1 million optic nerve fibers leave the eye; this indicates the occurrence of some sort of data compression. Functionally, the retina is organized in the form of receptive fields that are the basic units of visual coding [Hartline, 1940]. Each ganglion cell receive inputs from multiple photoreceptors from a specific area. Receptive fields have a very specific structure called center-surround: a central disk, the "center," and a concentric ring, the "surround," each responds in opposite manner to light [Kuffler, 1953].

There are two types of receptive fields: a) on-center field, and b) off-center field. Both of these configurations are shown in Fig. 2.8. An on-center ganglion cell fires rapidly when the center of its receptive field is stimulated with light, whereas it is inhibited when the surround is exposed to light; the response of such cell is highest for a white spot (filling the center of the receptive field) on a black background. [Hubel, 1995] mentions that *"if you listen to the discharges of such a cell over a loudspeaker, you will first hear spontaneous firing, perhaps an occasional click, and then, when the light goes on, you will hear a barrage of impulses that sounds like a machine gun firing."* An off-center cell has just the opposite behavior: firing when the surround is exposed to light, and inhibiting when the center is stimulated with light; the reponse of an off-center cell is maximum for a black spot (filling the center of the receptive field) on a white background. Both on-center and off-center ganglion cells are equally common in



(a) (b)

Figure 2.8: Typical center-surround receptive fields: (a) on-center and (b) off-center.



(a) (b)

Figure 2.9: Spectral-opponency in receptive fields of ganglion cells: (a) red-green opponency (b) yellow-blue opponency.

the retina and the receptive fields these cells form can be of different sizes and overlapped [Wandell, 1995], [Hubel, 1995]. The ganglion cells do not provide any response to a uniform stimulus when the negative and positive areas of the receptive fields (shown in Fig. 2.8) are balanced. This behavior of spatial opponency enables ganglion cells to perform data compression.

In addition to spatial opponency, ganglion cells behave in spectrally-opponent manner as well [Hurvich and Jameson, 1957], [Ferwerda, 2001]. As illustrated in Fig. 2.9, the red-green opponent cells take their input from the antagonism of L-cones and M-cones, whereas the yellow-blue opponent cells are excited by the S-cones and are inhibited by the combined response of L-cones and M-cones [Mollon and Cavonius, 1987]. The electrical signal leaving the retina via the optic nerve is basically the visual space representation of the real world scene, and this signal does not have a point-to-point relationship with photoreceptors,

but instead it carries information about the local relationships.

2.2.3 Thalamus and Visual Cortex

The electrical signal then reaches the thalamus which is a way station for sensory pathways. In the thalamus lies a sub-nucleus that is specifically related to vision and is called the lateral geniculate nucleus (LGN). This is the place where local relationships of the signal are maintained because ganglion cells terminate at this point after making one-to-one correspondences with LGN cells [Hubel, 1995]. Hence, LGN cells have the same receptive fields as those of ganglion cells. Finally, the signal arrives in the primary visual cortex where in the first level of processing, the visual cortex simply maps the input from the center-surround receptive fields. The second level of processing involves arranging these center-surround receptive fields in rows in order to find lines or edges. The visual cortex has several subdivisions that perform further processing to analyze the signal's features such as motion, form, orientation, color, depth perception, and many other cerebral modalities [Wandell, 1995]. Movement, form and depth perceptions are processed near the parietal lobe. Detail and color are processed near the temporal lobe. Eventually, all the processed information from the visual cortex's subdivisions are recombined into a coherent image that reaches our conscious recognition somewhere in the brain that has yet to be identified.

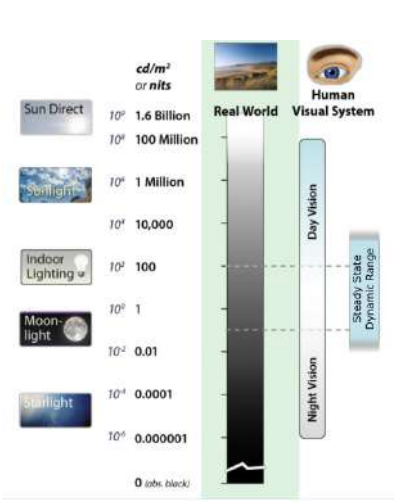
2.3 Characteristics of Human Vision

In the following sections, some of the basic global and local perceptual properties of the human visual system are discussed. All these sections are going to show how our perception of the light coming from an object is heavily influenced by factors such as the light coming from other objects in the scene and the ambient lighting conditions.

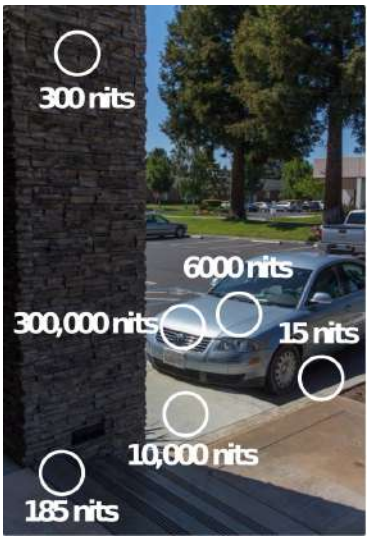
2.3.1 Effects of Ambient Illumination on Perception

Visual Adaptation

In the real world a large range of luminance levels are present, going from dim starlight to bright sunlight as it can be seen in Fig. 2.10a. During the course of the day, typically we experience scenes that, at the same time, have both very bright and very dark regions. For instance, in Fig. 2.10b, measured luminance levels of several spots are shown and we can see that below the car, the intensity is 15 nits (nit is a convenient name for the light unit cd/m^2), whereas on the hood of the car, we have specular highlights and reflections with luminance of $300,000$ nits. Even though our visual system is capable of perceiving a huge range of illumination levels (on the orders of 10^{10}), the issue is that it cannot perceive bright and dark regions of a scene at the same time. In such situation, the mechanism of *adaptation* starts which allows us to see only a *small part* of the full dynamic range present in a scene. The dynamic range of a scene is defined as



(a)



(b)

Figure 2.10: Visual Adaptation. (a) Dynamic range of real world and steady state dynamic range of human vision. (b) Example of dynamic range present in a scene.

a ratio of the brightest and the darkest intensity level present in it.

The adaptation process is controlled by the total retinal illumination and it is closely related to the mean luminance in the field of view. The process of adaptation is slow in dark conditions (often taking several minutes), whereas one could adapt in a matter of seconds in bright conditions [Kalloniatis and Luu, 2007]. The adapted small part of the full dynamic range is called the steady-state dynamic range. Thus, with this efficient phenomenon of the photoreceptors and the brain, we get to operate over the full dynamic range present in the real world. The response of photoreceptors to the incident light can be described using the Naka-Rushton (or Michaelis-Menten) equation:

$$\frac{R}{R_{max}} = \frac{Y^n}{Y^n + \sigma^n} \tag{2.3}$$

where Y denotes the input luminance and σ is the semi-saturation constant which forms the inflection point of the sigmoidal curve. R is the output response of the photoreceptor. Figure 2.11 presents photoreceptor response curves for several semi-saturation values. One may conclude from the same figure that the steady state dynamic range (in logarithmic base 10 scale) is 4 log units. However, in natural scenes, this range is much lower due to various factors such as light scattering in the eye, limited neurons in the primary visual pathway, etc. Therefore, the problem is that the boundary of the steady state dynamic range is not precisely defined. Different studies reported different steady state dynamic ranges, for example, 2 log units in [Myers, 2003], 3 log units in [Purves and Lotto, 2003],

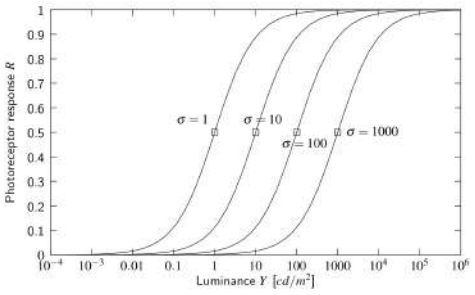
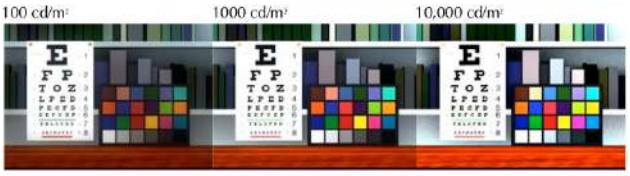


Figure 2.11: Photoreceptor response curves for several semi-saturation values; the x-axis is in logarithmic scale (base 10) and the y-axis is linear. Figure is taken from [Mantiuk, 2007].

(a)



(b)

Figure 2.12: (a) The Hunt effect: saturation decreases with decreasing illumination. (b) The Stevens effect: contrast decreases with diminishing illumination. Image from [Fairchild, 2004]

3.5 log units in [Normann and Perlman, 1979], and 3.6 orders of magnitude in [Kunkel and Reinhard, 2010].

Hunt Effect and Stevens Effect

Our perception of saturation depends on ambient lighting conditions: colorfulness of objects decreases when the illumination decreases. This phenomenon is known as the Hunt effect and it is illustrated in Fig. 2.12a. The Hunt effect indicates that a color present in dim viewing conditions would perceptually match to a much less saturated color under brighter lighting conditions. Figure 2.12b shows the effect of diminishing illumination on our contrast perception: light colors look darker and dark colors seem lighter. This is called the Stevens effect.

Bezold-Brücke (B-B) Effect

The B-B effect [Fairchild, 2013], [Pridmore, 1999] describes that when two stimuli of the same wavelength but having different luminance are viewed simultaneously or successively, they will appear to us as of having different hues. At high luminance levels, stimuli having wavelength below $500nm$ will appear bluer and

Figure 2.13: The simultaneous contrast effect. Image from [Bertalmío, 2014].

stimuli of wavelength above $500nm$ will appear yellower except for three so-called invariant hues: blue ($476nm$), green ($507nm$) and yellow ($574nm$) [Purdy, 1931].

2.3.2 Local Behavior of Visual System

In our daily life we rarely see colors in isolation. Our perception of a color is strongly influenced by other colors present in its vicinity. In this section we will review several phenomena related to the spatial behavior of the human vision.

Simultaneous Contrast

In Fig. 2.13, the intermediate squares have exactly the same shade of gray but the one with the dark immediate surround seems lighter than the other center square. Another way to test this phenomenon is by taking a sheet of paper that would otherwise seem white to us, but when it is used to shield the eyes while looking directly at the bright sky, it can appear totally black [Poynton, 2012].

Mach Bands

Mach bands is another striking phenomenon demonstrating that the human perception is not a simple function of luminance. When two solid gray stripes of slightly different shades are put together, our visual system performs undershoot and overshoot at the common edge of the stripes. Fig. 2.14 shows one such example where the visual system exaggerates contrast at the boundaries of the stripes of constant luminance.

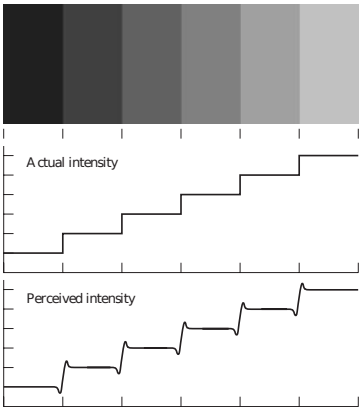


Figure 2.14: The simultaneous contrast effect. Image from [Gonzalez and Woods, 2006].

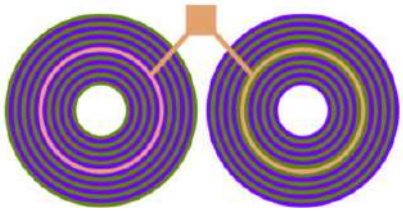


Figure 2.15: Chromatic induction: the centrally located test-rings are identical, but they appear different due to the influence of surrounding colors. Figure from [Monnier, 2008].

Chromatic Induction

Figure 2.15 shows the phenomenon of what we call chromatic induction: the light reaching our eyes from the inner test-rings in both sets of concentric rings is the same, yet they appear to us as of having different colors, because of the influence from their surrounding colors.

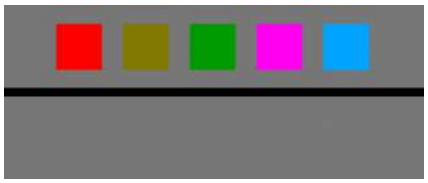


Figure 2.16: Example of Helmholtz-Kohlrausch effect. In the top row, each color square and the background has the same luminance and yet they appear different. Figure from [Wikipedia, 2017a].

Helmholtz-Kohlrausch Effect

Helmholtz-Kohlrausch Effect states that colored stimuli appears brighter to human observers than achromatic stimuli of the same luminance. An illustration of Helmholtz-Kohlrausch effect is shown in Fig. 2.16 where all the colored patches and the gray background in the top row has the same luminance but they appear to us differently. Furthermore, it is also noticeable that some colors (hues) do seem brighter than the others, but they all are perceived significantly brighter than the gray background. If the image in the top row is converted into grayscale we obtain an image that has the same shade of gray as illustrated in the bottom row of Fig. 2.16.

2.3.3 Human Color Constancy

As explained in section 2.1, the light $E(\lambda)$ that reaches our eyes is basically a product of the spectral power distribution of the illuminant $I(\lambda)$ and the reflectance property $R(\lambda)$ of the object. As illustrated in Fig. 2.17, when the same tomato is observed under three different scene illuminations, in theory the color of tomato should appear to us differently because in all three cases the object (tomato) is reflecting different composition of wavelengths to our eyes. But in reality we perceive the tomato's color nearly the same, i.e. the red. It implies that the human visual system somehow manages to 'discount' the effect of the illuminant and therefore the color of the object appears relatively same under different illuminations; this perceptual feature of our visual system is known as color constancy.

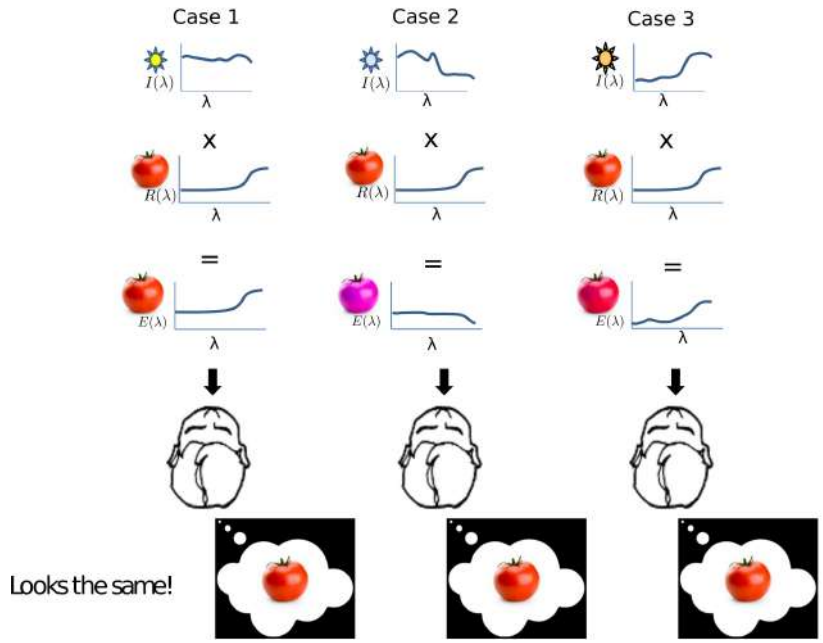


Figure 2.17: An example depicting the phenomenon of color constancy. Figure is adapted from [Brown, 2016].

Retinex Theory of Color Vision

In this section we briefly describe the main concepts of the Retinex theory [Land and McCann, 1971], [Land, 1977], a popular model of how the human visual system perceives colors. The Retinex model was developed starting from two main assumptions: a) that our perception of the color of an object is not completely determined by the light coming from it; instead it is heavily influenced by the light coming from other objects in the scene, and b) that the human visual system operates with three retinal-cortical systems that are responsible to process independently the corresponding short, medium, and long wavelength bands of the visible light and form three separate images with lightness L information that determine, by superposition, the perception of what we call colors.

In order to develop a method to compute lightness L, Land summerizes in [Land, 1977] a color-matching test for which the experimental setup is shown

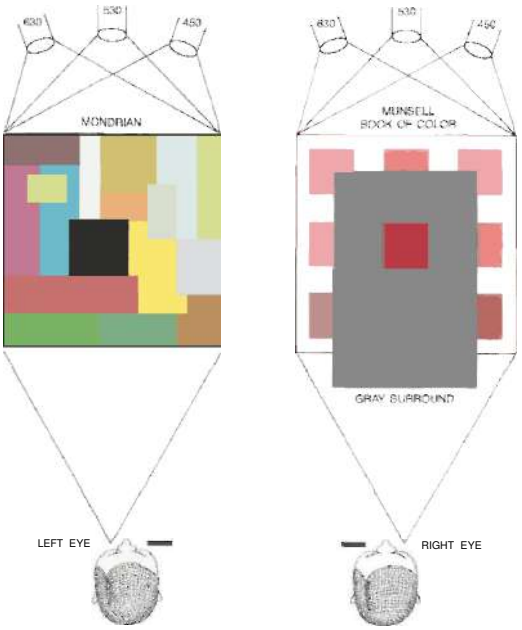


Figure 2.18: Color matching experiment of Retinex theory of color vision. Figure from [Land, 1977].

in Fig. 2.18. On the left side, we can see 17 colored sheets of paper arranged to form a panel to which Land calls the "Mondrian". To illuminate any color area of the Mondrian, three spectral lights with wavelengths $630nm$ for red, $530nm$ for green and $450nm$ for blue were used that can be mixed to any desired proportion to generate an experimental illuminant. On the right side, chips from *the Munsell Book of Color* were illuminated by a constant white illuminant obtained by adjusting the same set of spectral lights. The format of the experiment was to make observers use their left eye to view the Mondrian and their right eye to view chips.

In the experiment, the observer was asked to look at a given color of the Mondrian and find a chip from the Munsell book that best matches its color. In each run of the test, the experimental illuminant was adjusted such that the radiant energy coming to the observer's left eye from any colored sheet of the Mondrian is the same regardless of its reflectance: gray, yellow, red, green, and blue areas

were sending the same radiance and hence a photometer would have declared all of them identical. But in each and every case, the observer was able to match the Mondrian sheet to the closest Munsell chip, i.e. blue sheet to blue chip, yellow sheet to yellow chip, etc. The experiments showed that, in each matching pair, Mondrian sheet and Munsell chip had different radiances but the same *(scaled) integrated reflectances*, which is defined, for each waveband (short, medium and long) as a fraction: its numerator is the integral of the radiance of the patch over a given waveband, and its denominator is the integral (over the same waveband) of the radiance of a white object under the same scene illuminant.

The following text is reproduced from our article [Zamir et al., 2014]. " The *scaling* is a non-linear function that relates reflectance to lightness sensation. But this implies that in order to perceive the color of an object somehow our visual system is comparing the light coming from the object with the light coming from a reference white, and Land wondered how we are able to find this reference white *'in the unevenly light world without reference sheets of white paper'* [Land, 1977]. The sensation of white will be generated by the area of maximum radiance in all three bands (this is the von Kries' model or 'white-patch' assumption, although Land does not cite it); this area could be used as reference, but Land did not know how our visual system could *'ascertain the reflectance of an area without in effect placing a comparison standard next to the area'* [Land, 1977]. The solution he proposed consisted of comparing far-away points through paths: the ratio of the values at the two end-points of the path can be computed as the sequential product of the ratios of each point of the path with the following point. The Retinex algorithm consists of assigning, for each point and each waveband (long, middle, short), an estimate reflectance obtained as the average of the sequential products obtained on many paths, all ending in the given point. Land thought that this was a plausible explanation of how our visual system estimates reflectances but he did not want to venture where exactly this type of operations were being carried out, in the retina or at the cortex; therefore he chose the name 'Retinex' for his approach. The Retinex algorithm is directly applied to digital images in a straightforward way, where the pixels will be the points and the three chromatic channels play the role of the wavebands."

2.4 Colorimetry

Colorimetry is the branch of color science that deals with assigning numbers to the colors of a physical stimulus in such a manner that another stimuli with the same physical properties will have a complete color-match with it, under the same observing conditions [Wyszecki and Stiles, 1982].

2.4.1 Color Matching Experiment

Light is expressed by its spectral power distribution. In the *spectral reproduction* approach, a color is reproduced by directly reproducing its SPD. For example, one could divide the visible spectrum ($380nm$ to $740nm$) into 37 bands of $10nm$ each and then use these 37 components for coding each pixel in an image. While such system could be suitable to reproduce a single color or a few colors, it becomes highly impractical for image coding [Poynton, 2012].

The normal human eye contains three types of cone cells that respond to incident light to produce the sensation of color. This allows to generate any perceivable color by a proper mixture of any three given lights (as long as they are colorimetrically independent, i.e. that the mixing of two of them does not produce the same color as the remaining one), in what is known as the trichromacy property:

$$C = R_C(\boldsymbol{R}) + G_C(\boldsymbol{G}) + B_C(\boldsymbol{B}). \tag{2.4}$$

The way to interpret this equation is that any color C can be matched by mixing R_C units of the $\boldsymbol{R}$ primary, G_C units of the $\boldsymbol{G}$ primary and B_C units of the $\boldsymbol{B}$ primary; this is also known as Grassmann's first law of additive color mixture. The terms, $\boldsymbol{R}$, $\boldsymbol{G}$ and $\boldsymbol{B}$, define the particular primary set, which implies that for a different set of primaries, different mixing proportions will be needed to make a match. Two lights that produce the same cone response triplet (L, M, S from Eq. (2.2)) are perceived as having the same color, even if they have a different power spectrum; such lights are called metamers.

Cone sensitivity functions were not available in the 1920s. Owing to the trichromatic nature of the human vision, W. David Wright [Wright, 1929] and, independently, John Guild [Guild, 1932] conducted the same color matching experiment for which the setup is shown in Fig. 2.19. The observers were asked

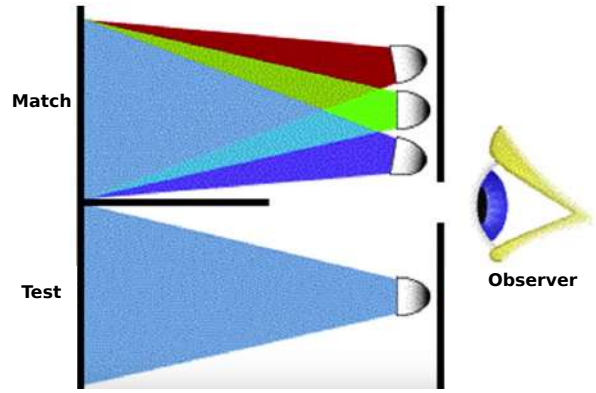


Figure 2.19: Color matching experiment.

to see through a circular split screen. On one half of the screen a test color (monochromatic light) was shown. Observers were given the control of three monochromatic lights (red = $700nm$, green = $546.1nm$, blue = $435.8nm$) and by varying the intensity of these lights, observers had to make a color on the other half of the screen that would match the test color. This process was repeated for each monochromatic spectral test light. Figure 2.20 shows, average over all observers, the amount of primary colors ($\overline{r}(\lambda)$ for the red, $\overline{g}(\lambda)$ for the green, and $\overline{b}(\lambda)$ for the blue) needed to match the test light of a specific wavelength of the visible spectrum. For example, by adding 0.16 units of the red and green lights, one could match the spectral light of wavelength $575nm$ (which we would perceive as yellow color). Note that in the same figure the red curve has negative values, because it turned out that it was not possible to match test stimulus in that region by simply mixing the three primary lights. Therefore, in this region, a certain amount of red had to be added into the test spectral color in order to obtain a match by mixing the blue and green primary lights.

Once we have the color matching functions, we can calculate the (R, G, B) tristimulus values for any light with spectral power distribution $E(\lambda)$ as:

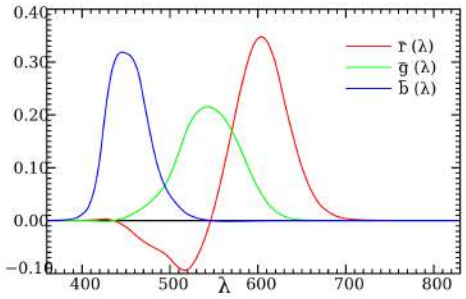


Figure 2.20: Color matching functions $\overline{r}(\lambda)$, $\overline{g}(\lambda)$ and $\overline{b}(\lambda)$. Figure is taken from [Wikipedia, 2007].

$$
\begin{aligned}
R &= \int_{380}^{740} \overline{r}(\lambda) E(\lambda) d\lambda \\
G &= \int_{380}^{740} \overline{g}(\lambda) E(\lambda) d\lambda \\
B &= \int_{380}^{740} \overline{b}(\lambda) E(\lambda) d\lambda
\end{aligned}
\tag{2.5}
$$

2.4.2 The First Standard Color Spaces

As it is noted in the previous section, any given color can be matched with a set of three primary colors. A color space is a method to describe and visualize colors. In 1931, CIE defined a system to classify colors based on the mean results obtained from Wright and Guild's data [Fairman et al., 1997]. Two sets of color matching functions were proposed: one is known as the CIE RGB colorimetry system which uses the same color matching functions as mentioned in Eq. 2.5, and the other is the CIE XYZ system which is still to this day a very popular method for specifying colors in the industry because this system is device independent, meaning that a set of parameters used to specify a color will produce the same color on whatever equipment they are applied [Ford and Roberts, 1998]. There were two main objectives behind defining the CIE XYZ color matching functions $\overline{x}(\lambda)$, $\overline{y}(\lambda)$ and $\overline{z}(\lambda)$: first, to eliminate the negative values from the functions $\overline{r}(\lambda)$, $\overline{g}(\lambda)$ and $\overline{b}(\lambda)$ (see Fig. 2.20), and second, to deliberately match one of

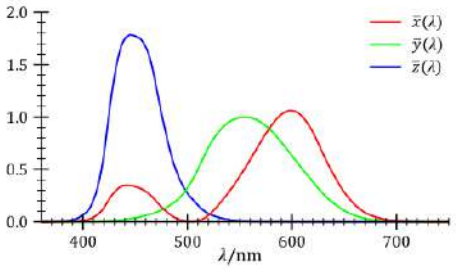


Figure 2.21: Color matching functions $\overline{x}(\lambda)$, $\overline{y}(\lambda)$ and $\overline{z}(\lambda)$. Image from [Wikipedia, 2007].

the color matching functions to the 1924 CIE standard luminosity function $V(\lambda)$, which describes the intensity of light seen as coming from a stimulus (however, not in a perceptually uniform way). Therefore, integrating the SPD of a color stimulus with $\overline{y}(\lambda)$ would correspond to its luminance.

The CIE XYZ color matching functions are shown in Fig. 2.21, and for any given color stimulus with spectral distribution $E(\lambda)$, the (X, Y, Z) tristimulus values can be computed in the same fashion as the (R, G, B) tristimulus values:

$$\begin{aligned} X &= \int_{380}^{740} \overline{x}(\lambda) E(\lambda) d\lambda \\ Y &= \int_{380}^{740} \overline{y}(\lambda) E(\lambda) d\lambda \\ Z &= \int_{380}^{740} \overline{z}(\lambda) E(\lambda) d\lambda \end{aligned} \tag{2.6}$$

CIE xyY Color Space and CIE Chromaticity Diagram

Numerically, it is possible to match a wide range of colors by mixing three properly chosen color primaries, resulting in a three-dimensional figure where each point represents a color. The sensation of a given color can be described with three main quantities:

- its hue, which is what we normally refer as perceived color (red, green, yellow, and so on).

- its brightness, which is an attribute according to which the color appears to have more or less light intensity.
- its saturation, that defines how pure the color is relative to its own brightness.

When describing colors it is convenient, for better conceptual understanding and clear visualization, to decouple luminance from the hue-saturation pair that is referred as **chromaticity** using the standardized normalization procedure defined by the CIE:

$$\begin{aligned} x &= \frac{X}{X+Y+Z} \\ y &= \frac{Y}{X+Y+Z} \\ Y &= Y \end{aligned} \tag{2.7}$$

Figure 2.22 shows the standard CIE xy chromaticity diagram, where the tongue-shaped region corresponds to the chromaticities of all the colors a standard observer can perceive. The horse-shoe shaped curve at the upper boundary

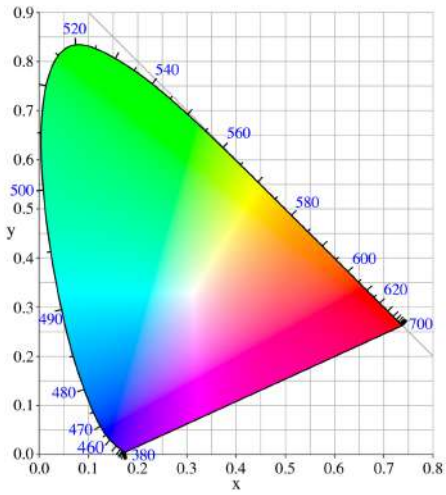


Figure 2.22: CIE xy chromaticity diagram. Image is from [Wikipedia, 2007].

represents monochromatic colors with their wavelengths labeled, and it is called spectral locus [Sharma, 2002]. The line at the lower boundary is called the line of purples and it corresponds to mixtures of blue and red lights from the two ends of the spectral locus. These are not spectral colors, which is why there is not any wavelength labeled with them.

A perfect white light with completely flat power spectrum lies at $x = y = 1/3$ where the saturation is 0%. Such white light is called equal energy white, and we will see in detail some other white lights in section 2.4.5. A pure monochromatic light has 100% saturation, and as we mix it with white, its chromaticity moves inward and therefore the saturation gets reduced. While the trichromacy property states that any color can be expressed as a linear combination of a given set of three primaries, it is important to note that any color that a device can create will always lie within the triangle determined by the chromaticities of its primaries. Furthermore, due to the convex shape of the spectral locus, any such triangle will not be able to cover the full chromaticty diagram. As a consequence tri-chromatic systems are unable to reproduce many colors that we can perceive.

2.4.3 Perceptually-uniform Color Spaces

Although CIE XYZ and CIE xyY color spaces are easy to implement [Poynton, 2012], they are perceptually non-linear. For example, the distance between two colors in XYZ color space (or in xyY color space) does not correspond to the perceptual difference between those two colors. Figure 2.23 shows a plot where colors within the contour of each ellipse are perceptually indistinguishable when compared with the color at the center of the ellipse. It is evident that the perceptual differences of the same magnitude has produced ellipses of different sizes, therefore affirming that these color spaces are perceptually non-uniform.

Several researchers have devoted their efforts to develop color spaces that have perceptual uniformity because it is an extremely important feature as it allows to define error tolerances in color reproduction systems [Bertalmío, 2014]. In 1976 the CIE defined two new color spaces: CIE L*a*b* (abbreviated CIELAB) and CIE L*u*v* (abbreviated CIELUV). These spaces are designed to be linear with visual perception. In both color spaces, L* represents lightness that is the standard estimate of the perceptual response to the luminance value Y, whereas the other two components denote chromaticity (albeit defined differently). u* and v* are defined so that just noticeably different colors are roughly equi-spaced. Whereas,

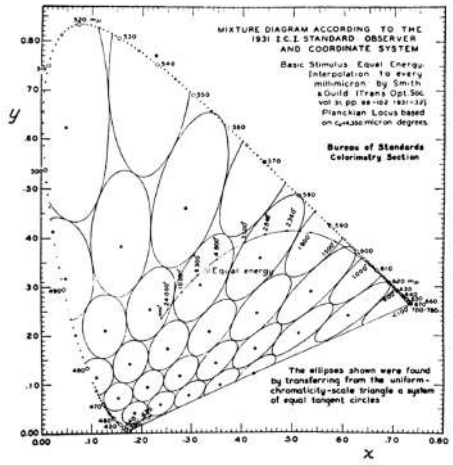


Figure 2.23: Ellipses representing the perceptual difference of the same magnitude. Figure from [Judd, 1979].

a* and b* are chosen such that the euclidean distance between two points match the perceptual difference between the colors corresponding to those points. The lightness coordinate L*, same for both spaces, is computed as:

$$L^* = \begin{cases} \left(\frac{29}{3}\right)^3 \frac{Y}{Y_n} & \text{if} \quad \frac{Y}{Y_n} \leq 0.008856 \\ \\ 116 \left(\frac{Y}{Y_n}\right)^{\frac{1}{3}} - 16 & otherwise \end{cases} \tag{2.8}$$

and the chromaticities for CIELAB are computed as:

$$\begin{aligned} a^* &= 500 \left(f\left(\frac{X}{X_n}\right) - f\left(\frac{Y}{Y_n}\right) \right) \\ b^* &= 200 \left(f\left(\frac{Y}{Y_n}\right) - f\left(\frac{Z}{Z_n}\right) \right) \end{aligned} \tag{2.9}$$

where (X_n, Y_n, Z_n) denote the tristimulus values of a reference white, more commonly referred as the "white point" which is the brightest point in the field

of view (more detail on this topic will be covered in section 2.4.5). The function $f(\cdot)$ is defined as

$$f(x) = \begin{cases} x^{\frac{1}{3}} & \text{if} \quad x > 0.008856 \\ \\ \frac{1}{3}\left(\frac{29}{6}\right)^2 x + \frac{4}{29} & otherwise \end{cases} \tag{2.10}$$

Whereas, the chromaticities for CIELUV are calculated as:

$$\begin{aligned} u^* &= 13L^*(u' - u'_n) \\ v^* &= 13L^*(v' - v'_n) \end{aligned} \tag{2.11}$$

where u'_n and v'_n are the chromaticity coordinates of the reference white and

$$\begin{aligned} u' &= \frac{4X}{X + 15Y + 3Z} \\ v' &= \frac{9Y}{X + 15Y + 3Z} \end{aligned} \tag{2.12}$$

From the CIELAB model, we can obtain another color space known as CIELCh (L for lightness, C for chroma, and h for hue). In this system, by using the Cartesian [a*, b*] pair, a polar-coordinate pair of chroma and hue can be defined as:

$$\begin{aligned} C^*_{ab} &= \sqrt{a^{*2} + b^{*2}} \\ h_{ab} &= tan^{-1}\left(\frac{b^*}{a^*}\right) \end{aligned} \tag{2.13}$$

Note that the CIE has not defined saturation term in the CIELCh color space. But by converting [u*, v*] pair into polar coordinates as:

$$\begin{aligned} C^*_{uv} &= \sqrt{u^{*2} + v^{*2}} \\ h_{uv} &= tan^{-1}\left(\frac{v^*}{u^*}\right), \end{aligned} \tag{2.14}$$

one can measure the psychometric saturation term as [Ford and Roberts, 1998]:

$$s_{uv} = \frac{C^*}{L^*} \tag{2.15}$$

Limitations of CIELAB, and CIELUV

Even though the CIELAB and CIELUV color spaces are more perceptual than tristimulus color spaces, they are not perfectly perceptual. CIELAB suffers from cross-contamination in some of its parts [Morovič, 2008]. For instance, in the blue region, making changes only in one attribute (such as hue) causes perceptual change in another attribute (such as saturation). CIELUV also have poor hue correlation in some regions. Although in 1976, it was difficult for the CIE to choose one color space over the other, recent studies recommend using CIELAB [Fairchild, 2013].

2.4.4 HSV Color Space

HSV (H for hue, S for saturation, and V for value) is a commonly used color space that is a linear transformation of RGB color model into cylindrical representation. Since RGB is not a perceptually uniform space, this system is also perceptually non-linear and device-dependent. However, its main advantage lies in the extremely intuitive way of specifying colors. This may be the reason that colorists have three sliders to control hue, saturation and luminance in their work stations, because color grading task becomes much simpler and comes naturally to artists in HSV (or alternatively HSL (lightness), HSI (intensity)) color space compared to RGB color space.

2.4.5 Color Temperature and Different Reference White Points

A light source is characterized by its power spectrum and as it is mentioned before, a perfect white light has a flat spectral power distribution known as equal-energy illuminant. However, in reality, white lights never have a completely flat spectrum. The CIE recommended a set of standard illuminants: A for incandescent light, B for sunlight, C for average daylight, D for phases of daylight, E is the equal-energy illuminant, while illuminants F represent fluorescent lamps of various compositions [Wikipedia, 2017b]. It indicates that there is no unique physical or perceptual white. Illuminants A and B have already become obsolete [Poynton, 2012]. The D series of illuminants are the most widely used. Since, many illuminants have, at their core, a heated object (black-body radiator that does not reflect light and emits radiation), it is useful to define illuminants simply by the

temperature in Kelvin (K) degrees. Although the CIE recommends using D65 illuminant that represents natural daylight with a temperature of approximately 6504K, the print industry uses D50, and in photography D55 is commonly used. The broadcast industry uses D65, whereas the cinema industry does not use D series illuminant as we shall see in chapter 3. Many computer displays and televisions have a default setting of 9300K which gives a bluish white, and this temperature is commonly used in consumer and studio displays in Asia due to cultural preferences [Poynton, 2012].

It is convenient to calculate the white point of an D-illuminant in terms of (x, y) coordinates of the CIE 1931 chromaticity diagram, and for this the CIE defined the following formulas:

$$x_D = \begin{cases} \dfrac{-4.6070 \times 10^9}{T^3} + \dfrac{2.9678 \times 10^6}{T^2} + \dfrac{0.09911 \times 10^3}{T} + 0.244063, \\ \qquad \text{for } 4000K \leq T \leq 7000K \\ \dfrac{-2.0064 \times 10^9}{T^3} + \dfrac{1.9018 \times 10^6}{T^2} + \dfrac{0.24748 \times 10^3}{T} + 0.237040, \\ \qquad \text{for } 7000K < T \leq 25000K \end{cases}$$
$$y_D = -3.000x_D^2 + 2.870x_D - 0.275 \tag{2.16}$$

where T represents the correlated color temperature in Kelvin. Chromaticity coordinates of several standard illuminants are summarized in Table 2.1.

Table 2.1: Standard illuminants by the CIE.

	A	C	D_{50}	D_{55}	D_{65}	E (equal-energy)
x	0.44757	0.31006	0.34570	0.33250	0.312727	0.333334
y	0.40745	0.31616	0.35870	0.34760	0.329024	0.333330

CHAPTER 3

Cameras, Displays, and Color Gamuts: The Need for Gamut Mapping

In this chapter we first discuss about the basic components of digital cameras and the common in-camera imaging pipeline that shows how these state of the art cameras are not just light measuring devices but apply several image processing algorithms based on the basic characteristics of visual perception in order to produce visually pleasant imagery. Later we will present various display technologies that are being used in computer monitors, televisions, regular projectors and digital cinema projectors. Particularly, we will explain what are the main factors that contribute in defining the color gamut of a display device. We will show why there is a significant gamut variation among displays and how gamut mapping can enable these devices to use their full color rendering potential. We will present details of standard color gamuts recommended for generating content for cinema, television broadcast and image exchange on the internet.

3.1 Color Image Processing Pipeline of a Digital Camera

Cameras capture incoming light and by processing it through a series of operations they provide us with a displayable image that represents the actual scene. Though the sequence in which these operations are performed may differ from manufacturer to manufacturer and usually they do not make this information public, Fig. 3.1 shows the basic steps that are common to all standard digital

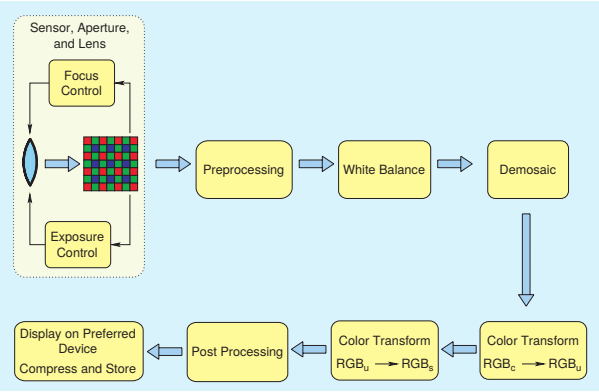


Figure 3.1: Image processing pipeline of digital cameras. Figure from [Ramanath et al., 2005].

cameras. In the following sections we briefly explain the color image processing pipeline of a camera. Particularly, we will discuss at which stage of the pipeline the process of gamut mapping is performed.

3.1.1 Lens, Aperture, Shutter, and Sensor

The incoming light rays first pass through the lens of the camera where they undergo a process of refraction and get focused onto the sensor. In order to regulate the amount of focused light that reaches the sensor, aperture and shutter are used. The aperture of a camera is a small opening that controls the amount of light: narrow aperture allows less light to enter, whereas wide aperture permits more light to pass. The shutter is a window in front of the sensor which basically opens for a specific time in order to expose the sensor to light; if the shutter speed is fast, the sensor will be exposed to light for less time, whereas slow shutter speed will allow light rays to strike the sensor for a longer period of time. Together, the lens, aperture and shutter of the camera assist in capturing the scene radiance onto the sensor.

An image sensor is a semiconductor device that contains an array of cells to capture the light photons and convert them into an electrical signal using the photoelectric effect [Nakamura, 2005]. There are two types of sensors: charged coupled device (CCD) and complementary metal oxide semiconductor (CMOS).

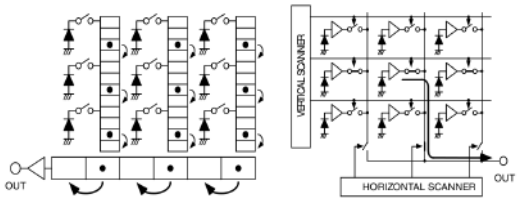


Figure 3.2: Classes of camera sensor. Left: CCD scanning. Right: CMOS charge transfer scheme. Figure is taken from [Nakamura, 2005].

Both families of sensors essentially perform the same task of converting light into voltage but they use different readout technologies. As illustrated on the left side of Fig. 3.2, the CCD sensor transfers the charge vertically, row-by-row, from each cell array to the serial shift register. The shift register then transfers the charge horizontally to be converted into analog voltage at just one output amplifier. On the right side of the same figure, it can be seen that the CMOS sensor has one amplifier at each cell location where charge gets converted directly into analog voltage. Finally, the analog-to-digital converter converts the analog voltage to digital values. At the image sensor, the real world scene gets discretized and therefore the number of cells (or pixels) present in a sensor defines the spatial resolution of the image that it will produce.

The process we have just mentioned would produce a grayscale image because the camera sensor captures only the light intensity, but then how do we obtain the color information? As we have previously noted, the human visual system has three types of cones (short, medium and long) that respond to different wavelengths of light in order to produce the color sensation. A digital camera works on a similar principle: it filters the incoming light rays into triplets of red, green and blue colors in order to produce a color image. To acquire these primary colors, the following two main configurations are in use [Adams et al., 1998]:

- **Three-sensor system:** uses a beam splitter (assembled by combining two dichroic prisms) to separate the incoming light into three different color channels. An image for each channel is then recorded on each sensor. See Fig. 3.3 (left).
- **Color filter array (CFA):** is added to the sensor of a single-sensor camera

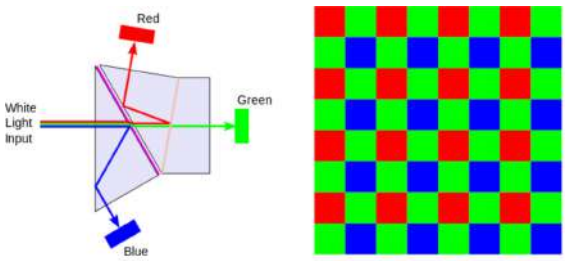


Figure 3.3: Classes of camera. Left: a three-sensor system that a beam splitter camera uses. Right: Bayer CFA that is used in a mosaic single-sensor camera. Image is taken from [Bertalmío, 2014].

either by a dye process or a pigment process. The CFA allows only a particular color (wavelengths) of light to enter into the sensor's cavity. The most popular and a very simple CFA is the Bayer pattern. As shown in the right column of Fig. 3.3, there are two green squares for each red and blue square because of the fact that the human visual system is more sensitive to the green color than to the blue or red color. Depending on the set of dyes used in the CFA, the color capturing capability varies among cameras. Camera makers use different set of dyes for different camera models. One may ask why camera manufacturers do not use dyes that would enable cameras to capture the content with the widest possible gamut. The main reason is the intended market of cameras in which they are going to be used [Bertalmío, 2014]. For example, if the camera content will be displayed on regular TVs (which follow the limited standard gamut of the broadcast industry), then there is no need to capture those colors that will never be reproduced by these display devices.

Virtually all digital cameras have a single sensor and use the CFA, therefore, the remaining steps of the image processing pipeline are explained according to the single-sensor camera system.

3.1.2 Preprocessing

At this stage the camera tries to compensate for a few problems that might have occurred due to several reasons such as light scattering in the lens, defective cells of the sensor, non-linearity introduced by amplifiers and other circuitry, and dark

current that occurs due to thermally-generated electrons and that can increase the output level corresponding to dark scene detail, thus reducing the dynamic range [Ramanath et al., 2005].

Some cameras provide an option to save data in RAW image format without applying any further image processing. RAW image generally represents the sensor values and this image has a camera-specific color space. Professional filmmakers (and photographers) prefer shooting content in RAW format and processing it offline (to make it ready for display) as it provides them more control, flexibility and freedom to make artistic and computational imaging related choices in order to enhance their storytelling [Bertalmío, 2014].

3.1.3 White Balance

Let us recall from section 2.3.3 that our visual system has an essential property that enables us to discount the effect of the illuminant and therefore the color of an object appears constant under different lighting conditions. For example, if we see a white sheet of paper in daylight, or in fluorescent light, or even under tungsten light, we continue to perceive the paper as being white even though the light reaching our eyes is different under different light sources. It indicates that if cameras do not attempt to replicate such behavior of human vision, they will produce images that would never seem correct in color appearance, and we may notice color casts in them. In order to perform white balance (also known as computational color constancy), cameras seek to estimate the scene illuminant and then discount its effects from the captured data [Buchsbaum, 1980], [Cardei and Funt, 1999], [Land, 1977], [Lam and Fung, 2008], [Vazquez-Corral et al., 2012].

3.1.4 Demosaicking

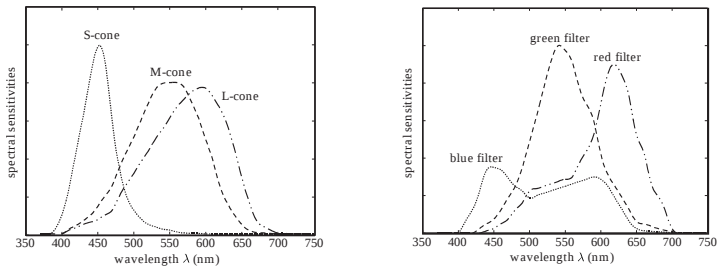
The image sensor produces a mosaic image as shown on the left side in Fig. 3.4. At each pixel of the mosaic image, the information of only one color is present because of the CFA, and the camera has to estimate the other two colors in order to generate a true color image. This approximation of the missing two colors is performed by using a demosaicking algorithm [Ramanath et al., 2002]. Most of the demosaicking algorithms [Hamilton and Adams, 1997] [Hibbard, 1995] [Cok, 1987], due to the configuration of Bayer filter and to obtain a high

Figure 3.4: Example of demosaicking. Left: mosaic image. Right: demosaic image.

resolution image, first interpolate the green channel using bilinear or anisotropic interpolation, and then they use the new green channel to interpolate the red and blue channels. However, these algorithms may introduce artifacts such as false colors and Moiré patterns. To deal with such issues, some approaches use a combination of a demosaicking method and a denoising method. For instance, [Buades et al., 2009] proposed a self-similarity driven demosaicking (SSD) algorithm where they first interpolate a mosaic image to obtain a demosaic image and then apply a non-local image denoising algorithm [Buades et al., 2005] to suppress the interpolation noise. The resultant image is a true color image that is free from artifacts as shown in Fig. 3.4.

3.1.5 In-camera Color Transformations

One may think that, ideally, a camera sensor should have spectral sensitivities identical to the cone sensitivity functions of the human vision, so that when it captures colors, they accurately match with the human perception. But this is not an appropriate option due to several practical reasons. Primarily, as Fig. 3.5a shows, the long and medium wavelength cone curves have a large amount of spectral overlap, and therefore making camera filters that could produce similar curves would be extremely difficult [Bertalmío, 2014] and very expensive [Poynton, 2010b]. Even if a perfect colorimetric filter was easy to develop, there is a good engineering reason not to use one; because reconstructing an RGB image from such sensor will introduce significant noise implications [Poynton, 2012], [Hubel et al., 1997]. An example of a typical digital camera's spectral



(a) (b)

Figure 3.5: Comparison of spectral sensitivities of: (a) the three types of cones in a human eye, and (b) a typical digital camera. Figure from [Lam and Fung, 2008].

sensitivity functions is shown in Fig. 3.5b, and it is clear that they are different compared to the cone sensitivities, and therefore the camera will not perfectly mimic the response of the human eye. While from a signal-to-noise point of view it is not practical to mimic cone responses for image capture, it is important to obtain a colorimetric estimate of the original scene. To achieve this, the image signal must be transformed from the sensor's RGB color space to the CIE XYZ color space (that uses color matching functions of a standard observer [Sharma, 2002]), and eventually into a desired output color space. In case the camera color space is larger than the standard output color space, a gamut mapping procedure is needed to reproduce images without spurious colors and artifacts.

3.1.6 Gamma Correction

Most RGB imaging systems encode images using a non-linear transfer function called *gamma correction*, and there are several reasons for doing this. First, in cathode ray tube (CRT) displays, it was observed that the luminance produced at the face of the screen has a non-linear relation with the device's input voltage: $L = \alpha V^{\gamma}$, where L denotes luminance, V represents the voltage, γ is the exponent of the power function with a value of 2.4, and α is the proportionality coefficient. Therefore, this was the original reason to perform gamma correction in order to pre-compensate the non-linearity of the CRT by adding the inverse exponent of the CRT's power function to the camera signal. Consequently, both non-linearities cancel each other out, and therefore CRT will provide the intended

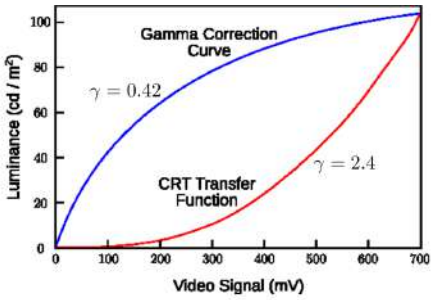


Figure 3.6: Gamma correction on a CRT.

scene luminance to viewers. The plot of gamma correction on a CRT is shown in Fig. 3.6.

Secondly, at that time it was well known that the relationship between perceived lightness and physical luminance is non-linear and it can be approximated as a pure power function with an exponent of 0.42, see Fig. 3.7. This indicates that our visual system is more perceptible to small changes in the dark parts of an image than in the bright parts. Given that over-the-air analog transmission introduced noise in the TV signal, linearly transmitting luminance captured by the camera would make noise very prominent in the dark regions of the image, but transmitting an image encoded with a power function of exponent 0.42 will reduce the visibility of noise. Thanks to the amazing coincidence, the non-linearity of the CRT turned out to be remarkably similar to the inverse of the perceptual luminance-to-lightness function. This indicates that even if we never had the need of gamma correction for physical reasons at the CRT, we still had to develop it for perceptual reasons [Bertalmío, 2014], [Poynton, 2012].

Although today when the TV technology is digital and CRTs are obsolete, we are still using gamma correction. The reason is that each pixel of an image is required to be coded with a limited number of bits. Since we are more sensitive to changes in the dark regions, it is better to allocate more number of bits to code low luminance values and use less number of bits for high luminance values where we are less perceptible to changes of the same amount, and for this the application of gamma correction is inevitable.

In color imaging, the primes on the R'G'B' triplet represent that the values

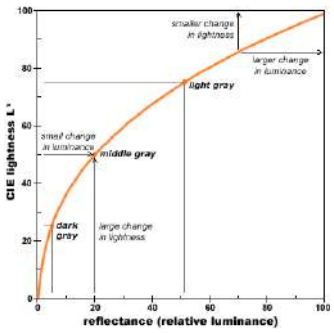


Figure 3.7: Perceived lightness as a function of luminance is approximately a power function of exponent 0.42. Plot from [MacEvoy, 2015].

are non-linearly coded. There exists several standardized color spaces intended for different display technologies: BT.709 for high-definition television (HDTV) [ITU-R, 2002], BT.2020 for ultra high-definition television (UHDTV) [ITU-R, 2012], DCI-P3 for digital cinema [SMPTE, 2011], and sRGB for computer monitors and internet [Stokes et al., 1996]. All of these color spaces use different transfer functions to implement gamma correction as we shall see in more detail later in this chapter.

3.1.7 Post Processing

It is worth mentioning that up until this stage, colors of the image are linearly related to the original scene light. But from this point forward cameras apply non-linear operations in order to add style and aesthetic adjustments in images such that they appear pleasant to viewers. Some commonly used post processing operations include edge enhancement, color-artifact removal, noise removal and tone mapping. All these operations are proprietary and very much camera-specific. Tone mapping is a non-linear operation to modify image values in order to achieve some preferred tone-reproduction; this is not the standard (sRGB) gamma correction which is publicly known. There is a large body of research [Debevec and Malik, 1997], [Nayar and Mitsunaga, 2000], [Farid, 2001], [Grossberg and Nayar, 2003], [Grossberg and Nayar, 2004], [Kim and Pollefeys, 2008], [Vazquez-Corral and Bertalmío, 2015] that aims at estimating this non-linearity so that the radiance of the scene can be recovered.

3.1.8 Compression

A 2-hour long movie with a spatial resolution of 2048×1080 pixels, temporal resolution of 24 frames per second and with 8 bits per color channel requires 1.15 terabytes of memory. Recording movies with such high data rate is well beyond the capabilities of regular cameras. And even if state of the art cameras are able to capture memory intensive content, this amount of data storage is not available in regularly used devices such as computers, mobiles, etc. Therefore, it is necessary to perform compression to reduce the size of the data in order to use memory efficiently. Image compression can be lossless or lossy. To store RAW images in CinemaDNG format or TIFF/EP format, lossless compression is generally used which is based on Huffman coding and run-length encoding. Lossy compression can be performed by using either the discrete cosine transform (DCT) or the wavelet transform. The most commonly used formats based on the DCT are JPEG for still images and MPEG-4/AVC H.264 for videos.

3.2 Color Gamuts

As we have mentioned in section 2.4.2, the physical spectra of an original scene produces sensation in our visual system, and this response can very well be estimated by integrating the spectra with the color matching functions, that provides us with XYZ tristimulus values representing colors of the scene. One may try to directly visualize these tristimulus values using a display device such as TV screen, digital projector, or computer monitor, but this will produce false colors because this is not how display devices reproduce colors. For them to accurately display media, we must convert the XYZ tristimulus values to a standard RGB color space.

A trichromatic device such as a digital projector uses three well chosen color primaries that can be mixed in proper proportions to create different colors. One such illustration of **additive color reproduction** is shown in Fig. 3.8 where it can be seen that each primary has an independent spectral power distribution that are physically summed to generate a new spectra that will cause sensation of the corresponding color in the human visual system.

It is very important to note that each display device has its own RGB color space, and we cannot just use the original RGB values coming from cameras on

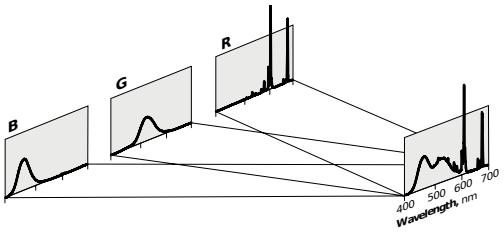


Figure 3.8: Example of Additive reproduction. Figure from [Poynton, 2012].

every type of display devices, rather we have to perform a transformation between the color spaces in order to obtain correct values. This transformation requires the white point and the xy chromaticities of the RGB primaries of both the input image and the target display device. We will see how to transform from the color space of one device to another device later in this chapter, but let us first explain more about color primaries and color gamuts.

The xy coordinates of the three RGB primaries of a display device define its gamut: which is the range of colors it can reproduce. In the xy chromaticity diagram, the set of RGB primaries forms a triangle with vertices in the points (x_R, y_R), (x_G, y_G) and (x_B, y_B). Since colors are created as a linear combination of the three primaries, they will all lie inside this triangle. In Fig. 3.9, we show an example of the color gamut of a device. It is evident that there is no set of three physical primaries that can create a gamut capable of covering the full visible spectrum; in fact no finite set of primaries can cover all colors due to the odd and squiggly shape of the spectral locus. This implies that there are many colors that we are able to see but display devices are not able to reproduce. Hence, a device gamut is a subset of the human vision gamut.

The range of colors (gamut triangle) shown in Fig. 3.9 are only achievable at low values of luminance. But as the luminance increases, the range of colors is reduced, and at the maximum relative luminance of 1, the gamut becomes a single point (being the white point) on the chromaticity diagram; an illustration of this behavior is presented in Fig. 3.10a. Hence, a 2D gamut representation is good for simplicity but a 3D color volume is a more complete way to describe the actual color palette that a device can reproduce. An example of color gamut in 3D is depicted in Fig. 3.10b.

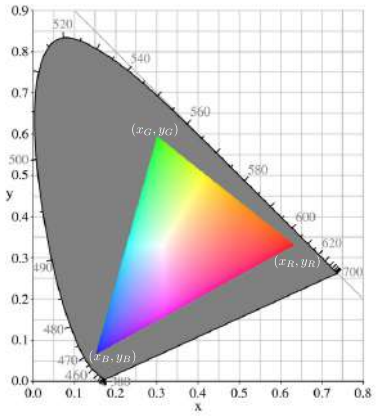
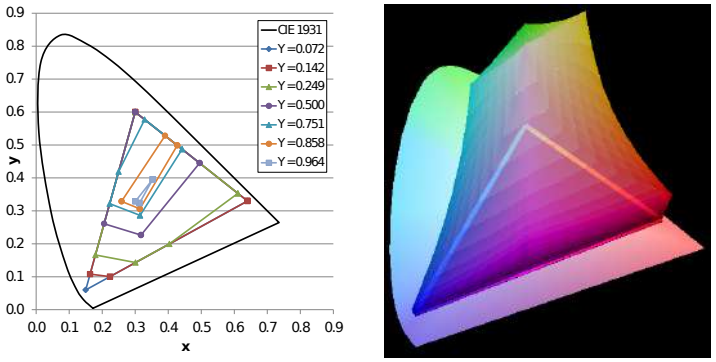


Figure 3.9: An example of color gamut in the CIE xy chromaticity diagram.



(a) (b)

Figure 3.10: (a) An example showing horizontal cross section of 3D color volume at several luminance levels. (b) 3D color gamut.

3.2.1 Pointer's Gamut

While working at Kodak Research, [Pointer, 1980] collected a database of 4089 samples of frequently occurring real surface colors. He made an analysis of this color set and derived a color gamut to which we commonly refer as Pointer's gamut. As shown in Fig. 3.11, Pointer's gamut covers 47.9% area of the visible

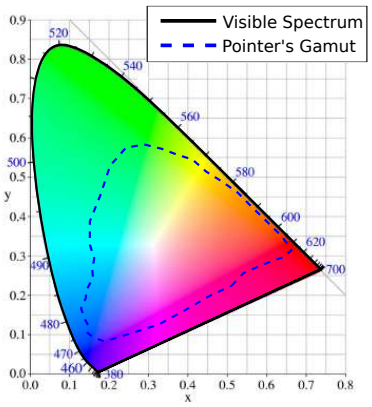


Figure 3.11: Pointer's gamut in the CIE xy chromaticity diagram.

spectrum (when plotted in the CIE xy chromaticity diagram). However, due to the irregular shape of Pointer's gamut, making a display that could reproduce nearly all colors present in the Pointer's gamut would require either three monochromatic RGB primaries or more than three regular primaries. Since many real world surface colors lie within Pointer's gamut, it may be a good reference gamut for display makers to aim for.

3.2.2 Broadcast Standard: BT.709

In 1990, participants in the ITU-R meeting reached to an agreement of standardizing a format for HDTV production and program exchange. The standard is commonly denoted either as Recommendation ITU-R BT.709, or as Rec. 709 [ITU-R, 2002]. We will be using BT.709 abbreviation throughout this document. BT.709 (and other standards that we will explain in the following sections) describes various technical aspects of image display, but we are mainly going to discuss and specify color primaries, white point, peak luminance and transfer function. The BT.709 standard is used in studio production, HD broadcast and cable TV, Blu-ray discs and DVD. The color primaries of BT.709 and its D_{65} white point chromaticities are mentioned in Table 3.1. The BT.709 gamut is visually illustrated in Fig. 3.12, from where it becomes immediately apparent that it covers a small range (33.5%) of colors that we can see, and 69.4% colors of Pointer's

Table 3.1: ITU-R BT.709 primaries and its white point.

	Red	Green	Blue	White, D_{65}
x	0.64	0.30	0.15	0.3127
y	0.33	0.60	0.06	0.3290

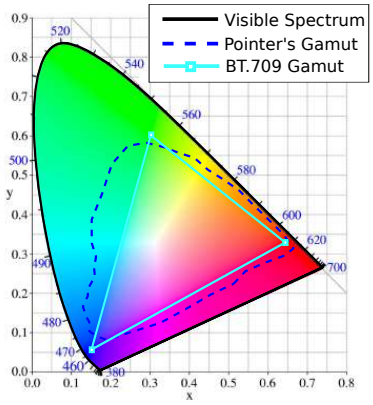


Figure 3.12: BT.709 gamut in the CIE xy chromaticity diagram.

gamut. Currently, HDTV content is mastered with peak luminance of 100 nits and it has minimum luminance below 0.01 nits [ITU-R, 2013].

The gamma correction, formally known as opto-electrical transfer function (OETF), that the BT.709 standard recommends to encode real world scene tristimulus values is:

$$V_{709} = \begin{cases} 4.5Z, & 0 \leq Z < 0.018 \\ 1.099Z^{0.45} - 0.099, & 0.018 \leq Z \leq 1 \end{cases} \tag{3.1}$$

where Z represents (linear-light) R, G or B component and V_{709} corresponds to gamma-corrected R', G' or B' value. It is important to recall from section 3.1.6 that, theoretically, a pure power function could be good enough for gamma correction, but notice in Fig. 3.13 that the OETF of BT.709 is not a pure power function: it has a linear segment with a slope of 4.5 near black. There is a practical reason for this: the slope of a pure power function with an exponent less than one is infinite at zero. Therefore, using a pure power function in a practical system

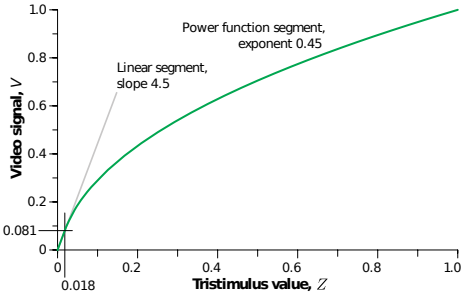


Figure 3.13: BT.709 OETF to map scene tristimulus to video code. Image from [Poynton, 2012].

(such as in a video camera) would produce infinite gain for values near black, and consequently, producing images with a significant noise in dark regions [Poynton, 2012], [Bertalmío, 2014].

Provided that the camera was operated with *factory* BT.709 settings, the encoded R'G'B' values can be converted into linear RGB signal by inverting Eq. (3.1) as:

$$Z = \begin{cases} \frac{V_{709}}{4.5}, & 0 \leq V_{709} < 0.018 \\ \left(\frac{V_{709}+0.099}{1.099}\right)^{\frac{1}{0.45}}, & 0.018 \leq V_{709} \leq 1 \end{cases} \tag{3.2}$$

Poynton mentions in his book [Poynton, 2012] that some people describe the OETF (or alternatively called as encoding gamma) of BT.709 as $\gamma_E \approx 0.45$ for picture rendering at the camera side, however, it is not accurate because the effect of linear segment for low values and offset terms make the overall transfer function very similar to a square root ($\gamma_E \approx 0.5$). Although television engineers approximated the non-linearity of CRT displays as a power law with an exponent of 2.4, it was never documented in BT.709 standard due to the fact that the physics of CRTs implicitly provided this behavior and there was not any need to explicitly specify the value of this non-linearity in the BT.709 standard. But with the popularity of digital flat panel displays that differ greatly in technology, ITU-R proposed a standard [ITU-R, 2011] in which the value for decoding gamma or electro-optical transfer function (EOTF) is specified as $\gamma_D = 2.4$, so that the modern displays produce consistent results among them and also emulate the EOTF of the legacy CRT.

In practice the effective power law with 0.5 exponent is used in the camera and

when it is combined with a display gamma of 2.4, we obtain an end-to-end gamma of 1.2, which is commonly known as **system gamma** or alternatively known as opto-optical transfer function (OOTF) [Borer, 2014]. The system gamma is a deliberate attempt of keeping a non-linearity in the signal in order to compensate for the subjective effects of viewing images in different ambient conditions.

Note that the values in Eq. (3.1) are described in the nominal range 0 to 1. For the 8-bit quantization in digital video standards, these values are scaled by 219 and offset by +16, thus giving 219 total steps between reference black to reference white. Values below 16 and above 235 are reserved for providing footroom and headroom that are necessary to preserve filter transients (over and under shoots) [Poynton, 2012].

3.2.3 sRGB system

With a joint effort of Hewlett-Packard and Microsoft in 1996, a new standard color space was proposed for computer monitors and image exchange on the internet. This standard, known as sRGB [Stokes et al., 1996], uses intentionally the same color primaries and white point as of BT.709, but a different transfer function for gamma correction. Because of its popularity and widespread use in operating systems, consumer software applications, internet browsers, desktop graphics, regular digital cameras and printers, it is considered as a default color space for image coding, meaning that if we receive an image without any information about its color primaries and white point, it is safe to assume that the image data is in sRGB format [Poynton, 2012]. To compensate the high levels of ambient luminance and flare that are typically found in office or home viewing conditions (unlike broadcast studios), the sRGB standard specifies an appropriate OETF to encode image values:

$$V_{sRGB} = \begin{cases} 12.92Z, & 0 \leq Z \leq 0.0031308 \\ 1.055Z^{\left(\frac{1}{2.4}\right)} - 0.055, & 0.0031308 < Z \leq 1 \end{cases} \tag{3.3}$$

Again, practically cameras use a pure power law to transform the tristimulus values of the original scene into the target display tristimulus values. And from the exponent ($1/2.4$) in Eq. (3.3), one may tend to use $\gamma_E = 1/2.4 \approx 0.42$, however, the correct approximation, due to the large offset and scale factor, is $\gamma_E \approx 0.45$. Such encoded image when displayed on a device that uses $\gamma_D = 2.4$ would

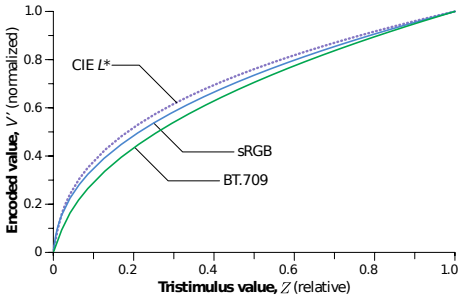


Figure 3.14: A comparison of BT.709, sRGB and CIE L* encoding functions. Image from [Poynton, 2012].

produce a system gamma of 1.1 that is significantly lower than the system gamma (1.2) of BT.709. In sRGB 8-bit image coding, each R, G or B component is encoded by utilizing the full range from 0 to 255, with no headroom or footroom. An illustration comparing the perceptual luminance-to-lightness function and the encoding functions of sRGB and BT.709 is depicted in Fig. 3.14 where it can be seen that all three curves are, although quite similar, not sufficiently close to be used interchangeably.

3.2.4 Cinema Standard: DCI-P3

In 2002, major motion picture studios formed an alliance, formally known as Digital Cinema Initiatives (DCI), to develop a standard architecture and document technical specifications for digital cinema. After arguing about what should be the appropriate set of color primaries to encode image data, it was mutually decided to use the widest gamut that could enclose the full visible spectrum in order to make a *future proof* standard [Kennel, 2007]. Hence, the DCI opted for the CIE XYZ color primaries having xy chromaticities [1, 0], [0, 1] and [0, 0] and equal-energy white reference with chromaticities [$1/3$,$1/3$].

Even though DCI recommended to use the XYZ primaries for image coding, practically all movies for digital cinema are mastered using the DCI-P3 color primaries that were proposed by Society of Motion Picture & Television Engineers (SMPTE) in [SMPTE, 2011] standard. This standard provides a comprehensive set of guidelines for movie encoding and environment conditions so that when

Table 3.2: DCI-P3 primaries and its white point.

	Red	Green	Blue	White
x	0.680	0.265	0.150	0.314
y	0.320	0.690	0.060	0.351

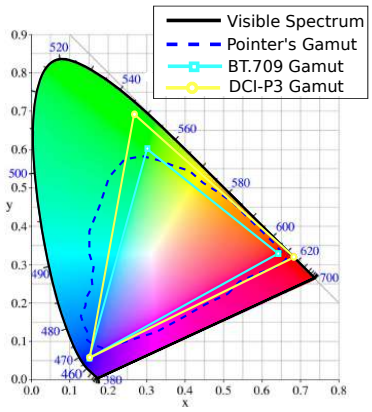


Figure 3.15: DCI-P3 gamut in the CIE xy chromaticity diagram.

the same movie is viewed using different projectors in different theaters, it has a similar and consistent visual appearance. The DCI-P3 standard uses a monochromatic ($615nm$) red primary, a non-monochromatic green primary, and it has the same blue primary as of BT.709/sRGB color space. In Table 3.2, the chromaticity coordinates of the DCI-P3 primaries and white point are listed. The DCI-P3 gamut is represented in Fig. 3.15 where it can be seen that it covers 45.5% area of the visible spectrum and 86.9% area of Pointer's gamut. The standard [SMPTE, 2011] specifies the exponent value of 2.6 for the decoding gamma transfer function and 12-bits per component are recommended for quantization. Movies that are inteded for cinema projection are graded with peak luminance of 48 nits and minimum luminance of 0.02 nits [Kennel, 2007].

3.2.5 Next-Generation Broadcast Standards: BT.2020 and BT.2100

In 2012, ITU-R recommended a new standard called BT.2020 [ITU-R, 2012] for the next generation ultra-high definition television. The key features of BT.2020 include two versions of spatial resolution (4K: 3840×2160 and 8K: 7680×4320), a larger bit depth of 12-bits per component, a higher frame rate frequency up to 120 Hz (progressive scan only), and a wider color gamut. The red, green and blue primaries of the BT.2020 standard are monochromatic. These primaries form a gamut that encompasses the DCI-P3 and BT.709/sRGB gamuts and covers 99.9% area of Pointer's gamut as shown in Fig. 3.16. It indicates that any display capable of realizing the full BT.2020 gamut will provide enhanced visual experience by reproducing nearly every color found in nature. BT.2020 [ITU-R, 2012] recommends converting scene radiance to digital code values using the same transfer function as used by BT.709, see Eq. (3.1). The color primaries and reference white of BT.2020 are summarized in Table 3.3.

Both BT.709 and BT.2020 provide images with a very limited dynamic range because they aim at emulating the characteristics of legacy CRTs: limited image brightness and less detail in dark regions. But state of the art display devices are capable of reproducing images with higher contrast ratio as they can produce very bright whites and extremely dark blacks on a pixel basis. Therefore, in 2016, ITU-R modified the BT.2020 standard and provided two important extensions: support

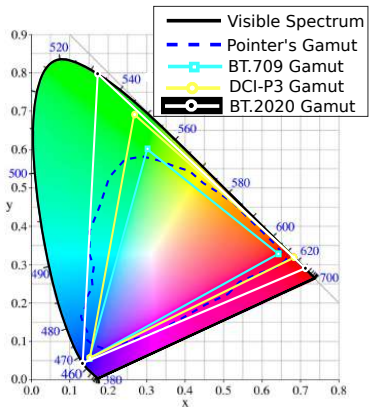


Figure 3.16: BT.2020 gamut in the CIE xy chromaticity diagram.

Table 3.3: ITU-R BT.2020 primaries and its white point.

	Red	Green	Blue	White, D_{65}
x	0.708	0.170	0.797	0.3127
y	0.292	0.797	0.046	0.3290

for full HD resolution (1920×1080) and support for high dynamic range (HDR) imagery. This standard is commonly known as BT.2100 [ITU-R, 2016]. There are two types of HDR image encoding methods specified in BT.2100. The first one is the Perceptual Quantizer (PQ) [Miller et al., 2013], [SMPTE, 2014] from Dolby that is developed through psychophysical experiments [Daly et al., 2013b] and therefore matches well with the human perception. The other one is called Hybrid Log Gamma (HLG) curve [Borer and Cotton, 2016] that is jointly developed by BBC and NHK through an engineering approach, so that it stays compatible with legacy display devices and matches closely with the already established TV transfer functions. Moreover, unlike PQ, the HLG transfer function takes into consideration the ambient luminance and produces images adjusted according to the viewing conditions. It is important to mention that, from production to display, one of these methods should be used consistently, and they should not be intermixed at any stage.

3.3 Significant Gamut Variation Among Displays

We have just seen that the standard color gamuts are significantly different from each other. There are a variety of display devices available in the market-place: each of which has its own color gamut. While each display device is capable of encompassing atleast one standard gamut defined for each specific market (broadcast, film, web, etc.), it may have color primaries that do not exactly match with any standard color gamut. In the following sections, we are going to discuss briefly some of the most common types of display devices.

3.3.1 What Makes the Gamut of a Display Wide?

There are two factors that contribute in defining the color gamut of a display device: its light source and its color filter array. The color filter array contains triplets of red, green and blue color filters, and each triplet of color filters corresponds to a pixel of the display. The job of these color filters is to separate the incoming light into the red, green and blue components. Thus, for example, the green color filter will allow only certain wavelengths of green light to pass, while attempting to block the undesired red and blue light components. In order to generate a highly saturated green color, either the filter needs to have a very narrow spectra, or the green component present in the incoming light must have a narrow spectrum that is finely aligned with the peak wavelength of the color filter. The same is true for reproducing highly-saturated red and blue primary colors. (TVs use color filter array, but in digital projectors either a prism or a color filtering wheel is used, as we shall see.)

The process of making filters with narrow spectrum is extremely difficult [Bertalmío, 2014] and very expensive [Poynton, 2010b]. Moreover, it is not optimum from the engineering point of view because such filters will incur significant noise penalty [Poynton, 2012], [Hubel et al., 1997] and loss of luminance [Jian et al., 2014]. This is why, all major display makers are investing efforts in developing light sources that can produce a pure white light containing the red, green and blue components that are either completely monochromatic or have a narrow spectrum.

3.3.2 Cathode Ray Tube (CRT)

The pioneering (and the most dominant in the past century) display technology was the CRT. A schematic in Fig. 3.17 shows some basic elements of a color CRT [Wandell and Silverstein, 2003]. The CRT uses a vacuum tube and at its neck lie three electron guns (one for each red, green and blue primaries) to which the input video signal is applied. Together the focusing electrode and the deflection plates modulate each electron beam that eventually strikes the inside face of the phosphorescent screen. The phosphorescent screen contains three different types of colored phosphors, each emits either red, green or blue light when stimulated by the electron beam. A metal sheet perforated with small holes is placed in front of the screen to ensure that when three electron beams sweep across the

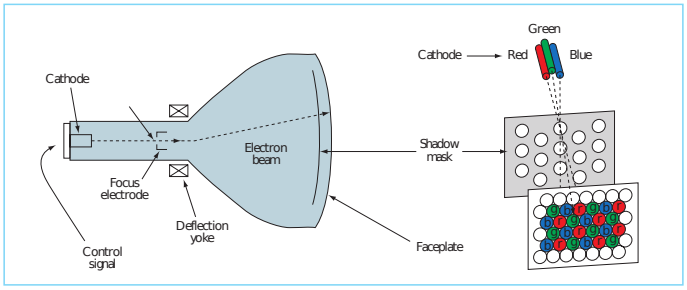


Figure 3.17: The basic architecture of a color CRT. Image is from [Wandell and Silverstein, 2003].

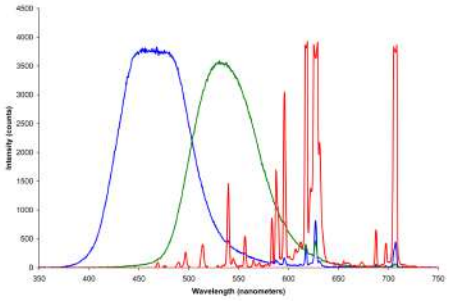


Figure 3.18: Spectra of color phosphors of a typical CRT. Image is from [Wikipedia, 2006].

entire screen, each beam illuminates only the intended phosphor and no others. Following a process called cathodoluminescence [Wandell and Silverstein, 2003], phosphors absorb electrons and emit visible light at the screen for us to perceive.

As illustrated in Fig. 3.18, the spectral power distributions of the red, green and blue phosphors used in a common CRT are significantly spread across the visible spectrum, therefore, it can only produce imagery with a limited gamut. In a pitch-dark room, CRT monitors are capable of producing an excellent black level of around $0.01\ cd/m^2$ [Soneira, 2004] but ordinary peak luminance of around $100\ cd/m^2$. Though this indicates that CRTs have a dynamic range of 4 orders of magnitude, this is not the case in typical viewing conditions. Since the screen of a CRT is coated with a fluorescent layer, it strongly reflects ambient light, and consequently elevates the display black level, therefore reproducing content with

a limited dynamic range of just 2 orders of magnitude [Daly et al., 2013a].

3.3.3 Liquid Crystal Display (LCD)

The first technology superseding CRTs, and currently evolving at a rapid pace, is the flat-panel LCD [Lueder, 2010]. Figure 3.19 shows a few basic optical components used in the imaging path of an LCD [Wandell and Silverstein, 2003]. Light from the illumination source first passes through a polarizer and acquires a certain polarization state. The principle feature of an LCD is the use of liquid crystal molecules (that act like shutters) between two polarized plates rotated by 90-degrees with respect to each other. In a normal state, these (nematic) liquid crystals are arranged in a helix or a twisted structure and are properly aligned with the second polarizer, so that when there is no voltage applied to these crystals, the polarized light reaches the screen to what is known as the *white mode* of an LCD. On the other hand, when a large electric charge is applied to the liquid crystals, they react and rearrange themselves in a such a way that the light gets completely blocked by the second polarization film and does not reach to the screen to which we normally call the *black mode* of an LCD. By regulating the amount of voltage applied to the liquid crystals, the intensity of the light that will reach to the screen is controlled. There is a layer of color filters that contains pixels, and each pixel further has the red, green and blue sub-pixels that filter out the incoming light and produce color images at the screen. Provided an ideal backlight, the dyes and pigments used for these color filters define the primary colors and therefore the

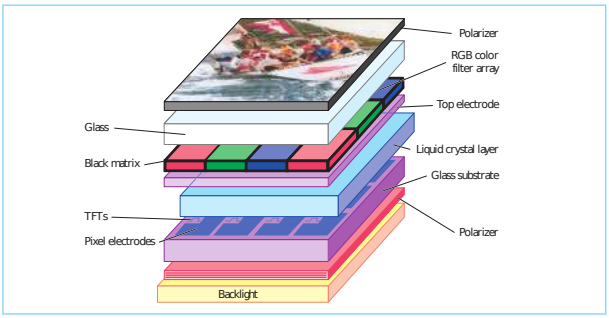


Figure 3.19: The basic architecture of a color LCD. Image is from [Wandell and Silverstein, 2003].

gamut of the display. But there is a limited variety of pigments and dyes that are compatible with the liquid crystals [Wandell and Silverstein, 2003].

As it may have become clear from the above explanation, liquid crystals are non-emissive; they respond to the light coming from an external light source to produce visible images. Some of the commonly used LCD backlight technologies are:

- **Cold Cathode Fluorescent Lamp (CCFL):** In general, small LCDs use two CCFLs placed at the opposite edges and large LCDs utilize multiple CCFLs arranged in parallel configuration for the backlighting. Most CCFLs are very bright and produce a fairly white spectra, but they are less energy efficient, and require costly and bulky apparatus to operate. Some display manufacturers make use of Hot Cathode Fluorescent Lamps (HCFLs) as they consume relatively less energy. LCDs with CCFL backlight are much brighter compared to CRTs and have a peak luminance typically in the range 100-500 cd/m^2. But their black levels are not as dim as of CRTs because of the powerful backlight that remains on all the time while operating, and even the best LCDs are not able to completely block the light reaching the screen when not needed [Yeo et al., 2008]. Therefore, conventional LCDs also provide dynamic range of 2 orders of magnitude.

- **Light Emitting Diode (LED):** LED-LCDs, or more commonly known as LED TVs, use light emitting diodes as a light illumination source in two different ways. The first conventional method of LED backlighting is to use an array of white LEDs placed at one or more edges of the display. The second way of implementing backlight is to use a grid of red, green and blue LEDs. Since RGB-LEDs provide a more purer white than both CCFLs and white-LEDs, they can produce a much wider color gamut. Apart from being thinner and consume less energy, the main advantage of the displays having LED backlighting is the ability to perform local dimming [Chen et al., 2006]. The local dimming technology enables displays to dim just those parts of the image that are required, while keeping the bright areas as bright as possible and in turn providing a greater contrast. LED TVs, due to local dimming, can provide a better dynamic range of around 3-4 orders of magnitude.

LED displays do not use one LED for each pixel because it is not physically

possible for consumer displays. Instead, they use a limited number of LEDs to create the global backlight and a diffuser plate is used to spread the light evenly; and the same is the case for the CCFL-LCD display devices.

3.3.4 Organic Light Emitting Diode (OLED)

Unlike LCDs, OLED displays [Mandle, 2010], [Hoffman et al., 2014] do not require an external backlight. An OLED panel has a layer of organic semiconductor that emits light upon receiving an electric current. An OLED display works on a pixel basis and it can completely switch off any pixel, therefore providing absolute blacks. In general, OLED displays are much brighter than CRTs and get nearly as bright as LCDs. These modern displays produce an incredible contrast ratio by reproducing deep blacks and bright whites simultaneously, with a dynamic range of 4 orders of magnitude, and therefore providing a better depiction of the scene. OLED TVs exhibit a wide color gamut, and some models (LG B6, LG C6, etc.) can reproduce almost full DCI-P3 color space [SMPTE, 2011]. Due to their high production costs, OLEDs are yet to become mainstream, and as of early 2017, only LG is actively releasing different models of OLED displays in the market-place, whereas Samsung has discontinued OLED production and moved to a new TV technology called quantum dots that we discuss in the following section.

3.3.5 Quantum Dots Technology

The light source at the back of a regular LCD has a wide spread power spectrum that is rich in blue and yellow, but has poor red and green components. When such light passes through the color filter array of the LCD, it generates suboptimal (broadband) primary colors that subsequently produce images having a limited gamut. To deal with this problem, Quantum Dot (QD) technology is employed in LCDs [Luo et al., 2014].

Quantum dots [Anikeeva et al., 2009] are semiconductor nanocrystals that emit colors when energized with a short-wavelength light i.e., blue light. The main physical property of QDs is that their emission color depends on their size rather than the material with which they are made. For example, Fig. 3.20 shows several QDs varying in size from $2nm$ to $6nm$, and when they are excited with the same blue light, each of them will emit the light of a specific wavelength (or

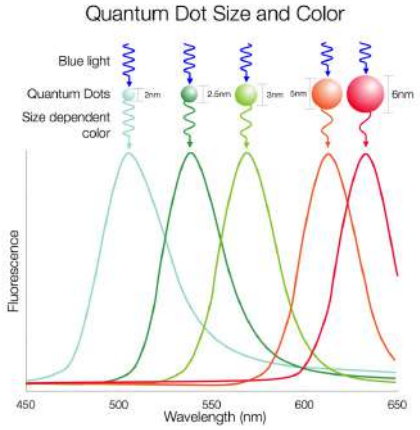


Figure 3.20: Spectral characterisitics of quantum dots: emission color of a quantum dot relates to its size. Image adapted from [Jian et al., 2014].

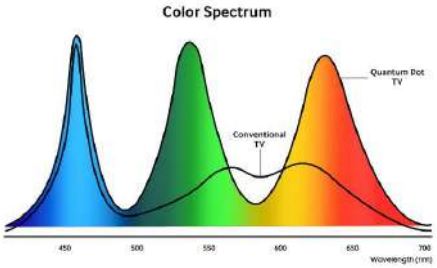


Figure 3.21: Spectral comparison of a conventional LCD with a QD TV. Image is from [Samsung, 2016].

color): the smallest QD will emit blue light, and the biggest QD will generate red light, with all other visible colors appearing in-between. However, it is important to mention here that the light these QDs emit is not monochromatic, rather it has a very narrow spectra in the range $30nm - 40nm$ [Jian et al., 2014]. This way display makers tune QDs to emit light precisely at the desired peak red, green and blue wavelengths, thus creating an optimal white light for LCDs that will subsequently produce images with a very wide color gamut [Zhu et al., 2015].

An example comparing the light spectrum of a conventional TV with a QD TV is illustrated in Fig. 3.21.

3.4 Digital Projectors

The primary components that virtually every digital projector needs to generate an image include a light source, a light processing engine, a set of color filters, a prism and a lens. As illustrated in Fig. 3.22, light from the lamp goes through an optical diffuser that focuses light onto a spatial modulator called the microdisplay. Then from the microdisplay, the light passes through a projection lens and forms a magnified image on the screen.

Before describing the various light sources used in digital projectors, we are first going to explain the process of modulating light to form an image on the projection screen. Currently, there are four major digital cinema projector manufacturers: Barco, Christie, Sony and NEC. Except for Sony, all the rest use the reflective Digital Light Processing (DLP) technology [Hornbeck, 1997] as a microdisplay in their projectors. The other options for the microdisplay are the LCD (explained above), and Liquid Crystal on Silicon (LCoS) that is a kind of hybrid technology between the DLP and the LCD.

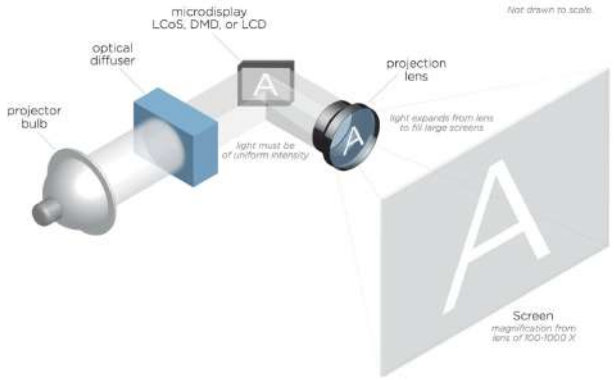


Figure 3.22: A highly simplified schemtic of a digital projector. Image is from the website of Laser Illuminated Projector Association [LIPA, 2014].

3.4.1 Digital Light Processing (DLP)

DLP technology [Hornbeck, 1997] by Texas Instruments modulates light using an optical semiconductor chip called the Digital Micromirror Device (DMD) [Dudley et al., 2003]. A DMD uses a rectangular array of microscopically small mirrors where the light from each mirror corresponds to a pixel on the projected image. Each DMD mirror has an electrode underneath which causes it to act as a light switch; the mirror faces either toward the projection lens or toward the heat sink (alternatively knows as *light dump*). In the former case, commonly referred as 'on-pixel', the mirror reflects light that then transmits through the lens and forms a white pixel on the projection screen. Whereas in the latter case, known as 'off-pixel', the reflected light does not reach the screen and therefore the corresponding pixel space remains dark. The intermediate positions between on and off states of the mirror correspond to 1024 different shades of gray, for which the mirror has to move exceptionally fast. A very simple illustration of the on-pixel and off-pixel mechanism is shown in Fig. 3.23.

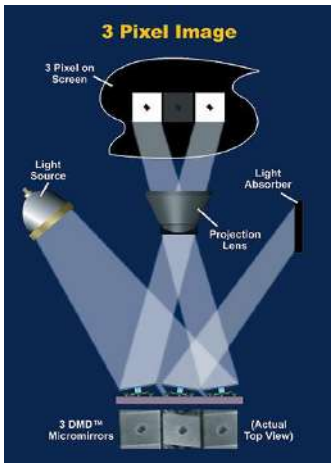
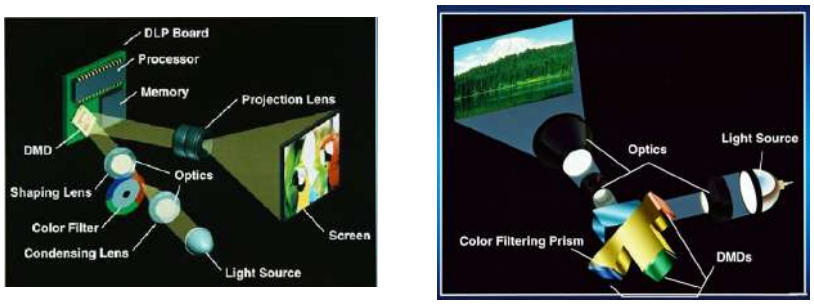


Figure 3.23: Working of DMD: on-pixel and off-pixel mechanism. Image is courtesy of Texas Instruments.

Color in DLP Systems

In each DLP projector, one out of the following two configurations is used to produce color images [Hornbeck, 1997].

- **Single-chip DLP projector:** A spinning wheel containing the red, green and blue primary colors is placed between the light source and the DMD as shown in the Fig. 3.24a. As the color wheel spins, depending on its color sector, the corresponding filtered light falls onto the DMD panel and subsequently reaches to the screen. For example, when the red filter is in front of the light source, the DMD reflects the red channel of the image at the projection screen. The same process is repeated for the green and blue channels. But this process occurs so rapidly that the viewer perceives a composite true color image. Single-chip DLP projectors are mostly deployed in academic institutes, boardrooms and home theaters. One disadvantage is that in the single-chip projectors two-third of the light is lost.
- **Three-chip DLP projector:** In digital cinema projectors, manufacturers use three DLP chips in order to provide higher brightness and greater color fidelity. Figure 3.24b shows that the light first passes through a prism that splits it into the red, green and blue primary colors. Each of these colored lights then strikes its own DMD. The colored lights reflected from three



(a) (b)

Figure 3.24: Two different type of DLP systems. (a) Single-chip DLP projection. (b) Three-chip DLP projection. Images are courtesy of Texas Instruments.

DMD chips are combined together before passing through the projection lens onto the screen.

3.4.2 Light Sources in Projectors

There are four types of light sources currently being used in digital projectors: lamps, LEDs (explained above), lasers and laser phosphors.

Lamps

In cinema projectors, still the most dominant solid-state illumination source is the lamp [Whelan and DeLair, 2012], mainly because of its price advantage over other technologies. From the 1900s to the 1960s, initially incandescent lamps and later carbon arc lamps were commonly used as a light source in movie projectors. These lamps were replaced by the Xenon short-arc lamp and the ultra-high pressure (UHP) lamp; current lamp-based projectors typically have either one of them, mainly because they can be excellent "day-light" simulators [Wyszecki and Stiles, 1982]. The brand new Xenon arc lamps can produce 33000 lumens (the lumen lm is a unit to measure the total amount of light that a projector emits) and they have an average lifetime of 500 hours [Belle, 2016]. But their brightness fades drastically over time, for instance, after about 200 hours of use, the brightness of a digital cinema projector drops to $22000\ lm$ and it falls below $15000\ lm$ within 800 hours [Beck, 2014]. The DCI compliant lamp-based digital movie projectors are capable of reproducing the full DCI-P3 color gamut.

Laser RGB

A laser (*l*ight *a*mplification by *s*timulated *e*mission of *r*adiation) is a solid-state light source with three main characteristics: it is monochromatic, very bright and coherent. Color primaries in a laser-RGB projector are generated by using the red, green and blue lasers. Since laser projectors [Silverstein et al., 2011], [Kusakabe et al., 2013] have monochromatic primaries with high color purity [Beck, 2014], they are able to cover the very wide BT.2020 gamut, and therefore reproducing nearly every color found in nature and providing the audience with a compelling color experience. Conventional lamp-based projectors are often too dim for 3D movies screening [Beck, 2014], meaning that the image may have a washed-out

appearance due to lack of contrast. However, digital projectors with RGB laser illumination can reach up to $60000\ lm$ [Stojmenovik, 2014], which is not only useful for 3D movies but for 2D movies too, because more brightness provides a large range of luminance and therefore a greater contrast.

One undesirable phenomena caused by laser light is called speckle. It is an interference pattern that occurs when coherent light gets scattered from an optically rough surface, such as a screen. The shimmery speckle pattern is mostly visible in uniform bright areas (e.g. sky) and it tends to decrease the perceived contrast by masking information in the image. For this reason, it needs to be suppressed. Common solutions to reduce speckle include the usage of an optical diffuser [Trisnadi, 2002] or a vibrating projection screen [McKnight and Curtis, 2014].

Laser Phosphor

As of 2016, red and green lasers are very expensive to make, but the blue lasers are inexpensive due to their mass production for blu-ray players. A laser-phosphor digital cinema projector uses an array of blue lasers and a yellow phosphor wheel. Light from the blue laser diode illuminates the phosphor wheel which then emits the yellow light. The yellow light passes through a dichroic filter that splits it into red and green primary lights. Later these red and green lights are combined with the direct blue laser light using a light integrator to obtain a broadband white light. Laser-phosphor projectors produce up to $6000\ lm$ [Stojmenovik, 2014].

Advantages of Laser projectors over Lamp Projectors

There are several advantages that laser projectors have over lamp projectors:

- **Image quality:** when compared to lamp-based projector, laser projectors reproduce images uniformly on the screen with better contrast, extremely wide color gamut and high brightness. Laser projectors do not suffer with the shortcoming of flicker as it is the case in the low-end lamp-based projectors.

- **Simplicity and high reliability:** Lasers have a lifetime of 20000 hours, which is a lot longer than a Xenon lamp. Lamp-based projectors need some

warm up time before showing any image, whereas laser projectors start up instantly. The brightness of laser projectors does not decay over time. One of the major advantages of laser projectors is the redundancy of laser diodes, meaning that if one or more laser diodes fail, the other diodes keep on working and we still have the image on screen. In contrast, if there is a sudden lamp failure (explosion), there would not be any image.

3.5 Color Transforms

Cameras yield content according to one of the following standard color gamuts: BT.709 for HDTV, BT.2020 for UHDTV, DCI-P3 for digital cinema, and sRGB for computer monitors. This standard procedure is necessary for cameras (and the same is true in the movie postproduction) to follow, so that later on, displays can decode the content accurately by adapting to these standard color gamuts. The adaptation procedure requires the transformation of the content to a device independent color space (e.g. CIE XYZ) and then to the native color space of the display.

3.5.1 Transformations between RGB and CIE XYZ

To transform an RGB color space to the CIE XYZ color space, we need to use a 3×3 transformation matrix M.

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} M \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \tag{3.4}$$

Computing Transformation Matrix: Given the chromaticity coordinates (x_r, y_r), (x_r, y_r), (x_r, y_r) of the color primaries and white point (X_W, Y_W, Z_W) of an RGB system, the transformation matrix M can be computed as:

$$M = \begin{bmatrix} S_r X_r & S_g X_g & S_b X_b \\ S_r Y_r & S_g Y_g & S_b Y_b \\ S_r Z_r & S_g Z_g & S_b Z_b \end{bmatrix} \tag{3.5}$$

where,

$$X_r = \frac{x_r}{y_r}; X_g = \frac{x_g}{y_g}; X_b = \frac{x_b}{y_b}$$
$$Y_r = Y_g = Y_b = 1 \tag{3.6}$$
$$Z_r = \frac{1 - x_r - y_r}{y_r}; Z_g = \frac{1 - x_g - y_g}{y_g}; Z_b = \frac{1 - x_b - y_b}{y_b}$$

and

$$\begin{bmatrix} S_r \\ S_g \\ S_b \end{bmatrix} = \begin{bmatrix} X_r & X_g & X_b \\ Y_r & Y_g & Y_b \\ Z_r & Z_g & Z_b \end{bmatrix} \begin{bmatrix} X_W \\ Y_W \\ Z_W \end{bmatrix} \tag{3.7}$$

For example, we can convert RGB values of an image represented in BT.709 color gamut (see primaries and white point in Table 3.1) into CIE XYZ values as:

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} 0.412453 & 0.357580 & 0.180423 \\ 0.212671 & 0.715160 & 0.072169 \\ 0.019334 & 0.119193 & 0.950227 \end{bmatrix} \begin{bmatrix} R_{709} \\ G_{709} \\ B_{709} \end{bmatrix}$$

The second row of this transformation matrix provides the luminance coefficients of BT.709 color space, and they sum to one because white is normalized to one in RGB systems. Since the column vectors of the M matrix are the XYZ tristimulus values, we can recover xy primaries from them using Eq. (2.7). Furthermore, to recover white point, first transform RGB = [1, 1, 1] to XYZ triplet and then calculate x and y chromaticities.

Similarly, the CIE XYZ values can be transformed into RGB values using the inverse of the M matrix as:

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = \big[M\big]^{-1} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \tag{3.8}$$

3.5.2 Transformations among RGB Systems

Values from one RGB system to another RGB system can be transformed as:

$$\begin{bmatrix} R_D \\ G_D \\ B_D \end{bmatrix} = \big[M_D\big]^{-1} \big[M_S\big] \begin{bmatrix} R_S \\ G_S \\ B_S \end{bmatrix} \tag{3.9}$$

where M_S and M_D are the transformation matrices for the source and destination color spaces, respectively; both of them are computed from their corresponding primaries and white points. (Let us note that in Eq. (3.9) the product of M_S and RGB triplet produce a CIE XYZ triplet.)

Transformation among RGB systems may produce values that are negative or greater than unity, meaning that these colors will fall outside the destination gamut. One simple method to fix this issue is to set negative values to zero and values greater than unity to one. However, this process of clipping may produce hue shifts in the image. Therefore, instead of using this lossy and problematic method, it is better to apply a gamut mapping procedure to obtain results perceptually faithful to the original content and without artifacts.

While performing these transformations, we need to pay attention at two important aspects:

- Conversion among RGB systems produce correct results only when it is applied to linear-light tristimulus values. However, practically in all standard RGB color spaces, each value is encoded using a non-linear transfer function. Therefore, before applying transformations we need to take out the effect of the transfer function in order to obtain linear-light values.

- If the source and destination systems have the same white point, only then the Eq. (3.9) will produce accurate results. In case white points differ (for example, as among BT.709 and DCI-P3 color spaces), the process of chromatic adaptation needs to be applied between ${M_D}^{-1}$ and M_S in Eq. (3.9).

Chromatic Adaptation

Frequently it happens that the illumination source under which the image was captured is different than that of the display device on which it is going to be displayed. In fact it is a very common situation in gamut mapping, image editing, etc. For example, the white point used in BT.709 image is ($x = 0.3127, y = 0.3230$) but if we display this image as it is using a digital cinema projector that expects a white point of ($x = 0.314, y = 0.351$), then there will be a noticeable color-shift from the actual captured colors. Content capturing is another situation

during which keeping white points the same between different scenes (and even within the same scene) is important. To deal with such issues a Chromatic Adaptation Transform (CAT) [Susstrunk et al., 2000] is required that allows us to perform illuminant to illuminant mapping in order to preserve the appearance of image colors.

Given the source reference white (X_{WS}, Y_{WS}, Z_{WS}) and the target reference white (X_{WD}, Y_{WD}, Z_{WD}), a source color (X_S, Y_S, Z_S) can be transformed into a destination color (X_D, Y_D, Z_D) as:

$$\begin{bmatrix} X_D \\ Y_D \\ Z_D \end{bmatrix} = \left[M_{CAT}\right]^{-1} \begin{bmatrix} \rho_D/\rho_S & 0 & 0 \\ 0 & \eta_D/\eta_S & 0 \\ 0 & 0 & \lambda_D/\lambda_S \end{bmatrix} \left[M_{CAT}\right] \begin{bmatrix} X_S \\ Y_S \\ Z_S \end{bmatrix} \tag{3.10}$$

where M_{CAT} denotes the chromatic adaptation matrix. The responses (ρ, η, λ) are calculated as:

$$\begin{bmatrix} \rho_S \\ \eta_S \\ \lambda_S \end{bmatrix} = \left[M_{CAT}\right] \begin{bmatrix} X_{WS} \\ Y_{WS} \\ Z_{WS} \end{bmatrix}$$

$$\begin{bmatrix} \rho_D \\ \eta_D \\ \lambda_D \end{bmatrix} = \left[M_{CAT}\right] \begin{bmatrix} X_{WD} \\ Y_{WD} \\ Z_{WD} \end{bmatrix}$$

The two most widely used M_{CAT} matrices based on Bradford and Von-Kries methods are:

$$M_{Bradford} = \begin{bmatrix} 0.8951 & 0.2664 & -0.1614 \\ -0.7502 & 1.7135 & 0.0367 \\ 0.0389 & -0.0685 & 1.0296 \end{bmatrix}$$

$$M_{Von\,Kries} = \begin{bmatrix} 0.3897 & 0.6890 & -0.0787 \\ -0.2298 & 1.1834 & 0.0464 \\ 0 & 0 & 1 \end{bmatrix}$$

3.6 The Need for Gamut Mapping

In this section, let us briefly remark the need for gamut mapping in different application domains. As we mentioned in section 3.1.5, different cameras have

different color spaces, but they must produce content according to standard color gamuts (e.g. BT.709 for home video, DCI-P3 for cinema, etc.). If cameras opt to apply this process of color conversion using a 3×3 matrix, the reproduced footage may have artifacts such as hue shifts, loss of spatial detail and spurious colors. Therefore, there is a need for GMAs that are fast, accurate and automatic. This is also the case in post-production, regardless of the color capabilities of the cameras (in movies, the color gamut is chosen during grading). We have also seen in previous sections how different technologies enable display devices to generate significantly different color gamuts. Though some displays have wider color gamuts than others, all these devices receive input signals encoded with standard color gamuts. On one hand, gamut reduction is needed when the colors of the input image fall outside the display's gamut, in order to reproduce colors accurately. On the other hand, gamut extension is required when the target display's gamut is wider than the gamut of the input image, so as to exploit the full color rendering potential of the display device. Even though gamut extension is considered as an enhancement procedure at the moment [Morovič, 2008], it is gaining importance and becoming highly relevant with the introduction of wide-gamut TVs and cinema projectors in the market-place. GMAs can also be employed in printers, websites, photo-sharing online platforms, computer graphics, video games, etc.

CHAPTER 4

Literature Survey

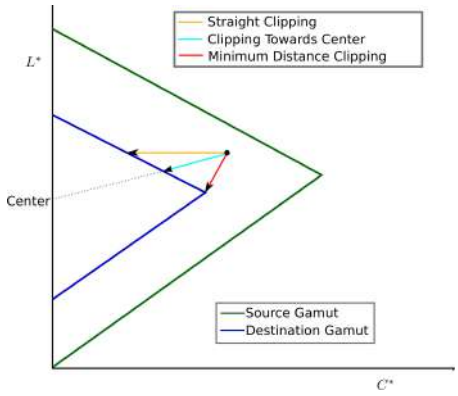
A plethora of gamut mapping algorithms exists in the literature; the majority of these methods perform gamut reduction, whereas only a few deal with the gamut extension problem. For a detailed explanation of gamut mapping algorithms, we refer the interested reader to the comprehensive book [Morovič, 2008].

In general, GMAs are divided into two broad categories: global GMAs and local GMAs. Global (also called non-local or non-adaptive) GMAs modify each color of an image independently, meaning that these methods completely ignore the spatial color distribution in the image. On the other hand, local GMAs modify pixel values by taking into account their neighborhoods.

4.1 Gamut Reduction Algorithms (GRAs)

4.1.1 Global GRAs

The global GRAs are further classified into two sub-classes: clipping and compression methods. Gamut clipping is the simplest approach to perform gamut mapping where colors that lie inside the destination gamut are left unmodified while those colors that fall outside are projected onto the destination gamut boundary by following a mapping criterion. An illustration of commonly used clipping techniques is presented in Fig. 4.1a. [Sara, 1984] described GRAs using four different clipping strategies. The first method is the LCLIP that clips



(a)

(b)

Figure 4.1: (a) An illustration of different commonly used clipping techniques. (b) An example showing artifacts due to gamut clipping. Top: input image. Bottom: reproduced image by using HPMINDE [Murch and Taylor, 1989]; notice the loss of texture on the head of the red parrot and on the neck of the yellow parrot. Original image is from [Kodak, 1993].

the chroma of out-of-gamut (OOG) colors to the destination gamut boundary along lines of constant hue and lightness (see example of straight clipping in Fig. 4.1a). The second method is called SCLIP that maps the OOG colors to the target gamut boundary towards mid gray point ($L^* = 50$), while preserving the hue (see clipping towards center in Fig. 4.1a). Both LCLIP and SCLIP are used for benchmarking GRAs in recent studies of [Froehlich et al., 2013] and [Masaoka et al., 2016]. The third approach (minimum ΔE) involves clipping of the OOG color to that color of the destination gamut boundary which has the smallest distance from it (see minimum distance clipping in Fig. 4.1a). The fourth method uses 26 points and for each of them a center of gravity is defined towards which the clipping is performed. [Montag and Fairchild, 1998] performed the psychophysical evaluation of the first three GRAs of [Sara, 1984] and two other clipping methods and they reported that the LCLIP method was preferred over the other competing GRAs. The results obtained using the third method of [Sara, 1984] causes hue shifts. To fix this issue, [Murch and Taylor, 1989] proposed

Hue Preserving Minimum ΔE (HPMINDE) method that involves clipping of the OOG color to the closest color, in terms of ΔE error, on the boundary of destination gamut along lines of constant hue. [Katoh and Ito, 1996] suggested an algorithm that uses a weighted color difference formula in CIELAB color space to map colors from source to destination gamuts. [Marcu and Abe, 1996] reviewed several previously published GRAs and presented a new GRA that first determines a center point on the lightness axis for each individual OOG color, and then performs a clipping operation on the OOG colors towards their corresponding centre points. [Masaoka et al., 2016] presented a GRA that aims at mapping colors from a very wide gamut (BT.2020) to a small gamut (BT.709), while preventing excessive chroma loss; this involves dealing differently with colors of low luminance and high luminance, specifically the authors map bright colors without respecting the constant hue lines in order to avoid a blown-out appearance.

All gamut clipping methods project the whole OOG color segment to a single point on the destination gamut boundary, and this may produce a gamut mapped image with a visible loss of texture and color gradients. In Fig. 4.1b, we present an example of mapping the input image to an extreme color gamut using the HPMINDE method [Murch and Taylor, 1989]. In the reproduced image the massive impact of gamut clipping can be seen in various regions, particularly on the head of the red parrot and on the neck of the yellow parrot. To overcome the issues associated with gamut clipping, gamut compression algorithms modify all the colors present in an input image. [Johnson, 1979] presented a GM method that linearly compresses the lightness and chroma. [Gentile et al., 1990] proposed two approaches: one performs compression along lines of constant hue and lightness, and the other involves compression along lines of constant saturation. [UGRA, 1995] uses chroma compression, and [Herzog and Müller, 1997] introduced linear and non-linear compression functions to obtain gamut mapped images. Sigmoidal Gaussian Cusp Knee (SGCK), recommended in [CIE, 2004], is a combination of GCUSP method presented in [Morovič, 1998] and the sigmoidal lightness mapping and cusp knee scaling proposed in [Braun and Fairchild, 1999]. At a given hue, the color that has the maximum chroma is called the **cusp**. SGCK, while keeping the perceived hue constant, compresses lightness and chroma along lines toward the point on the lightness axis having the same lightness as the cusp of the destination gamut, using a knee function. Gamut compression algorithms map a larger OOG color segment to a smaller in-gamut color segment

(a) (b)

Figure 4.2: Impact of gamut compression: (a) input image, and (b) reproduced image. Original image is from [Kodak, 1993].

and therefore they may cause a significant loss in saturation, especially when the difference between the source gamut and the target gamut is large. One such example is presented in Fig. 4.2 where the effect of strong desaturation due to linear gamut compression is evident. The hybrid approach of [Yuan et al., 2015] uses a combination of a gamut clipping method and a gamut compression method in order to perform gamut reduction. The methods of [Pytlarz et al., 2016] and [Schweiger et al., 2016] make use of a soft-clipping operation that squeezes colors near the target gamut boundary in order to accommodate OOG colors.

Both clipping and compression GRAs operate pixel-wise and completely ignore the spatial color configuration of the source image. Consequently, such GRAs produce gamut mapped images that may be contaminated with artifacts and are perceptually off from the original images.

4.1.2 Local GRAs

Local GRAs are also known as ‘spatial’ methods. [Meyer and Barth, 1989] proposed a spatial GRA where the lightness of the source image is first compressed in the Fourier domain using a low pass filter followed by the chroma compression along lines of constant hue and (mapped) lightness. Then, the high frequency detail of the image is added back to the gamut compressed image. [Bala et al., 2001] described a similar method where the source image is processed using a gamut clipping algorithm, and as a second step the difference of luminance between the original and the mapped image is calculated. Subsequently, a spatial

filter is applied to the difference and the result is added back to the gamut mapped image. Another similar approach is recently presented by [Zhu et al., 2016]. Their method first decomposes the input image into a base layer (which contains large-scale image-edge information) and a detail layer (that contains image-texture information). Then a gamut clipping method is applied to the base layer to obtain a gamut mapped image, to which the detail layer is added. In the algorithms of [Meyer and Barth, 1989], [Bala et al., 2001], and [Zhu et al., 2016] another stage of gamut clipping is integrated to process the resulting image in case the spatial filtering operation places a few pixels outside the destination gamut. [Morovič and Wang, 2003] introduced a multilevel, full color GRA that first decomposes the image into a number of spatial frequency bands. Secondly, at the lowest frequency band, the lightness compression is applied followed by the application of initial gamut mapping. Then, the next higher frequency band is added to the gamut mapped image and again gamut mapping is applied to the resulting image. This step is repeated until the highest frequency band is reached. The core idea was to transform the gamut of the original image into the destination gamut while preserving the original image's overall information into the reproduced image. [Kimmel et al., 2005] presented a variational approach where the GM problem is formulated as a quadratic programming optimization form, which is guaranteed to have a unique optimal solution provided that the target device has a convex gamut. They defined a functional to measure both the image difference and its derivatives to perform gamut reduction operation. [McCann, 1999], [McCann, 2002] proposed a Retinex-inspired framework that performs spatial comparisons to preserve the local gradients to obtain the final gamut mapped image. A similar multi-resolution GR approach that adapts the original image is suggested by [Farup et al., 2007]. [Zolliker and Simon, 2007] presented an unsharp masking scheme that combines the spatial filtering with the global GRA to retain the local contrast information in the reproduction. [Alsam and Farup, 2012] proposed an iterative GRA that at iteration level zero behaves as a gamut clipping algorithm, whereas, by increasing the number of iterations, the solution approaches spatial gamut mapping. [Gatta and Farup, 2017] introduced a Retinex-inspired method that performs gamut mapping in two stages. First the relative luminance of the input images is mapped to the target gamut. Then a clipping operation is applied on the saturation component of the image in order to obtain the final gamut mapped image. [Nakauchi et al., 1999] defined gamut mapping as an optimization problem where they use a perceptual metric to

minimize the perceived differences between the original and the reproduced image in order to obtain the final reduced-gamut image, and many other algorithms [Lau et al., 2011], [Preiss and Urban, 2012], [Preiss et al., 2014] followed a similar idea of optimization. The functionality of modifying pixel values taking into account their local neighborhoods certainly makes GRAs adaptive and flexible but at the same time far more complex and computationally expensive than global GRAs. Spatial GRAs are often based on many assumptions, and may produce halo artifacts.

4.2 Gamut Extension Algorithms (GEAs)

The case of gamut extension is different from that of gamut reduction: only a handful of GEAs exist in the literature. One could think of simply taking a GRA and use it to perform one-to-one mapping in the reverse direction in order to obtain a gamut extended image, as commented in [Morovič, 2008]. However, the key struggle is to produce gamut extended images that are natural, pleasant and perceptually as similar as possible to the original images. Same as with GRAs, GEAs are also classified into two broad categories: global GEAs and local GEAs.

4.2.1 Global GEAs

[Hoshino, 1991] proposed the first non-local GEA that extends colors from the limited-gamut printed images to the color gamut of high-definition TV. His algorithm first maps the lightness using a non-linear tone reproduction curve, and then the chroma is mapped along lines of constant lightness and hue. However, the gamut extended images were unnatural in appearance and therefore a revised version of that GEA was presented in [Hoshino, 1994]. [Kang et al., 2003] performed numerical fitting of data obtained by allowing a group of observers to manually extend lightness and chroma in a linear manner. [Anderson et al., 2007] introduced a user assisted method where an expert was asked to expand the gamut of some key frames and the rest of frames were corrected accordingly.

While all these aforementioned GEAs treat each color without analyzing the content of the input image, [Pan and Daly, 2008], [Casella et al., 2008], [Heckaman and Sullivan, 2011] introduced methods that first classify the colors of the input image according to some criterion and then extend them. In particular,

the work of [Pan and Daly, 2008] labels each color of a given image as skin or non-skin, and then applies two sub-GEAs; out of which one is designated to render skin colors and the other one expands the non-skin colors. Finally, the wide-gamut image is obtained by blending the output of both sub-GEAs using a linear combination. The GEA of [Casella et al., 2008] uses a sigmoid function into linear mapping techniques to deal with objects of low chroma and high chroma differently. And the method of [Heckaman and Sullivan, 2011] identifies certain memory colors such as green grass and blue sky, and renders them independently to make full use of the target device's gamut while preserving the flesh tones.

The algorithm of [Liu et al., 2010] uses the CIELUV color space to expand the color gamut of the input image from an anchor point while respecting the hue lines. [Kim et al., 2004] described a GEA with three types of extension strategies: chroma mapping, mapping along lines from the origin (which they called vector mapping), and adaptive mapping that is a compromise between the first two strategies. [Laird et al., 2009] proposed and evaluated the following five GEAs:

- **True-color:** The true-color algorithm maintains the color information of the input image into the destination gamut without applying any sort of extension, meaning that the output of true-color is nothing but a representation of the input image in a wide-gamut color space:

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix}_{true-color} = M^{-1}_{Destination} M_{Source} \begin{bmatrix} R \\ G \\ B \end{bmatrix}_{Source}, \qquad (4.1)$$

  where both M_{Source} and $M_{Destination}$ are 3×3 transformation matrices to convert the values of RGB to XYZ color space.

- **Same Drive Signal (SDS):** The most commonly used GE algorithm is the SDS, where the RGB primaries of the input material are linearly mapped to the RGB primaries of the display device enabling the SDS algorithm to make full use of the wide-gamut of the display.

- **Hybrid Color Mapping (HCM):** The HCM algorithm linearly combines the output of the true-color and SDS algorithms based on the saturation of

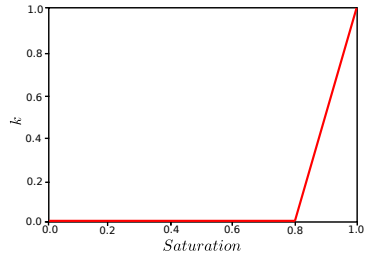


Figure 4.3: Mixing function for HCM.

the input image:

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix}_{HCM} = (1-\kappa) \begin{bmatrix} R \\ G \\ B \end{bmatrix}_{true-color} + \kappa \begin{bmatrix} R \\ G \\ B \end{bmatrix}_{SDS}, \qquad (4.2)$$

where κ is a mixing factor that works as a function of saturation (see Fig. 4.3).

- **Chroma Extension (CE):** The SDS algorithm works by extending the input signal to the destination gamut. However, hue differences between the source and destination gamuts may lead SDS to produce an image that exhibits hue shifts. To tackle this problem, the chroma extension algorithm maps colors of the source gamut to the reproduction gamut along lines of the chroma axis in the CIELCh color space, while keeping lightness and hue constant.

- **Lightness Chroma Adaptive (LCA):** The lightness-chroma adaptive algorithm performs GE by altering both lightness and chroma while keeping the hue constant. Both CE and LCA algorithms make use of the so called high chroma boost (HCB) function which smoothly maps colors of an input image in a manner that the high chroma objects get more boost in saturation than the low chroma ones. This approach aims to preserve memory color objects as they often have less saturated values, and extends colors of artificial objects (such as plastics) that have high chroma.

There are a few other GEAs [Meng et al., 2013], [Song et al., 2014a], [Song et al., 2014b] that mainly aim at preserving the skin tones in the reproductions. In

three different thesis works [Braun, 1999], [Ling, 2001], [Chen, 2002], authors use GRAs in the reverse direction to obtain images with extended gamut.

4.2.2 Local GEAs

To the best of our knowledge, apart from the GEAs that we have developed in this thesis, there is only the following previous work that extends colors taking into account their local neighborhoods. [Li et al., 2011] presented a multilevel framework where first a non-linear hue-varying extension function is applied globally to expand the color gamut of the input image. Then in the second stage they make use of image dependent chroma smoothing to avoid an over-enhancement of contrast and to retain in the result the local information of the input image.

4.3 Reproduction Intent: Accuracy or Pleasantness

Every GMA aims at reproducing content according to the specific application in which it is going to be employed. For instance, a GMA that is intended to be used in the cinema industry needs to reproduce images that are faithful to the vision of the content's creator. Whereas in the television industry, TV makers prefer distorting image attributes such as tones and colors in ways that they think consumers may find visually pleasant [Poynton, 2010a].

On one hand, GMAs with *accurate reproduction intent* aim at reproducing images that are perceptually faithful to the originals. On the other hand, GMAs with *pleasant reproduction intent* reproduce images that a viewer deems pleasant. The latter may imply departing from the original content as much as needed, and incorporating steps that could modify the aesthetics (contrast, sharpness, etc.) of images. In this study, our goal is to develop GMAs that comply with the accurate reproduction intent, and the results these algorithms produce are evaluated with the criterion of accuracy. One could question the choice of evaluation criteria: why ask users to choose the most accurate result instead of the one they find most pleasant? The reason is that, for a gamut mapping technique to be adopted by the movie industry, it must yield gamut mapped results that preserve as much as possible the artistic intent of the content's creator. Designing GMAs for pleasantness does not guarantee this, usually quite the opposite: apart from

general inter-subject variability in terms of preference, there is also a strong correlation between colorfulness and perceived image quality [Fedorovskaya et al., 1997], so user tests based on subject preference would rank higher those GMAs that produce saturated images even if that implied a departure from a subdued color palette chosen by the movie artist. User tests based on accuracy, on the other hand, are much less subjective (as users are simply asked to estimate differences with respect to the provided reference) and the preservation of the artistic intent is in a sense "built-in" in the procedure, since the ground truth acts as a stand-in for the content's creator intent.

4.4 Evaluation of GMAs

There are two possible ways of evaluating GMAs: subjective and objective evaluation.

4.4.1 Subjective Evaluation

In the case of subjective evaluation, subjects take part in psychophysical experiments where they have to choose or rate the reproductions based on a criterion (preference or accuracy). Psychophysical studies to evaluate GMAs follow closely the guidelines of [CIE, 2004] in which the conditions of the experimental setup are detailed: for example, viewing conditions, method of comparison, minimum number of observers and test images, maximum time duration per experiment session, methods for subject's color vision testing, etc.

Pair Comparison

The most common psychophysical method is the pair comparison, where two different gamut mapped versions of an original image are shown to observers in isolation or alongside the original image. Observers are then asked to select the gamut mapped image which exhibits more of the property (pleasantness, naturalness, or accuracy) being evaluated.

In the pair comparison evaluation, in order to calculate differences among n chosen GMAs, observers need to compare $n(n-1)/2$ number of pairs for each test image. In such experiments, observers are either given an option of tie in case

the visual appearance of reproductions is equally close to the original image, or forced to pick one of the reproductions anyway. For a given pair of reproductions, a score of 1 is given to the reproduction which is selected by an observer, and a score to 0 to the other reproduction. And when there is a case of tie, both reproductions receive a score of 0.5. The responses of an observer are typically stored in a $n \times n$ raw matrix where the value in column i and row j denote the score given to GMA i as compared with GMA j.

In this study, we use the pair comparison method to gather raw experiment data. And to analyze the raw data we use the same approach as used in the work of [Morovič, 1998] (see chapter 5 of his thesis), that is based on Thurstone's law of comparative judgement [Thurstone, 1927]. There are several other studies that make use of the pair comparison scheme to evaluate GMAs such as the works of [Montag and Fairchild, 1997], [Bonnier et al., 2006], [Muijs et al., 2006], [Hardeberg et al., 2008], [Dugay et al., 2008], and [Laird et al., 2009].

Category Judgment

In a category judgement experiment the observer is shown, one at a time, several images reproduced by different GMAs and asked to evaluate their pleasantness by assigning them descriptive names (such as excellent, good, fair, poor, bad) or just numbers from an integer scale (e.g., 1-10). This method is based on the law of categorical judgement [Torgerson, 1954]. Before starting the main experiment, it is recommended by [CIE, 2004] to show observers for training purposes a pair of images, one of which serves as an example of the best quality image and the other one as a worst case example.

Rank Order

In a rank order test, observers are asked to rank a given set of images from the worst to the best (or the other way round) according to a perceptual attribute.

4.4.2 Objective Evaluation

Subjective evaluation is time consuming, expensive and often unreliable if the number of observers is not sufficient. Therefore, an alternative is to use objective quality metrics that are capable of finding specific distortions in reproduced

images. There exists a vast variety of image quality metrics [Berns, 1993], [Luo et al., 2001], [Barańczuk et al., 2009], [Lissner et al., 2013], [Eskicioglu and Fisher, 1995], [Preiss et al., 2014] in the literature that could in principle be used to quantify the results of GMAs. [Hardeberg et al., 2008] and [Barańczuk et al., 2009] presented psychophysical studies where they identify the best performing image quality metric for the gamut reduction problem. It is important to note that the ranking of color metrics for gamut reduction may not be consistent in the context of gamut extension if the metrics are not trained to predict well the distortions found in gamut extended images. In the following section, we discuss the image quality metrics that are widely used to evaluate GMAs.

CIEΔE_{76}

The CIE XYZ color space is non-uniform in terms of visual perception. In contrast, CIELAB is a more perceptual color space and it became popular in the color industry mainly because it is straightforward to compute the color difference between two images simply by using the Euclidean distance. The ΔE distance between two colors x and y is defined as:

$$\Delta E_{76}(x,y) = [(L_x - L_y)^2 + (a_x - a_y)^2 + (b_x - b_r)^2]^{\frac{1}{2}} \tag{4.3}$$

Even though the CIELAB color space is more perceptual than CIE XYZ color space, it is not perfectly perceptual. Therefore, many attempts [Berns, 1993], [Berns, 2000], [Luo et al., 2001] were made to develop sophisticated error metrics that can predict the color distortions more accurately.

CIE94

[Berns, 1993] proposed an error metric which is derived from a detailed analysis of a large set of psychophysical data. It was later adopted by the CIE and given a name CIE94;

$$\Delta E_{94}(x,y) = \left[\left(\frac{L_x - L_y}{k_L S_L}\right)^2 + \left(\frac{C_x - C_y}{k_C S_C}\right)^2 + \left(\frac{H_x - H_y}{k_H S_H}\right)^2\right]^{\frac{1}{2}} \tag{4.4}$$

where L, C and H denote lightness, chroma and hue, respectively. The parametric variables k_L, k_C and k_H are set to unity, and the weighting functions S_L, S_C and

S_H as defined as

$$\begin{aligned} S_L &= 1 \\ S_C &= 1 + 0.045C_x \\ S_H &= 1 + 0.015C_x \end{aligned} \tag{4.5}$$

CIEDE2000

The CIE94 comes with two major drawbacks; first it is only reliable for large color differences ($\Delta E > 5$), and second its performance is suboptimal for blue colors. To deal with these issues, [Luo et al., 2001] presented a method that estimates the distortions not only in the lightness, chroma and hue components but also incorporates an interactive term between chroma and hue differences. [CIE, 2001] named it as the CIEDE2000 color difference metric:

$$\Delta E_{00}(x,y) = \left[\left(\frac{L_x - L_y}{k_L S_L} \right)^2 + \left(\frac{C_x - C_y}{k_C S_C} \right)^2 + \left(\frac{H_x - H_y}{k_H S_H} \right)^2 + \Delta R \right]^{\frac{1}{2}} \tag{4.6}$$

where,

$$\Delta R = R_T f(\Delta C \Delta H) \tag{4.7}$$

For more details regarding parametric variables k_L, k_C and k_H , weighting functions S_L, S_C and S_H, and interactive term between chroma and hue differences R_T, see [Luo et al., 2001] and [Sharma et al., 2005].

Combination of a ΔE Metric and a Contrast Measure

[Barańczuk et al., 2009] show that the ΔE metric and its variants do not measure accurately enough the quality of GMAs since they neglect the structural details of images. To better predict the quality of GMAs, [Barańczuk et al., 2009] introduced a metric which is a linear combination of Delta-E measure $Q_{\Delta E}$ and contrast preserving measure $Q_{\Delta LC}$.

The image quality metric $Q_{\Delta E}$ is the average ΔE over the pixels of the two images;

$$Q_{\Delta E}(X, Y) = \frac{1}{mn} \sum_{i=1}^{m} \sum_{j=1}^{n} \Delta E(x_{ij}, y_{ij}) \tag{4.8}$$

where ΔE could be replaced with any of its flavors (ΔE_{76}, ΔE_{94} or ΔE_{00}).

The image quality metric $Q_{\Delta LC}$ computes the Michelson contrast [Michelson, 1927] on a $k \times k$ patch $P_X \subset X$ of the image X as follows:

$$LC(P_X) = \frac{x_{max} - x_{min}}{x_{max} + x_{min}} \tag{4.9}$$

where in a patch P_X, x_{max} and x_{min} represent the highest and lowest luminance values, respectively. Similarly, the local contrast value $LC(P_Y)$ for the corresponding patch P_Y in the image Y is computed. Finally, the change in contrast is defined as

$$\Delta LC(P_X, P_Y) = |LC(P_X) - LC(P_Y)| \tag{4.10}$$

Subsequently, the image quality measure $Q_{\Delta LC}$ finds the overall contrast distortion by averaging the ΔLC over all possible patches in the images.

Color Image Difference Metric

The structural similarity index (SSIM) [Wang et al., 2004] is a widely used perceptual metric to assess the quality of a reproduction by comparing it with a reference image. Despite of its excellent prediction accuracy against distortions such as lossy compression, noise, blur, and channel fading, the SSIM completely ignores the changes in color since it operates on grayscale images only. But in the case of gamut mapping the chromatic components (chroma and hue) are the ones mainly affected, and therefore cannot be ignored in the quality assessment process. To this end, [Lissner et al., 2013] presented a perceptual color image difference (CID) metric based on SSIM. The CID metric is particularly tailored to assess the quality of gamut mapped images and it predicts the quality by using the following combination of five image difference features (IDFs): lightness-difference L_L, lightness-contrast C_L, lightness-structure S_L, chroma-difference L_C and hue-difference L_H. Mathematically,

$$CID = 1 - L_L C_L S_L L_C L_H. \tag{4.11}$$

Note that the arguments are omitted from Eq. (4.11) for the sake of brevity, and for the detailed explanation, see [Lissner et al., 2013].

Later, [Preiss et al., 2014] show that if the CID metric is used in the gamut mapping optimization problem [Preiss and Urban, 2012], the results are often contaminated with artifacts such as lightness inversion, chromatic ringing, chromatic edges, and lightness banding. To overcome these challenges, [Preiss et al., 2014] proposed an improved color image difference (iCID) metric in which, apart from the IDFs in [Lissner et al., 2013], two additional IDFs are introduced: chroma-contrast C_C and chroma-structure S_C. The Eq. (4.11) is extended as

$$iCID = 1 - L_L C_L S_L L_C L_H C_C S_C. \tag{4.12}$$

CHAPTER 5

Gamut Mapping Algorithms in RGB Color Space

In this chapter we present our spatial gamut mapping algorithms that rely on a perceptually-based variational framework. The text of this chapter is mainly reproduced from our journal article titled "Gamut mapping in cinematography through perceptually-based contrast modification" [Zamir et al., 2014], and our conference publication titled "Considering saliency in a perception inspired gamut reduction algorithm" [Vazquez-Corral et al., 2014].

The first half of this chapter provides detail of our GMAs (both GEA and GRA) published in [Zamir et al., 2014]. Our algorithms adapt an image energy functional [Bertalmío et al., 2007] whose minimization leads to image enhancement and contrast modification. We show how by varying the importance of the contrast term in the image functional we are able to perform gamut reduction and gamut extension. We propose an iterative scheme that allows our algorithms to successfully map the colors from the gamut of the original image to a given destination gamut while keeping the perceived colors close to the original image.

In the second half of this chapter, we present a saliency inspired GRA by making modifications in the GRA of [Zamir et al., 2014] in order to better respect the saliency of the original image in the reproduced one. The results show that the proposed approach [Vazquez-Corral et al., 2014] produces a gamut mapped image whose saliency is closer to that of the original image with a minor loss of accuracy.

5.1 Perceptually-based Color and Contrast Enhancement

In this section we briefly recap the main concepts of the Retinex theory of Land, arguably the most popular approach for perceptually-based color correction, and its close relationship to the variational method for contrast enhancement which we will adapt for our gamut mapping purposes (for details we refer the interested reader to [Bertalmío, 2014]).

[Land, 1977] in his paper makes a very clear and detailed explanation of his Retinex theory and the experiments that led to its postulation. After scores of perceptual matching tests, his conclusion was that our perception of the color of an object had a physical correlate in what he called *scaled integrated reflectance*, which is defined, for each waveband (long, medium and short, corresponding to cone response sensitivities,) as a ratio: the integral of the radiance of the object over the waveband, divided by the integral over the same waveband of the radiance of a white object under the same scene illuminant. The scaling is a non-linear function that relates reflectance to lightness sensation. But this implies that in order to perceive the color of an object somehow our visual system is comparing the light coming from the object with the light coming from a reference white, and Land wondered how we are able to find this reference white *"in the unevenly light world without reference sheets of white paper"* [Land, 1977]. The sensation of white will be generated by the area of maximum radiance in all three bands (this is the von Kries' model or "white-patch" assumption, although Land does not cite it); this area could be used as reference, but Land did not know how our visual system could *"ascertain the reflectance of an area without in effect placing a comparison standard next to the area"* [Land, 1977]. The solution he proposed consisted of comparing far-away points through paths: the ratio of the values at the two end-points of the path can be computed as the sequential product of the ratios of each point of the path with the following point. The Retinex algorithm consists of assigning, for each point and each waveband (long, middle, short), an estimate reflectance obtained as the average of the sequential products obtained on many paths, all ending in the given point. Land thought that this was a plausible explanation of how our visual system estimates reflectances but he did not want to venture where exactly this type of operations were being carried out, in the retina or at the cortex; therefore he chose the name "Retinex" for his approach. The Retinex algorithm is directly applied to digital images in a straightforward way, where the pixels will be the points and the three color channels R, G and B play

the role of the wavebands. As proved in [Provenzi et al., 2005], Retinex always increases brightness so it cannot be directly applied to overexposed pictures; also, if the algorithm is iterated the results may improve but the convergence image is flat white, so there is some "sweet spot" of the number of iterations yielding the best output [Bertalmío et al., 2009]. Another major source of problems is Retinex's reliance on paths: their length, shape and number condition the results and many works have been proposed trying to optimize the selection of these variables.

The Automatic Color Enhancement (ACE) algorithm of [Rizzi et al., 2003] is also based on perception, and its relationship with Retinex will become clear shortly. ACE is designed to mimic some basic characteristics of the human visual system, like the white patch and the grey world mechanisms, lateral inhibition, the independence of chromatic channels, or the influence of spatial relationships in the scene. [Rizzi et al., 2003] perform experiments that show how ACE has several excellent properties: it allows to obtain very good color constancy, it increases the dynamic range of the input and, unlike Retinex, it can deal both with under- and over-exposed pictures, it can perform de-quantization (eliminating quantization artifacts produced by encoding an image with an insufficient number of bits per channel), and it can reproduce some visual perception illusions. Its main limitation is its computational complexity, $O(N^2)$ where N is the number of pixels.

In the work titled *"Perceptual color correction through variational techniques,"* [Bertalmío et al., 2007] start by recalling the variational histogram equalization method of [Sapiro and Caselles, 1997], in which it is shown that the minimization of the energy functional

$$E(I) = 2\sum_x \left(I(x) - \frac{1}{2}\right)^2 - \frac{1}{AB}\sum_x \sum_y |I(x) - I(y)| \tag{5.1}$$

produces an image I with a flat histogram. The range of I is $[0, 1]$, x, y are pixels and A, B are the image dimensions.

In [Bertalmío et al., 2007] the energy in Eq. (5.1) is interpreted as the difference between two positive and competing terms

$$E(I) = D(I) - C(I), \tag{5.2}$$

the first one measuring the dispersion around the average value of $\frac{1}{2}$ (as in the gray world hypothesis), the second term measuring the contrast as the sum of

the absolute value of the pixel differences. But this measure of contrast is global, not local, i.e. the differences are computed regardless of the spatial locations of the pixels. This is not consistent with how we *perceive* contrast, which is in a localized manner, at each point having neighbors exert a higher influence than far-away points. Therefore, [Bertalmío et al., 2007] propose an adapted version of the functional of Eq. (5.1) that complies with some very basic visual perception principles, namely those of grey world, locality and not excessive departure from the original data:

$$\begin{aligned} E(I) =& \frac{\alpha}{2} \sum_x \left(I(x) - \frac{1}{2} \right)^2 - \frac{\gamma}{2} \sum_x \sum_y w(x,y) |I(x) - I(y)| \\ &+ \frac{\beta}{2} \sum_x \left(I(x) - I_0(x) \right)^2 , \end{aligned} \tag{5.3}$$

where α, β and γ are constant and positive weights, I is a color channel (R, G or B), $w(x,y)$ is a normalized Gaussian kernel of standard deviation σ, and $I(x)$ and $I(y)$ are two intensity levels at pixel locations x and y respectively.

By minimizing the image energy in Eq. (5.3) the aim is to maximize the contrast (second term of the functional), while not departing too much from the original image (third term) and also preserving the gray world hypothesis (first term). We can intuitively explain how the contrast term operates by considering the following. Since $\gamma > 0$, in order to minimize $E(I)$ we need to increase $\sum_x \sum_y w(x,y)|I(x)-I(y)|$, i.e. the local contrast, a weighted sum of local pixel differences. So if a pixel has a larger value than all the pixels in its neighborhood, then increasing the local contrast is achieved by enlarging these differences, increasing the pixel value even more. Conversely, when a pixel has a smaller value than all the pixels in its neighborhood, then increasing the local contrast is achieved by making the pixel value even smaller. If $\gamma < 0$, then the minimization of Eq. (5.3) *reduces*, not increases, the contrast, as pointed out in [Bertalmío et al., 2009]. The radius σ of the kernel $w(x,y)$ controls how local the contrast enhancement is: a very large σ corresponds to the global case of Eq. (5.1), whereas with a small σ value the contrast term is computed over small pixel neighborhoods.

It is formulated in [Bertalmío et al., 2007] that the solution to the minimization

of Eq. (5.3) can be found as the steady state of the evolution equation

$$I^{k+1}(x) = \frac{I^k(x) + \Delta t \left(\frac{\alpha}{2} + \beta I_0(x) + \frac{\gamma}{2} R_{I^k}(x)\right)}{1 + \Delta t(\alpha + \beta)} \tag{5.4}$$

where the initial condition is $I^{k=0}(x) = I_0(x)$. The function $R_{I^k}(x)$ indicates the contrast function:

$$R_{I^k}(x) = \frac{\sum_y w(x,y) s\left(I^k(x) - I^k(y)\right)}{\sum_y w(x,y)} \tag{5.5}$$

where x is a fixed image pixel and y varies across the image. The slope function $s(\cdot)$ is a regularized approximation to the sign function, which appears as it is the derivative of the absolute value function in the second term of the functional; in [Bertalmío et al., 2007] they choose for $s(\cdot)$ a polynomial of degree 7.

In [Bertalmío et al., 2007] it is shown that Eq. (5.3) has a single minimum and that it is also a fixed point of ACE. In other words, we can say that ACE is a numerical implementation of the gradient descent of Eq. (5.3). The minimization of Eq. (5.3) yields very good color constancy results and this method shares all the good properties and possible applications of ACE, plus the numerical implementation in [Bertalmío et al., 2007] has a reduced complexity of $O(NlogN)$, where N is the number of pixels.

This method can be used for contrast enhancement since it produces good results without halos, contrast reversals, spurious colors or any other kind of visual artifact. The main reason for this, formally proven in [Bertalmío et al., 2007], is that unless σ is too small, the minimization of Eq. (5.3) preserves the ordering of the level lines of the original image. Furthermore, the algorithm has a local but not pixel-wise nature: the value at each pixel is updated according to comparisons with all its neighbors. This comparison procedure has a sort of anchoring behavior, preventing pixel values from having sudden changes, which helps to explain why the same algorithm applied independently to the frames of a video produces results without temporal artifacts; we have corroborated this in our experiments, as we shall see.

There is a very close connection between the formulation of [Bertalmío et al., 2007] (Eq. (5.3)) and Retinex. In their kernel-based Retinex (KBR) formulation, [Bertalmío et al., 2009] take all the essential elements of the Retinex theory (channel independence, the ratio reset mechanism, local averages, non-linear

correction) and propose an implementation that is intrinsically 2D, and therefore free of the issues associated with paths. The results obtained with this algorithm comply with all the expected properties of Retinex (such as performing color constancy while being unable to deal with overexposed images) but do not suffer from the usual shortcomings such as sensitivity to noise, appearance of halos, etc. The KBR algorithm, as Retinex, always increases the values of the image it is applied to, and therefore it is not suited to overexposed images; this issue is linked in [Bertalmío et al., 2009] to the fact that there is not any energy that is minimized by the iterative application of the KBR algorithm. Using the analysis of contrast performed by [Palma-Amestoy et al., 2009], [Bertalmío et al., 2009] are able to determine how to modify the basic KBR equation so that it can also handle overexposed images, and the resulting, modified KBR equation turns out to be essentially the gradient descent of the energy of Eq. (5.3). In this way, the connection between Retinex, ACE, and the perceptual color correction of [Bertalmío et al., 2007] (Eq. (5.3)) becomes explicit.

5.1.1 Gamut Mapping via Energy Functional Adaptation

In this section, we adapt the image energy functional defined in Eq. (5.3) in order to perform gamut mapping from the gamut of the original image to the target gamut of a given device. Firstly, we replace the first term of the functional (5.3) as:

$$\begin{aligned} E(I) =& \frac{\alpha}{2} \sum_x \left(I(x) - \mu\right)^2 - \frac{\gamma}{2} \sum_x \sum_y w(x,y) |I(x) - I(y)| \\ &+ \frac{\beta}{2} \sum_x \left(I(x) - I_0(x)\right)^2, \end{aligned} \tag{5.6}$$

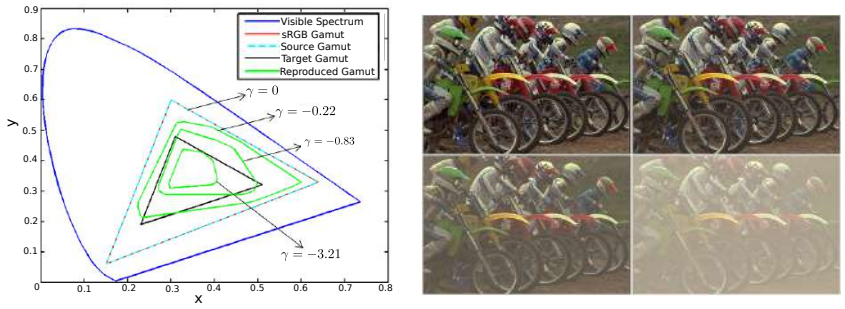
where μ is the mean average of the original image I_0. In this way we are penalizing departures from the original mean, instead of imposing the gray world assumption, which could cause a change in the color palette of the result. The resulting evolution equation is

$$I^{k+1}(x) = \frac{I^k(x) + \Delta t \left(\alpha\mu + \beta I_0(x) + \frac{\gamma}{2} R_{I^k}(x)\right)}{1 + \Delta t(\alpha + \beta)} \tag{5.7}$$

Also, we recall from [Bertalmío et al., 2009] that $\gamma \in \mathbb{R}$ in the functional is positive or negative depending on whether we want to maximize or minimize the contrast, respectively, and we will use positive values for gamut extension and negative values for gamut reduction.

Gamut Reduction Algorithm (GRA-RGB)

In order to perform a gamut reduction operation we minimize the contrast of the source image using Eq. (5.7) until the colors of the larger gamut converge to the smaller destination gamut. In cases where we need to apply GR, the contrast coefficient γ will be negative and the value of α will be small. The evolution Eq. (5.7) has a steady state for each particular set of values for α, β, Δt and γ. For example, in Fig. 5.1a, a chromaticity diagram is shown with different gamuts (visible spectrum, sRGB gamut, source gamut, target gamut and reproduced gamut). It can be seen that when $\beta = 1$, $\alpha = 0$, and $\gamma = 0$ the steady state of the evolution equation is equivalent to the original image. In the same figure we show that as γ decreases and α increases following a $\frac{|\gamma|}{20}$ slope, the steady state of Eq. (5.7) has a gamut which is gradually smaller. Fig. 5.1a shows that, just by selecting a small enough value for γ ($\gamma = -3.21$ in this case) we are already performing gamut reduction. However, in this case, colors that were originally inside the target gamut move inwards too much, and the appearance of the image



(a) (b)

Figure 5.1: Perceptual GR Approach. (a) Gamuts on chromaticity diagram. (b) Contrast reduction results. Top left: original image. Top right: $\gamma = -0.22$. Bottom left: $\gamma = -0.83$. Bottom right: $\gamma = -3.21$. As gamma becomes smaller the image becomes more washed out.

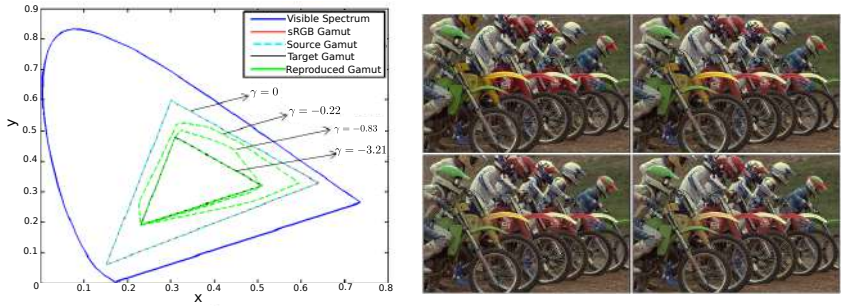
(a) (b) (c) (d)

Figure 5.2: Gradual mapping of colors. Out-of-gamut colors (in green) when (a) $\gamma = 0$, (b) $\gamma = -0.22$, (c) $\gamma = -0.83$, (d) $\gamma = -3.21$. As gamma decreases the number of out-of-gamut pixels is reduced.

becomes washed-up, as Fig. 5.1b shows.

To improve the previous result, we present an iterative method in terms of the contrast coefficient γ. At each iteration, we run Eq. (5.7) for some particular α, β, and γ until we reach the steady state. The steady state of each iteration will provide us with some pixels of the final result. At iteration 1, we set $\beta = 1$, $\alpha = 0$, and $\gamma = 0$, and therefore the original image is obtained as the steady state. We select the pixels that are inside the destination gamut for the final image and leave them untouched for the following iterations. We move to iteration 2, where we decrease γ (for example, setting $\gamma = -0.05$) and increase α in relation to γ by $\frac{|\gamma|}{20}$. We run again Eq. (5.7) until steady state, and we check whether any of the colors that were outside the gamut at the previous iteration have been moved inside the destination gamut. If this is the case, we select them for the final image and leave them untouched for the following iterations. We keep iterating by decreasing γ (and increasing α accordingly) until all the out-of-gamut colors come inside the destination gamut. An example of this iterative procedure is shown in Fig. 5.2, where green pixels represent out-of-gamut pixels remaining in that iteration. It can be seen in Fig. 5.3a that the reproduced gamut is covering a much wider range of colors than previously. It is shown in Fig. 5.3b that the colors are better preserved as compared to the previous example (see Fig. 5.1b).

We also want to note that the standard deviation σ of the Gaussian kernel w is of great importance; we observe in Fig. 5.4 that a small value of σ leads to the preservation of colors but may introduce a few artifacts, whereas for the larger values of σ each color pixel is strongly influenced by the surrounding colors. Therefore, our method computes several gamut mapped images $\mathcal{I}_\sigma$ by using four different values of standard deviations $\sigma \in \{\sigma_1, \cdots, \sigma_4\}$. Subsequently, in



(a) (b)

Figure 5.3: Modified Perceptual GR Approach. (a) Gamuts on chromaticity diagram. (b) Top left: original image. Top right: $\gamma = -0.22$. Bottom left: $\gamma = -0.83$. Bottom right: $\gamma = -3.21$. In the modified perceptual GR approach, the reproduced image covers a wider range of gamut and appears pleasant in appearance.

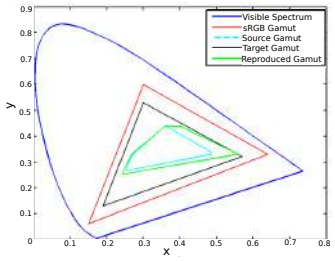
(a) (b) (c) (d) (e)

Figure 5.4: Effect of standard deviation (σ). (a) Original image. (b) I_σ, $\sigma = 25$. (c) I_σ, $\sigma = 100$. (d) I_σ, $\sigma = 200$. (e) I_{final}, using (5.8). The small values of standard deviation (σ) of the Gaussian kernel (w) preserve colors but introduce artifacts. The larger values of σ produce the gamut reduced images that are free from artifacts, but where colors are less saturated.

order to obtain a final gamut mapped image $\mathcal{I}_{final}$, we select for each pixel x, a value out of four gamut mapped images $\mathcal{I}_\sigma(x)$ which has the minimum Lab ΔE distance with compared to the original image value $\mathcal{I}_{orig}(x)$

$$\mathcal{I}_{final}(x) = \arg\min_{\mathcal{I}_\sigma} \left(Lab(\mathcal{I}_\sigma(x)) - Lab(\mathcal{I}_{orig}(x))\right)^2, \quad \forall x, \sigma \in \{\sigma_1, \cdots, \sigma_4\}. \tag{5.8}$$

This procedure increases the computational cost, but the quality of the reproduced image improves to a great extent.



(a)

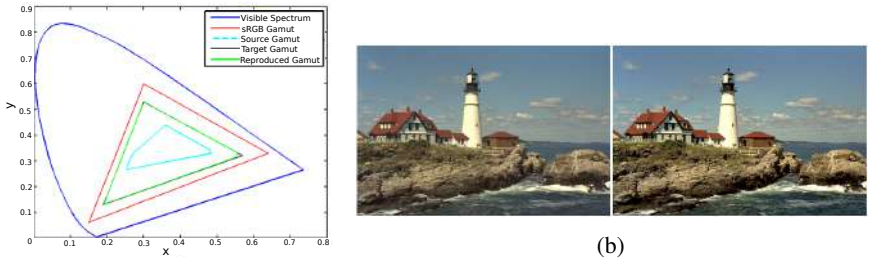
(b)

Figure 5.5: Perceptual GE Approach. (a) Gamuts on chromaticity diagram. (b) Left: input image. Right: result of first trial of perceptual GEA. As any of the colors from the source image touches the boundary of the target gamut, the reproduction is achieved.

Gamut Extension Algorithm (GEA-RGB)

In this section, the goal is to develop a gamut extension algorithm that is capable of accurately expanding the colors from a smaller source gamut to a larger destination gamut. To this end, we enhance the contrast of the source image by running the evolution Eq. (5.7) to the steady state while keeping the sign of γ positive and defining α as $\frac{|\gamma|}{20}$. To select the value of γ that is adequate for the extension, we pick the minimum γ for which any of the colors touches the boundary of the target gamut as it is presented in Fig. 5.5a. The drawback in this case is that the reproduced image (Fig. 5.5b) is less saturated, so we cannot realize the full potential of the reproduction device in terms of colors. To overcome this problem, we apply Eq. (5.7) to larger values of γ and let the original gamut exceed the destination gamut up to a certain threshold level (T_e). This threshold T_e controls the level of saturation; a large value of T_e indicates a higher saturation level, whereas a small value of T_e yields a less saturated output (however, the higher saturation level comes at the cost of less perceptually valid results). After this, the colors that were placed outside the destination gamut in the previous stage are mapped back inside using our gamut reduction algorithm. It is shown in Fig. 5.6b that the reproduced image is more pleasing and exhibits a wider gamut (Fig. 5.6a) as compared with the gamut extended image in Fig. 5.5b.

A preprocessing might be needed in cases where the source image has a medium-to-large dynamic range, since in this case when we apply the GEA-RGB



(a) (b)

Figure 5.6: Modified Perceptual GE Approach. (a) Gamuts on chromaticity diagram. (b) Left: input image. Right: result of proposed perceptual GEA. The result of GEA are pleasingly saturated (see the grass, water, roof of the house) and covers a wider range of gamut.

(a) (b) (c) (d)

Figure 5.7: Effect of preprocessing stage in GEA-RGB. (a) Input image. (b) GE without preprocessing. (c) Histogram shifted image. (d) GE with preprocessing. Notice the trees in the background; the colors are preserved when GE is applied after preprocessing.

the colors with low intensity levels tend to go towards black, something which is evident in Fig. 5.7b (see trees and forearms of the man). This indicates that the image content has a significant impact on the results of the GEA. To cope with this, we introduce a preprocessing step in our GEA where we give a shift to the histogram of each channel of the source image, from low towards high intensity values. The shift in histograms depends on the difference between the source and the target gamut; the shift is larger if the difference between gamuts is large, whereas a small shift is required when the difference between gamuts is smaller. As we mentioned earlier, we let our GE method extend the color gamut of the source image outside a certain threshold (T_e) of the destination gamut. Once the threshold T_e is achieved, we then take the corresponding γ and estimate the shift value as

$$shift(c) = \frac{\gamma \overline{I_c}}{\gamma_{ref}} \qquad c \in \{R, G, B\} \tag{5.9}$$

where $\overline{I_c}$ denotes the mean of the color channel I_c. In our GE experiments we observe that even for a large difference between the source and the target gamuts, a maximum amount of contrast coefficient $\gamma = 2$ is required, therefore we fix the reference gamma value $\gamma_{ref} = 2$ to calculate the shift factor for the histograms. We perform preprocessing on each channel of the image independently by normalizing its values between the shift value (obtained using Eq. (5.9)) and the maximum image channel value. Even though this improvement in the model comes at the expense of some computational complexity, nonetheless the results are much more enhanced as shown in Fig. 5.7d.

5.1.2 Experiments and Results

As stated in the introduction, good GMAs should not only reproduce the colors accurately but they should also generate results that match the inputs in terms of perception. In this section, we present both the qualitative and the quantitative evaluation of our GMAs on some very commonly used color images in the literature for gamut mapping applications. The original sRGB images are illustrated in Fig. 5.8 and these images exhibit a variety of spatial and chromatic characteristics. Furthermore, we perform experiments on videos, confirming that our GMAs can be applied independently on a frame-by-frame basis without producing temporal artifacts.

To compute the results for our GMAs (GRA-RGB and GEA-RGB), we work in the RGB color space by fixing the parameters $\beta = 1$, $\Delta t = 0.10$ and $\alpha = \frac{|\gamma|}{20}$. We apply Eq. (5.7) to the source image until convergence, which we assume that has been reached when the difference between two consecutive steps falls below 0.5%. In case of GR, we decrease the gamma value with the change $\Delta\gamma = 0.03$ in order to obtain the final gamut mapped image as explained in section 5.1.1. In the case of GE we need to select a γ value which is enough to extend the gamut of the source image to a certain threshold (T_e).

The primaries of gamuts that are used in this chapter are summarised in Table 5.1, and we will refer the gamuts using the naming convention mentioned in the same table.

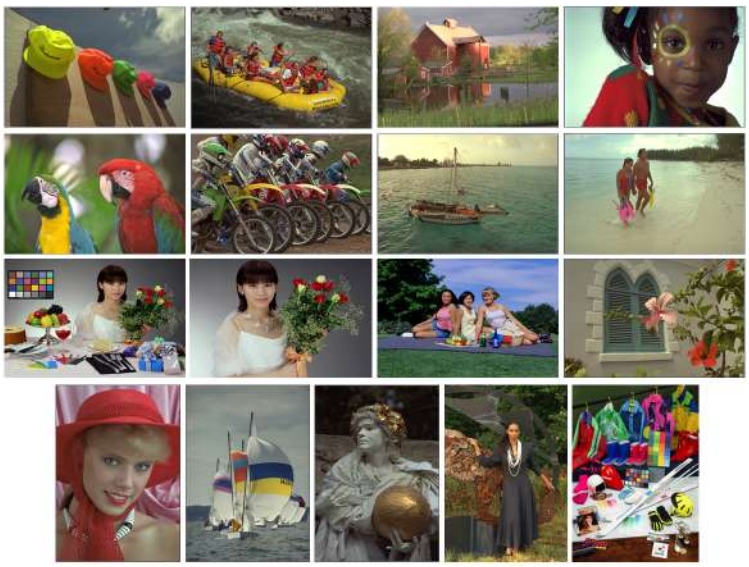


Figure 5.8: Original sRGB images. The first 2 images in row 3 are from [ISO, 2004], the third image in row 3 and the last image of row 4 are from [CIE, 2004]. Rest of the images are from [Kodak, 1993].

Results of Gamut Reduction

To perform GR, we apply the procedure explained in section 5.1.1 to map the colors of the original image inside the target gamut. We perform GR using four different values for the standard deviation $\sigma \in \{50, 100, 150, 200\}$ of the Gaussian kernel w, and combine all the four gamut reduced versions into a final gamut mapped image using Eq. (5.8).

Visual Quality Assessment: In order to validate the efficiency of our GRA-RGB, we apply our method on a rather challenging target gamut named 'Toy' in Table 5.1. Given an image in sRGB, our algorithm maps the gamut of the original image into the gamut 'Toy'. The results presented in Fig. 5.9 show that our proposed framework works well in preserving the colors, texture and color gradients from the out-of-gamut regions while staying faithful to the perception of the original image. For example, in Fig. 5.10, rows 1 and 4, it can be seen that the colors reproduced by our GM algorithm (fifth column) are much more

Table 5.1: Primaries of gamuts.

Gamuts	Red Primaries		Green Primaries		Blue Primaries	
	x	y	x	y	x	y
BT.709/sRGB	0.640	0.330	0.300	0.600	0.150	0.060
DCI-P3	0.680	0.320	0.265	0.690	0.150	0.060
Simulated BT.709	0.610	0.330	0.330	0.530	0.150	0.060
Simulated DCI-P3	0.640	0.330	0.300	0.600	0.150	0.060
Toast	0.570	0.320	0.300	0.530	0.190	0.130
Toy	0.510	0.320	0.310	0.480	0.230	0.190

Figure 5.9: Results of GRAs on still images. Column 1: original images. Column 2: output of HPMINDE [Murch and Taylor, 1989]. Column 3: output of [Lau et al., 2011]. Column 4: output of [Alsam and Farup, 2012]. Column 5: output of our GRA-RGB.

saturated than those of HPMINDE [Murch and Taylor, 1989] (second column), and the state of the art algorithms of [Lau et al., 2011] (third column) and [Alsam and Farup, 2012] (fourth column). Similarly, in Fig. 5.10, row 2, our algorithm not only reproduces the color efficiently but also preserves a great amount of texture. In Fig. 5.10, row 3, we can see that our method accurately represents the difference in the lightness of identical hue (see the pink socks and pink beanie).

In Fig. 5.11 we show the results for our gamut reduction method and for the

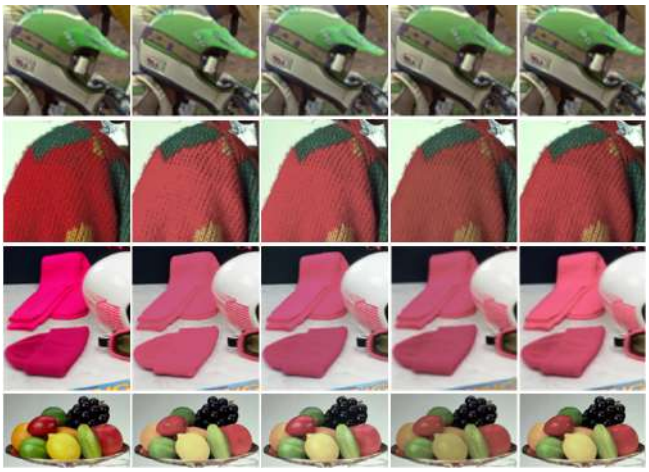


Figure 5.10: Detail preservation using GRAs on still images. Column 1: original cropped regions. Column 2: output of HPMINDE [Murch and Taylor, 1989]. Column 3: output of [Lau et al., 2011]. Column 4: output of [Alsam and Farup, 2012]. Column 5: output of our GRA-RGB. This is a closer (zoomed-in) view of the regions cropped from Fig. 5.9.

state of the art algorithm of [Alsam and Farup, 2012] on two different image sequences, one professional and the other amateur, from which we have selected challenging shots that are colorful and have noticeable camera and object motion. Both GRAs are applied to each frame independently. We used the target gamuts 'Toy' and 'Toast' to reduce the gamut of image sequence 1 and image sequence 2, respectively. It can be seen that, for both sequences (selected frames) shown in the figure, our algorithm produces more saturated results as compared with the method of [Alsam and Farup, 2012]. For example, in Fig. 5.12, the color reproduction of the girl's skin (column 1) is perceptually more faithful in our results than in those of [Alsam and Farup, 2012]. Furthermore, in columns 2 and 6, we can see that the method of [Alsam and Farup, 2012] has problems preserving the color of the very bright regions, whereas our algorithm efficiently reproduces those regions. It can be seen in column 4 that the algorithm of [Alsam and Farup, 2012] introduces a shift in the hue (see blue color), while on the other hand our method represents colors more faithfully. Moreover, the colors reproduced by our algorithm are more saturated and pleasant as compared with [Alsam and Farup, 2012] (see column 5).

(a)

(b)

Figure 5.11: Results of GRAs. (a) Image sequence 1 (professional footage). (b) Image sequence 2 (amateur video). In both image sequences: top row, original frames; middle row, output of [Alsam and Farup, 2012]; bottom row, output of our GRA-RGB. In the first 3 columns of image sequence 1, notice the color of the girl's skin, her shirt and the poster on wall, whereas in the last 2 columns notice the colors of the door, cushions on sofas and the lamp. In image sequence 2, differences in perception and color reproduction can be seen in almost every region of the frames such as floor, inside the shop, and clothes of pedestrians and street artists. Original image sequence 1 is property of CBS Interactive and taken from [CBS, 2013]. Original image sequence 2 is from [Kupaev, 2012].

Figure 5.12: Detail preservation using GRAs on video, all regions are cropped from Fig. 5.11. Top row: original cropped regions. Middle row: output of [Alsam and Farup, 2012]. Bottom row: output of our GRA-RGB.

Since we are using an extreme target gamut to map the colors of image sequence 1, we experience a few spatial artifacts in our results; similar artifacts are observed in the results obtained with the method in [Alsam and Farup, 2012]. In cinematography, the content of digital video needs to be mapped from the camera gamut to the BT.709 gamut, and these gamuts exhibit not so extreme differences with respect to each other. Therefore, to test our algorithm under a real scenario, we obtain a new target gamut by simulating the difference between DCI-P3 and BT.709 gamuts inside the BT.709 gamut. The simulation of gamuts is shown in Fig. 5.13. It turns out that our GRA performs efficiently under realistic settings without introducing any spatial artifacts as it can be seen in Fig. 5.14. Moreover, we observed that it makes sense to apply our method on each frame independently, since there are no temporal artifacts and therefore no coherence between frames is required.

Objective Quality Assessment: Visually, the results presented so far underline the good performance of our GR method in terms of visual quality. This subjective outcome is backed by using the perceptual color quality measure presented in [Lissner et al., 2013]: the Color Image Difference (CID) metric estimates

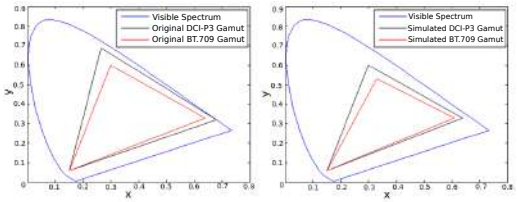


Figure 5.13: Simulation of relationship between DCI-P3 and BT.709 gamuts. Left: original gamuts. Right: simulated gamuts.

Figure 5.14: Spatial artifacts due to GRA-RGB. Top row: original images. Middle row: effect of extreme target gamut. Bottom row: effect of realistic target gamut. The relationship of target gamut and spatial artifacts can be seen at the forehead of the person, cup and on the pants.

the perceptual differences given by the changes, from one image to the other, in features such as hue, lightness, chroma, contrast, and structure.

Comparisons using the CID metric in the still images of Fig. 5.8 are provided in Fig. 5.15. In this figure, we can see that our algorithm outperforms the other methods in 15 out of 17 test images. Moreover, the CID statistical data (mean, median and root mean square) is also presented in Table 5.2, which justifies the reliable performance of our algorithm over the other approaches. In addition to objective quality tests on still images we also compare, using the CID metric,

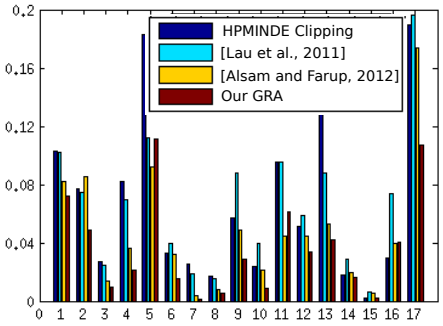


Figure 5.15: Quality assessment of GRAs on the still images from Fig. 5.8 using the CID [Lissner et al., 2013]. Horizontal axis: row-wise image number of Fig. 5.8. Vertical axis: perceptual error.

Table 5.2: Quality assessment of GRAs on still images (Fig. 5.8): CID perceptual error.

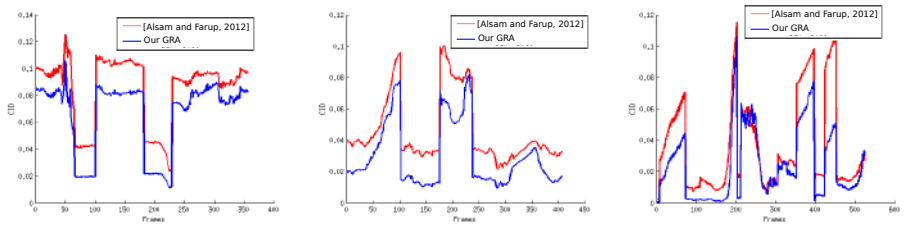
	Mean	Median	RMS
HPMINDE [Murch and Taylor, 1989]	0.0674	0.0514	0.0873
[Lau et al., 2011]	0.0665	0.0695	0.0807
[Alsam and Farup, 2012]	0.0472	0.0398	0.0627
Our GRA-RGB	**0.0368**	**0.0285**	**0.0495**

Table 5.3: Quality assessment of GRAs on videos: CID perceptual error.

	Mean	Median	RMS
[Alsam and Farup, 2012]	0.0528	0.0413	0.0623
Our GRA-RGB	**0.0372**	**0.0221**	**0.0474**

the perceptually-based efficiency of our algorithm and [Alsam and Farup, 2012] on the image sequences shown in Fig. 5.11 and on an animated video (where we perform reduction from the BT.709 to the 'Toast' gamut). The graphs shown in Fig. 5.16 and the CID statistical data presented in Table 5.3 depict that our method produces results with less perceptual error than the other method.

In conclusion, both the visual inspection and the objective assessment show that our GR method produces a gamut reduced video which is, perceptually, more faithful to the original video as compared with the other methods.



(a) (b) (c)

Figure 5.16: Quality assessment of GRAs on videos using the perceptual difference measure CID [Lissner et al., 2013]. (a) Image sequence 1 (Fig. 5.11a). (b) Image sequence 2 (Fig. 5.11b). (c) Animated video. Horizontal axis: time (in frame numbers). Vertical axis: perceptual error.

Results of Gamut Extension

Visual Quality Assessment: To quantify the performance of our GE approach, we use 'BT.709' as a target gamut and set the saturation threshold $T_e = 10\%$ of the total number of pixels. We apply the proposed GEA-RGB on videos that have the starting color gamut equal to the 'Toast' gamut (for the sake of fair comparisons, these videos are reduced to the 'Toast' gamut using clipping in the CIELAB color space). We compare our algorithm with the SDS algorithm [Laird et al., 2009], which is similar to the industrial method of performing gamut extension in new TV screens [Poynton, 2007]. We evaluate results of our GE method and SDS [Laird et al., 2009] by comparing them with the ground truth in order to analyze their efficiency. In Fig. 5.17, we present the results of our gamut extension procedure on videos where it can be seen that our expansion algorithm generates images that are pleasant and realistic in appearance. Fig. 5.18 shows zoomed-in details from Fig. 5.17. In columns 1 and 5 of Fig. 5.18, it can be seen that our GEA enhances the color of trees while keeping the sky color closer to the ground truth, whereas SDS [Laird et al., 2009] noticeably changes the color of sky regions. It is shown in columns 2 and 4 that the reproduction of skin color (face of the man and woman's arms) by our GEA is realistic in appearance, but the SDS algorithm [Laird et al., 2009] reproduces flesh tones very much departed from the ground truth. The results also show that the SDS approach [Laird et al., 2009] over-saturates some regions, which in turn looks artificial. For example, in columns 3 and 6, it is clearly noticeable that our GEA achieves very good color

(a)

(b)

Figure 5.17: Results of GEAs. (a) Image sequence 1. (b) Image sequence 2. In both image sequences; top row, ground truth (BT.709 gamut); middle row, output of Same Drive Signal algorithm [Laird et al., 2009]; bottom row, output of our GEA. In both image sequences, the results of our GEA are more pleasing and perceptually faithful to the ground truth as compared with the output of the SDS algorithm [Laird et al., 2009]; see zoomed-in details in Fig. 5.18. Both image sequences are the property of Paramount Pictures and taken from [Pictures, 2013].

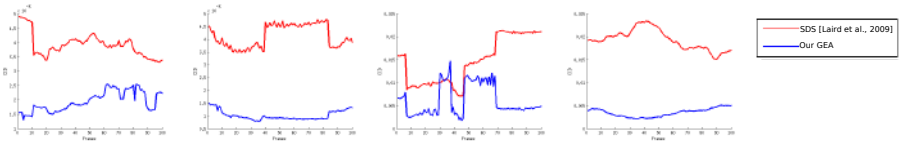
reproduction (wall, hood of the shed and hat) without introducing any unnatural colors. On the other hand, the results produced by [Laird et al., 2009] are highly saturated and quite different from the original content. We have noticed that our GEA not only extends the color gamut perceptually but also performs contrast enhancement, which is evident in Fig. 5.18 (columns 3 and 6); if this were a problem, the original contrast could easily be restored with a very small amount

Figure 5.18: Comparison of GEAs, details from Fig. 5.17; top row, ground truth; middle row, output of Same Drive Signal algorithm [Laird et al., 2009]; bottom row, output of our GE method. The results of the SDS algorithm [Laird et al., 2009] are highly saturated and off from the ground truth. See the reproduction of sky, trees, skin tone, wall, clothes, fruit bucket and hat.

of simple blurring.

Our GE approach shows a robust performance in both image sequences. It can also be seen that the results obtained using our algorithm are less saturated than those of SDS [Laird et al., 2009] in a few regions. If we increase the saturation threshold (T_e) our algorithm will achieve more saturated colors, at the expense of increasing the perceptual error of the results.

Objective Quality Assessment: The primary purpose of our gamut extension algorithm is to process the colors of a digital video in order to display it with digital cinema projectors. In this case, a video should be mapped from the BT.709 source gamut to the DCI-P3 target gamut. However, proper evaluation of the results would require a wide gamut digital cinema projector to display the mapped video under special (cinematic) lighting conditions and/or a way to present the DCI-P3 results in this document, as we are limited by sRGB gamut for the paper. Therefore, to quantify the performance of our gamut expansion method, we take an experimental approach where we simulate the difference between DCI-P3 and BT.709 gamuts inside the BT.709 gamut. Thereafter, we convert the BT.709 videos to the simulated sRGB gamut using a clipping on the xy components in the xyY color space while keeping the luminance component unchanged. Finally, we map the videos from the simulated BT.709 gamut to the simulated DCI-P3 gamut. Since in this case we have a ground truth to compare with (simulated



(a) (b) (c) (d)

Figure 5.19: Quality assessment of GEAs on videos using the perceptual metric CID [Lissner et al., 2013]. (a) Sequence 1, Fig. 5.11a. (b) Sequence 2, Fig. 5.11b. (c) Sequence 1, Fig. 5.17. (d) Sequence 2, Fig. 5.11b. Horizontal axis: time (in frame numbers). Vertical axis: perceptual error.

Table 5.4: Quality assessment of GEAs on videos: CID perceptual error.

	Mean	Median	RMS
SDS [Laird et al., 2009]	0.0105	0.0060	0.0128
Our GEA-RGB	**0.0031**	**0.0024**	**0.0041**

DCI-P3 gamut is equal to the original BT.709 gamut), we can assess our gamut extension algorithm quantitatively using the CID metric. For this experiment we set $T_e = 3\%$ since source and target gamuts have small color differences. The results are presented in Fig. 5.19, where it can be seen that our reproduction is perceptually more faithful to the original content than the results of SDS. The CID statistical data is summarized in Table 5.4, which justifies the good the performance of our GEA-RGB on videos.

We have observed in our experiments that the quality of the input signal also matters; if there exist any artifacts due to compression or noise in the input video they may become prominent once passed through the GE module. A possible way to deal with such kind of input material is to perform preprocessing using the approach reported in [Palma-Amestoy et al., 2009]. Finally, we want to stress that, as it was the case with our GRA, no temporal artifacts appear in the results of our GEA either.

5.2 Saliency-based Gamut Reduction Algorithm (SBGRA)

Image saliency predicts the attentional gaze of observers viewing a scene [Murray et al., 2013], [Seo and Milanfar, 2009]. It has been used as a cue to aid in the performance of both image processing and computer vision applications such as color to gray conversion [Gooch et al., 2005], [Ancuti et al., 2011], image detail visibility [Pedersen et al., 2013], and motion-compensated frame interpolation [Jacobson and Nguyen, 2011]. In a domain closer to that of gamut mapping, it has been used to decide whether a black point compensation algorithm is needed when printing an image [Lindner et al., 2011]. However, despite the fact that the goal of gamut reduction is to emulate the experience felt by the viewer in the original image, there is a limited amount of research relating saliency and gamut mapping. In the work of [Chen and Beghdadi, 2011] the authors considered saliency in their Image Gamut Boundary Reduction (IGBR) algorithm. The IGBR method was proposed as a preprocessing stage for gamut mapping algorithms. The idea is that the gamut boundary in the non-salient regions can be reduced without losing any discriminative power. Saliency models have also been proposed as gamut mapping artifact detectors by [Cao et al., 2010] and [Raja and Pedersen, 2013].

The goal of this particular work is to bring saliency one step closer to gamut mapping by introducing it in the gamut reduction process. The hypothesis of SBGRA is that it is possible to modify our GRA-RGB [Zamir et al., 2014] to better represent the saliency of the original image without any significant loss of the image reproduction quality.

5.2.1 Considering Saliency in Gamut Mapping

By using the GRA-RGB (explained in section 5.1.1), we can obtain a map of γ values used at each pixel accounting for the amount of contrast reduction that is necessary at that pixel. Let us use Γ_{GM} to denote this map of γ values. In this map all values are negative: the smaller the value is, the larger is the contrast reduction. Since GRA-RGB follows an iterative approach based on the γ parameter, it may result in a large difference of the contrast reduction performed in a region in comparison to its surrounding regions. This will affect the saliency of the mapped

image since contrast is known to be an important cue for human saliency, and therefore, saliency in those regions where a small value of γ has been used might be lost.

The idea pursued in this work is to modify the GRA-RGB [Zamir et al., 2014] to obtain a gamut mapped image with a saliency map closer to that of the original image. SBGRA is divided into three parts: detection of the regions where saliency is lost, creation of a new map Γ_{SGM} of γ values, and combination of the initial gamut mapped image (obtained using the GRA of [Zamir et al., 2014]) and the one obtained by using the new Γ_{SGM} map. Figure 5.20 presents the work-flow of our SBGRA which is explained in the following sections.

5.2.2 Detecting Regions of Lost Saliency

To compute the saliency of an image, in this work we use the saliency by induction mechanisms (SIM) method [Murray et al., 2013]. Let us call I_{or} the original image and I_{GM} the gamut mapped image obtained by using our GRA-RGB [Zamir et al., 2014]. We can measure the amount of saliency loss at each pixel by

$$L(I_{GM}, I_{or})(x) = \begin{cases} \dfrac{SIM(I_{or})(x)}{max(SIM(I_{or}))} - \dfrac{SIM(I_{GM})(x)}{max(SIM(I_{or}))} & \\ & \text{if} \quad SIM(I_{or})(x) > SIM(I_{GM})(x) \\ 0 & otherwise \end{cases} \tag{5.10}$$

In order to obtain binary regions to work with, we perform a threshold operation on $L(I_{GM}, I_{or})$ as

$$L_{thresh}(I_{GM}, I_{or})(x) = \begin{cases} 1 & \text{if} \quad L(I_{GM}, I_{or})(x) > \epsilon_1 \\ 0 & otherwise \end{cases} \tag{5.11}$$

where ϵ_1 is a real number. We are interested in those regions where the saliency might have been lost due to a large difference between their gamma values and the gamma values of their surrounds. Therefore, to refine those regions given by L_{thresh} we compute a new map $R(\cdot)$ as

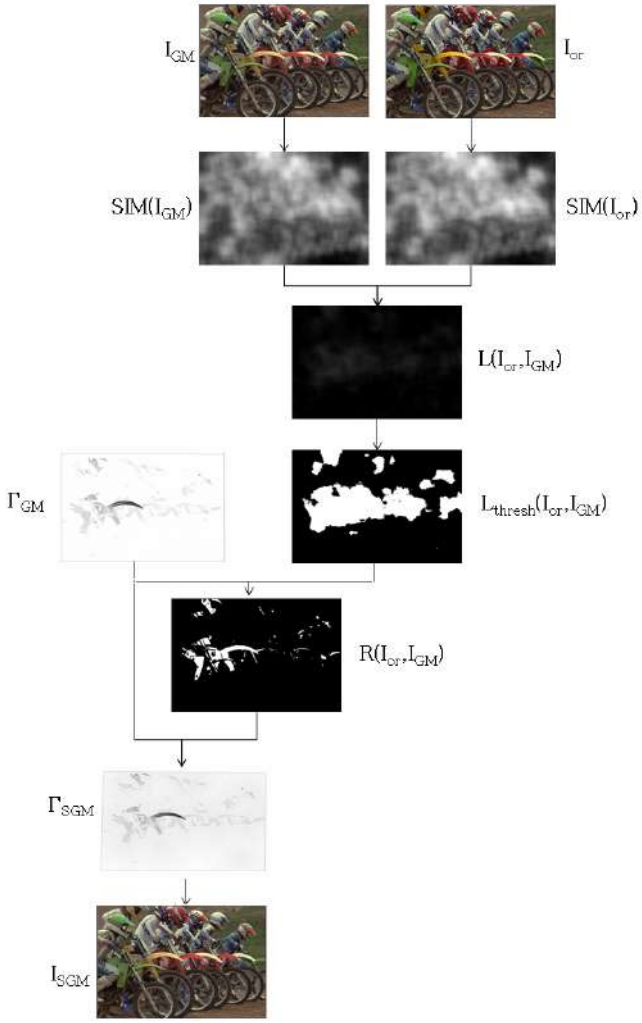


Figure 5.20: Workflow of our method. From the original image and the output of GRA-RGB [Zamir et al., 2014] method we obtain the saliency maps and the difference between them. We binarize this image and combine it with the γ value map obtained using GRA-RGB [Zamir et al., 2014], thus defining the set of interesting regions. Then, we compute the new map and apply the evolution Eq. (5.7) in order to generate a new gamut mapped image.

$$R(I_{GM}, I_{or})(x) = \begin{cases} 1 & \text{if} \quad L_{thresh}(I_{GM}, I_{or})(x) \cdot |\Gamma_{GM}(x)| > \epsilon_2 \\ 0 & otherwise \end{cases} \tag{5.12}$$

where ϵ_2 is a real number. In other words, we select only those regions where the value of γ is small enough so that the surrounds can be decreased.

The map $R(\cdot)$ is already defining those regions that we need to deal. However, we have to be sure that the contours of the regions are consistent with the original image data. To that end, we fill in the map $R(I_{GM}, I_{or})$ with the adjacent pixels that have a similar γ value.

5.2.3 Creating New γ Map

Our goal here is to reduce the difference in the γ values between the regions where the saliency has been lost and their surrounds. But at the same time we do not want to sacrifice reproduction accuracy to a great extent. Therefore, SBGRA takes into account the distance of each pixel from a region of interest. By taking this into consideration, our new map Γ_{SGM} is defined as

$$\Gamma_{SGM}(x) = \begin{cases} \Gamma_{GM}(x) + \tau \cdot (1 - d(x, R)) \cdot max(R(I_{GM}, I_{or}) \cdot \Gamma_{GM}) \\ \qquad\qquad\qquad\qquad\qquad\qquad\qquad \text{if} \quad R(x) = 0 \\ \Gamma_{GM}(x) + \frac{\tau}{2} \cdot max(R(I_{GM}, I_{or}) \cdot \Gamma_{GM}) \qquad otherwise \end{cases} \tag{5.13}$$

where $d(x, R)$ is the distance between the pixel x and its closest region of interest in R (that is, the closest pixel with $R = 1$) and τ is a real number between 0 and 1 that controls the decrements in the surrounding pixels.

Then, from the new map Γ_{SGM} we can run evolution Eq. (5.7) for each of the values and obtain a new gamut mapped image I_{SGM}.

5.2.4 Combining Gamut Mapped Images

Let us note here that the image I_{SGM} obtained by the map Γ_{SGM} will have a saliency similar to the original image in our regions of interest, but may lose some

Table 5.5: Saliency difference between the original image and the gamut mapped image.

Images [Kodak, 1993]	GRA-RGB ($\times 10^6$)	SBGRA ($\times 10^6$)	Improvement (%)
Caps	3.02	2.67	11.77
Raft	2.63	2.59	1.71
Barn	2.03	1.98	2.61
Girl	1.10	1.06	3.51
Birds	2.56	2.53	1.22
Motorbikes	3.24	2.79	14.20
Boat	1.58	1.55	1.67
Beach	0.77	0.67	13.22
Party	8.20	8.14	0.81
Portrait	7.85	7.37	6.17
Picnic	10.90	9.29	15.27
Window	4.02	3.95	1.86
Woman with Hat	3.32	3.27	1.49
Sailing Boats	0.84	0.84	0
Statue	3.08	2.98	3.32
Model	5.76	5.70	1.09
Ski	2.88	2.62	8.99

saliency in other regions. Thus, the last step of our SBGRA combines both I_{GM} and I_{SGM}. Our final image is then defined as

$$I_{final}(x) = \begin{cases} I_{SGM}(x) & \text{if} \quad SIM(I_{SGM})(x) > SIM(I_{GM})(x) \\ I_{GM}(x) & otherwise \end{cases} \tag{5.14}$$

5.2.5 Results

We apply SBGRA on 17 still images that are presented in Fig. 5.8. After testing different values of τ we select the image that better preserves saliency. $d(x, R)$ is defined as the normalization between 0 and 1 of the Euclidean distance in the pixel domain. ϵ_1 is obtained by using the method of [Otsu, 1979], while ϵ_2 is defined as the value given by the Otsu's method divided by 4.

In order to measure the performance of SBGRA we define a measure of

Table 5.6: CID error for GRAs.

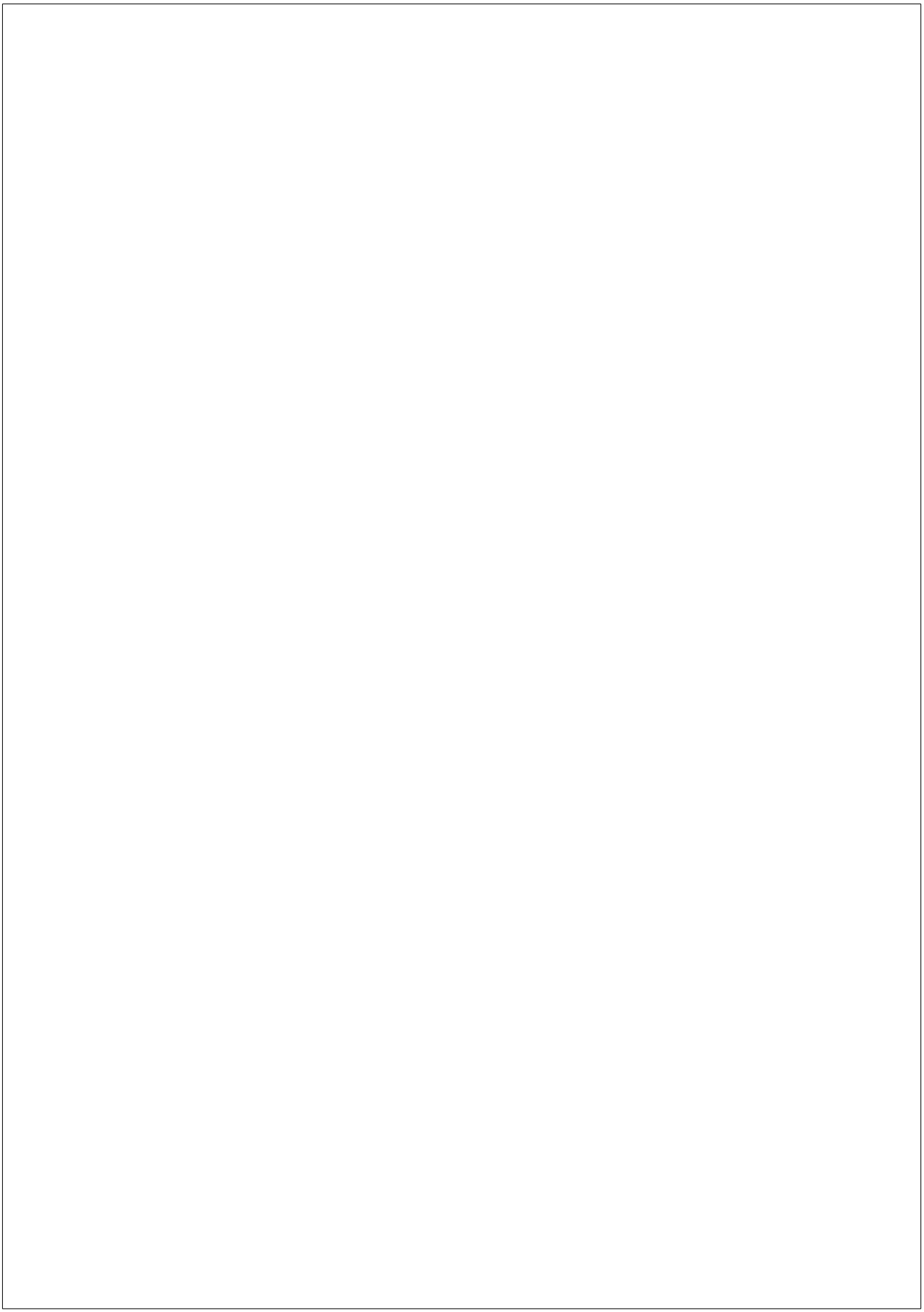
Methods	Mean	Median	RMSE
Our SBGRA	0.0414	0.0299	0.0576
Our GRA-RGB	0.0360	0.0280	0.0485
HPMINDE [Murch and Taylor, 1989]	0.0674	0.0514	0.0873
[Alsam and Farup, 2012]	0.0472	0.0398	0.0627
[Lau et al., 2011]	0.0665	0.0695	0.0807

saliency difference. This measure compares the saliency of a gamut mapped image to the saliency of the original image as follows

$$S(I_{GM}, I_{or}) = \sum_{x} |(SIM(I_{or})(x)) - (SIM(I_{GM})(x))| \tag{5.15}$$

Based on this formula, we compare the results of SBGRA to those of GRA-RGB, and the saliency differences are presented in Table 5.5. We can see that there is only one image where SBGRA fails to improve the saliency. In particular, this method achieves a remarkable improvement of more than 10% in 4 of the images.

Once it is proven that the SBGRA does reduce the saliency difference with respect to the original image, a new question arises: does this improvement come at the cost of bad imaqe quality? To answer this question, we use the Color Image Difference (CID) metric [Lissner et al., 2013]. Results of the CID metric in the form of mean, median and root mean squared error (RMSE) for the 17 images are presented in Table 5.6. In this Table we compare our SBGRA to our GRA-RGB method as well as to other methods: HPMIMDE [Murch and Taylor, 1989], [Alsam and Farup, 2012], and [Lau et al., 2011]. Results for SBGRA are slightly poorer than those obtained by the GRA-RGB, but they are still better than the results of all other competing methods. In other words, we achieve better saliency reproduction at the expense of a slight loss in image quality, but this loss is affordable since none of the other methods outperforms our SBGRA.



CHAPTER 6

Gamut Extension in CIELAB Color Space

This chapter presents some of our contributions to gamut extension problem. The text of this chapter is derived from our publication in Electronic Imaging conference titled "Gamut extension for cinema: psychophysical evaluation of the state of the art, and a new algorithm" [Zamir et al., 2015], and our article published in the IEEE Transactions on Image Processing titled "Gamut extension for cinema" [Zamir et al., 2017].

6.1 One-Shot Gamut Extension Algorithm (GEA-LAB1)

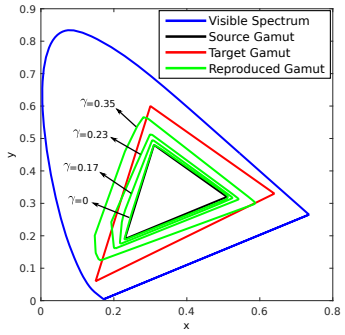
Our GEA-RGB [Zamir et al., 2014] presented in the previous chapter (section 5.1.1), due to its inherent behavior of expanding colors by increasing the contrast of the image, produces results with over-enhanced contrast, which in turn makes a few colors go towards black (loss of saturation), as it is visible in Fig. 6.1b. It can be seen that the overall contrast of the reproduction is increased noticeably, making it depart from the original image. Also, the over-enhancement of contrast causes loss of color details as it is shown in the area highlighted by a bounding box in Fig. 6.1b. To overcome these problems, we propose a new GEA based on [Zamir et al., 2014]; this method performs gamut extension using the same energy functional (Eq. 5.6) and its corresponding evolution equation (Eq. 5.7) as used by [Zamir et al., 2014], but uses the CIELAB color space. This key modification eliminates not only the problems with saturation and contrast in the reproduced images, but also the need to perform any sort of preprocessing as it

(a) (b) (c)

Figure 6.1: Gamut extension example. (a) Input image. (b) Result of GEA-RGB [Zamir et al., 2014]. (c) Result of GEA-LAB1 [Zamir et al., 2015]. Original image is courtesy of [Kodak, 1993].

was the case with the GEA-RGB [Zamir et al., 2014].

To perform gamut extension, our GEA-LAB1 [Zamir et al., 2015] first converts the RGB input image to the CIELAB color space, and then only maximizes the contrast of the chromatic components 'a' and 'b' using Eq. (5.7), while keeping the lightness component constant. To show how the evolution Eq. (5.7) extends the color gamut, an example with several different gamuts (visible spectrum, source gamut, target gamut and reproduced gamut) on a chromaticity diagram is shown in Fig. 6.2a. It is important to note that for each set of values for α, β, and γ the evolution Eq. (5.7) has a steady state. For example, it is shown in Fig. 6.2a that when $\beta = 1$, $\alpha = 0$, and $\gamma = 0$ we obtain the original image as the steady state of the evolution equation. Moreover, it can be seen in the same figure that as we increase γ the steady state of Eq. (5.7) has a gamut which is gradually larger. Fig. 6.2a also shows that the colors of the source gamut can be expanded to the destination gamut just by using a large enough value for γ ($\gamma = 0.35$ in this case). And to select an adequate γ value, we keep increasing the γ value and running evolution Eq. (5.7) to steady state until the gamut of the input image exceeds the target gamut up to a certain threshold T. This threshold T defines a stopping criteria according to which if $T\%$ pixels of the original image move out of the target gamut we should stop performing extension. Additionally, the threshold T controls the amount of saturation; a large value of T provides a higher level of saturation, whereas a small value of T produces a less saturated output. For each γ value, the corresponding reproductions are shown in Fig. 6.2b. After this, the colors that were placed outside the target gamut in previous iterations are mapped back inside using our GRA-RGB [Zamir et al., 2014].



(a)

(b)

Figure 6.2: Gamut extension approach. (a) Gamuts on chromaticity diagram. (b) Gamut extension results. From left to right: input image ($\gamma = 0$), gamut extended image with $\gamma = 0.17$, $\gamma = 0.23$ and $\gamma = 0.35$. As the γ value increases the gamut becomes larger; notice the increment in saturation of helmets, socks, ski suits and shoes. Original image is courtesy of [Kodak, 1993].

6.2 Gamut Extension Algorithm Driven by Hue, Saturation and Chroma Constraints (GEA-LAB2)

Since our GEA-LAB1 [Zamir et al., 2015] used a fixed value of threshold T for all the images, the results were often off from the ground truth. Moreover, the GEA-LAB1 [Zamir et al., 2015] causes hue shifts in some of the reproductions. In this section we present a new method that, unlike GEA-LAB1 [Zamir et al., 2015], works iteratively with added constraints to perform gamut extension in terms of the contrast coefficient γ. The general structure of our algorithm (GEA-LAB2), that also operates on the 'a' and 'b' components in the CIELAB color space, is as

follows.

At each iteration, we run the evolution Eq. (5.7) for some particular α, β, and γ (in the first iteration the values are $\beta = 1$, $\alpha = 0$, and $\gamma = 0$) until we reach the steady state. For each pixel of this steady state image we check that it simultaneously satisfies three constraints on saturation, hue and chroma: if it does not, then the pixel is marked as "done" and subsequent iterations will not modify its value (i.e. this pixel is now part of the final output, the gamut extended image). We move to the next iteration, by slightly increasing the value of γ and setting $\alpha = \frac{\gamma}{20}$. We run again evolution Eq. (5.7) until steady state and check whether any of the pixels violated any condition: those pixels are left unmodified for the following iterations. We keep repeating this process until the gamut of the original image exceeds the destination gamut up to a threshold T, at which point the iterations stop and all pixel values are final except for colors that now lie outside the destination gamut: these are mapped back inside using the GRA-RGB [Zamir et al., 2014].

In the following sections we describe in detail the constraints that are checked after each iteration and the way the destination gamut is constructed.

Constraints

These are the three constraints that pixels must satisfy in order to continue with the iterative process, otherwise we leave them untouched for the following iterations:

- *Saturation Constraint*: We define a condition to avoid pixels from getting desaturated, therefore each pixel must satisfy this constraint

$$S_{ae} - S_{be} > 0, \tag{6.1}$$

  where S_{ae} and S_{be} denote saturation after extension and saturation before extension, respectively.

- *Hue Constraint*: The human visual system is highly sensitive to the changes in hue, and one major goal in GE is to produce images with unnoticeable hue shifts. In order to make our GEA-LAB2 robust against unpleasant color shifts, we make each pixel to comply with the hue constraint

$$|\theta_{ae} - \theta_{be}| < \epsilon_h, \tag{6.2}$$

where θ_{ae}, θ_{be} and ϵ_h represent hue angle after extension, hue angle before extension and hue threshold, respectively.

- *Chroma Constraint*: One of the main challenges in GE is that, after extension, colors of less saturated natural objects, skin tones, and shades of memory colors may appear too colorful and unrealistic. To deal with it, we manually segment a set of images that contain shades of sky and skin colors. We then define the crucial region on the chromaticity diagram in such a way that it encloses the chromaticity values of all the colors of these segments of images. The crucial region is shown in Fig. 6.6, and colors that fall inside this region should satisfy this chroma condition

$$|C_{ae} - C_{be}| < \epsilon_c, \tag{6.3}$$

where C_{ae}, C_{be} and ϵ_c indicate chroma after extension, chroma before extension and chroma threshold, respectively.

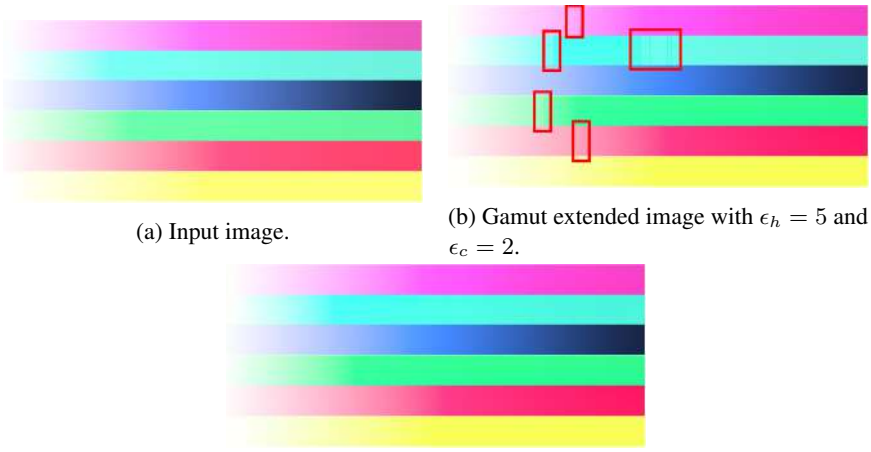
We want to mention that the values of the parameters ϵ_h and ϵ_c are of great importance; an example of an image with smooth color gradients is shown in Fig. 6.3 where it is evident that too small values of ϵ_h and ϵ_c produce an image with noticeable artifacts, whereas for slightly larger values of ϵ_h and ϵ_c the reproduction is free from false edges but may have hue shifts. Therefore, these values should be adjusted in a way that the artifacts (false edges and hue shifts) stay below the visibility level. In the experiment section we propose values for these parameters that are suitable even for large differences between source and destination gamuts.

Scaled Destination Gamut Computation

One of the problems in gamut extension is that the majority of GEAs are either image dependent or perform well for a few combinations of source and destination gamuts [Morovič, 2008]. To reduce this dependency issue, we compute a scaled version of the original image gamut and use it as the destination gamut.

Given two points p_0 and p_1, the following parametric representation describes the line passing through them

$$p = p_0 + \eta(p_1 - p_0), \tag{6.4}$$



(a) Input image.

(b) Gamut extended image with $\epsilon_h = 5$ and $\epsilon_c = 2$.

(c) Gamut extended image with $\epsilon_h = 15$ and $\epsilon_c = 4$.

Figure 6.3: Effect of parameters ϵ_h and ϵ_c. The red bounding boxes indicate the regions where artifacts (false edges) appear.

where η is a scaling factor and $\eta > 1$ will provide a point p on the line further from p_0 than p_1.

To obtain the scaled target gamut, given an original image I_0 and a target gamut (TG), we first convert the image colors into luminance (Y) and chromaticity values (x and y) [Sharma, 2002]. Next we calculate a reference point p_{r1} by taking the mean of those chromaticity points (x and y) that make the two-dimensional convex hull of the image I_0. Both the reference point p_{r1} and the convex hull of I_0 are shown in Fig. 6.4a. We then define a set of lines (L) by substituting the reference point $p_0 = p_{r1}$ and each vertex of the convex hull of I_0 as p_1 in Eq. (6.4). Finally, we generate new points (one from each line in L) using a same value of scaling factor ($\eta_1 > 1$) that is chosen in such a way that none of the new points falls outside the TG and at least one of them touches the boundary of the TG as shown in Fig. 6.4a. Similarly, we calculate another scaling factor η_2, but this time using the mean of all chromaticity values of the image as reference point p_{r2} and vertices of the convex hull of I_0.

Once we have the scaling factors η_1 and η_2, we apply them using Eq. (6.4) on the xyY triplets that make the three-dimensional (3D) convex hull of the original

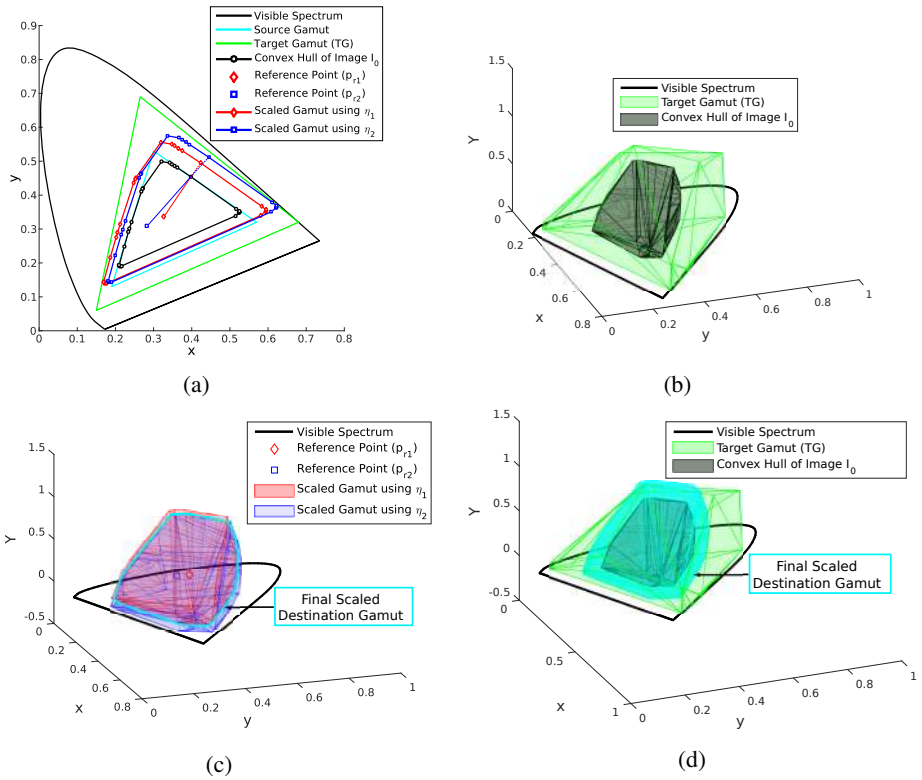


Figure 6.4: Scaled destination gamut computation.

image I_0 (shown in Fig. 6.4b) to obtain two 3D scaled gamuts. Note that these two scaled gamuts, depicted in Fig. 6.4c, are obtained using two reference points (p_{r1} and p_{r2}); one of which is the mean of the xyY points that make the 3D convex hull of the input image, and the other is the mean of all xyY values of the image. Finally, we create the final scaled destination gamut by computing the intersection of both 3D scaled gamuts as illustrated in Fig. 6.4c. An example with all the relevant gamuts is shown Fig. 6.4d.

6.3 Qualitative Experiments and Results

6.3.1 Methodology

One major goal of this study is to investigate psychophysically the performance of GEAs in cinematic conditions using a digital cinema projector (Barco DP-1200 [Barco, 2007]). A GEA must respect as much as possible the artist's vision, and in this sense it is completely different from a process like colorization of black and white pictures. In general, and not only in the case of legacy material, there never is a ground truth. The original material in a reduced gamut is all there is, and this material has to have its gamut extended while preserving as much as possible its original appearance. In our tests we actually have a ground truth because for evaluation purposes we start from a wide gamut material, reduce its gamut, extend it with a GEA and then ask users to compare the result with the wide gamut ground truth, but in general a wide-gamut ground truth is never available. Nonetheless, by showing that a GEA performs well as evaluated using ground truth data, we expect that it also performs well when ground truth data is not available, which is a most common approach in image processing (e.g. denoising algorithms are evaluated in terms of PSNR by comparing with a ground truth "clean" image that is never available in real scenarios, segmentation algorithms are evaluated by comparing their results to those of manual segmentations, etc.). One could question the choice of evaluation criteria: why ask users to choose the most accurate result instead of the one they find most pleasant? The reason is that, for a GE technique to be adopted by the movie industry, it must yield gamut extended results that preserve as much as possible the artistic intent of the content's creator. Designing GEAs for pleasantness does not guarantee this, usually quite the opposite: apart from general inter-subject variability in terms of preference, there is also a strong correlation between colorfulness and perceived image quality [Fedorovskaya et al., 1997], so user tests based on subject preference would rank higher those GEAs that increased color saturation even if that implied a departure from a subdued color palette chosen by the movie artist. User tests based on accuracy, on the other hand, are much less subjective (as users are simply asked to estimate differences with respect to the provided reference) and the preservation of the artistic intent is in a sense "built-in" in the procedure, since the ground truth acts as a stand-in for the content's creator intent.

We show all the stages of our experimental framework in Fig. 6.5. The first

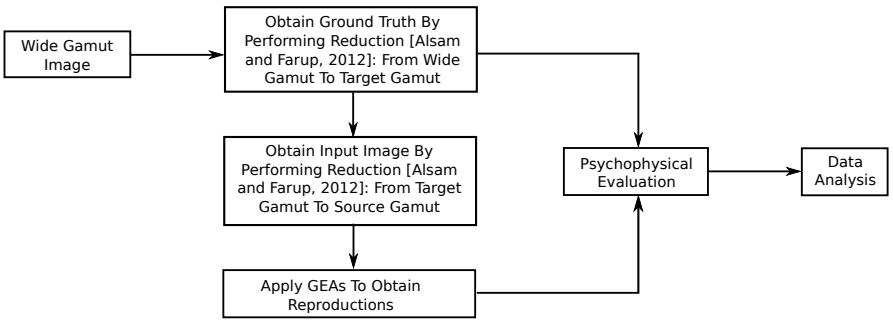


Figure 6.5: A schematic of the evaluation for GEA-LAB2.

task is to obtain both the wide-gamut ground truth images and the limited-gamut input images from the given wide-gamut test images. The wide-gamut test images may have colors that fall outside the gamut of our cinema projector, therefore, to create the ground truth, we map the colors of the test images to the gamut of the projector by using the state of the art gamut reduction algorithm of [Alsam and Farup, 2012]. It is important to note that in our experiments we make use of a cinema projector that has the same red and green primaries as those of the DCI-P3 standard but has a slightly different blue primary. Nonetheless, both the DCI-P3 gamut and the projector's gamut cover almost the same amount of area on the chromaticity diagram as it can be seen in Fig. 6.6. Therefore, we refer to the projector's gamut as DCI-P3 in the rest of this chapter for the sake of simplicity. Next, to create the limited-gamut input images, we apply the gamut reduction method of [Alsam and Farup, 2012]. Once we have the input images ready, we apply to them our GEA-LAB2 and four competing GEAs [Laird et al., 2009] to generate reproductions for the following two different experimental setups:

1. *Mapping from small gamut to DCI-P3 gamut:* as laser displays with their extended gamut capabilities are becoming popular, in the near future the common case will be to have large color differences between the standard gamuts and displays' gamuts. Therefore, we create this setup to investigate the behavior of GEAs when the difference between source and target gamut is large. To this end, we map the source images from the small 'Toy' gamut (slightly smaller than the BT.709 gamut, see primaries in Table 6.1) to the large DCI-P3 gamut. On the chromaticity diagram, the difference in

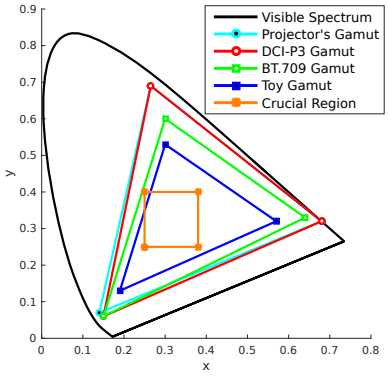


Figure 6.6: Gamuts on chromaticity diagram.

Table 6.1: Primaries of gamuts.

Gamuts	Red Primaries		Green Primaries		Blue Primaries	
	x	y	x	y	x	y
BT.2020	0.708	0.292	0.170	0.797	0.131	0.046
BT.709/sRGB	0.640	0.330	0.300	0.600	0.150	0.060
DCI-P3	0.680	0.320	0.265	0.690	0.150	0.060
Projector	0.680	0.320	0.265	0.690	0.140	0.070
Toy	0.570	0.320	0.300	0.530	0.190	0.130

gamuts for this setup is almost equal to the difference between BT.709 and BT.2020.

2. *Mapping from BT.709 to DCI-P3 gamut:* in this setup we mimic the practical situation where the source material has BT.709 gamut and we map the source colors to the colors of the DCI-P3 gamut.

Next we performed a subjective evaluation (for both experimental setups) with 15 observers of which ten were male and five female and their age was between 27 and 44 years, with average of 32 years. All observers were tested for normal color vision using the Ishihara color blindness test. To mimic the lighting conditions of a real cinema, we created a low light ambience where the illumination measured

at the screen was around 750 lux and the ambient illuminance was 1 lux. During the experiment there was not any strong colored object in the field of view of the observers. The glare-free screen used in our experiments was 3 meters wide and 2 meters high. Each observer was instructed to sit approximately 5 meters away from the screen.

In the psychophysical experiments, we used a forced-choice pairwise comparison technique where observers were shown three images on the projection screen: the target gamut ground truth image (in the middle) and a pair of reproductions (one image on the left side and the other on the right side of the ground truth). For both setups, observers were asked to make selections according to the instructions: a) if there are any sort of artifacts in one of the reproductions, choose the other, and b) if both of the reproductions have artifacts or are free from artifacts, choose the one which is perceptually more closer to the ground truth. Since there were 30 test images, 5 algorithms and 2 experimental setups, each participant had 600 comparisons to judge in total, but to avoid fatigue, we split these comparisons in four sessions that were conducted on four different days. Although there was no time restriction to make choices, each observer took approximately 30 minutes to complete one session. Finally, we analyze the psychophysical data using the same work as used by [Morovič, 1998] (see chapter 5 of his thesis), that is based on Thurstone's law of comparative judgement [Thurstone, 1927].

Moreover, in order to validate the robustness of the competing GEAs explicitly, 9 experienced observers (who belong to the image processing community and participated in various psychophysical studies) were shown a pair of images on the projection screen; one of which was the ground truth and the other was the reproduction. We asked them to examine artifacts and hue shifts in the reproductions as compared with the originals.

6.3.2 Creation of Wide Gamut Test Images

To the best of our knowledge, there is not any publicly available colorful wide-gamut dataset that can be used to assess the performance of GEAs. Therefore, we create wide-gamut images using a camera that is capable of capturing images in RAW format. Once we have RAW images, we associate with them the wide-gamut color space (ProPhoto RGB) to obtain true color images in Adobe Lightroom. Along with the 21 images shown in Fig. 6.7, we use 9 other test images that come from professional feature films such as 'Amazing Spider-Man', '127 Hours',

Figure 6.7: Some of the wide-gamut images used in our tests. Note that only the central part of the images is shown.

'Tangled', 'Rio', and 'Requisitos para ser una persona normal'.

6.3.3 Settings for Proposed Method

We work in CIELAB color space and the parameter values that we use in all our experiments are $\beta = 1$, $\epsilon_h = 15°$, $\epsilon_c = 4$, $\Delta t = 0.10$, $T = 1\%$ of the total number of pixels, and the gamma increment of $\Delta\gamma = 0.01$ between two consecutive iterations. The proposed GEA in a non-optimized MATLAB implementation running on a machine with 8 cores 3.4-GHz CPU takes (on average) 4.5 minutes to process an image of resolution 656×1080 pixels.

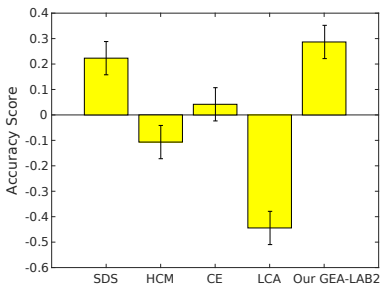
6.3.4 Results

In order to compute the accuracy scores from the raw psychophysical data, we use the same data analysis procedure as used by [Morovič, 1998] in his gamut mapping study. The analysis of psychophysical data of the setup 1 presented in

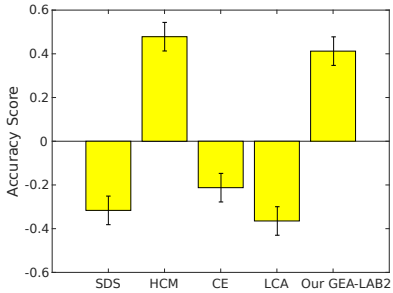
Fig. 6.8a shows that, when there is a large difference among the source-target gamut pair, our GEA-LAB2 produces images that are perceptually more faithful to the originals as compared with the other competing algorithms. The four state of the art GEAs [Laird et al., 2009] that we compare with our algorithm are briefly described in section 4.2.1 of chapter 4. The observers declared LCA [Laird et al., 2009] as the least accurate algorithm, whereas both the CE [Laird et al., 2009] and the HCM [Laird et al., 2009] algorithms rank third and fourth, respectively.

In Fig. 6.9, we report the percentage of reproductions for which 9 experienced observers, on average, noticed the visual distortions: artifacts or hue shifts. For the setup 1, the LCA and CE algorithms produce images with loss of texture due to over-saturation and it can be seen in Fig. 6.9a that the subjects noticed artifacts in 25% of the reproductions obtained using the LCA algorithm and in 12% of the images in the case of the CE algorithm. The observers confirmed that the GEA-LAB2 produces images with very low error rate, around 2%. One such example is shown in Fig. 6.10 where it is clearly visible that the colors reproduced by our method are artifact free and perceptually more faithful to the ground truth than those of LCA and CE algorithms. Moreover, it can also be seen in the said figure that the LCA and CE algorithms turn subtle spatial color details of the woman's apron into noticeable unpleasant color gradients.

Even though both the SDS [Laird et al., 2009] and HCM algorithms do not introduce much noticeable artifacts, their reproductions show strong hue shifts. We show an example in Fig. 6.11 where the hue shifts are evident on the floor in

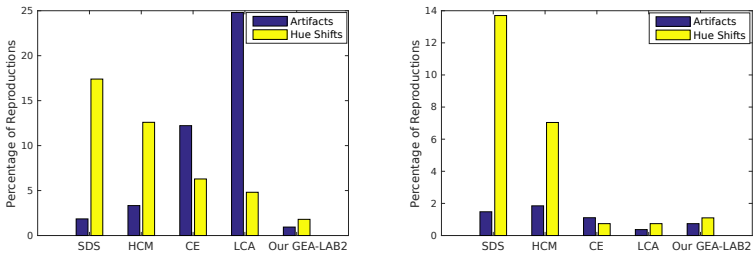


(a) Setup 1.



(b) Setup 2.

Figure 6.8: Accuracy scores using 15 observers and 30 images.



(a) Setup 1. (b) Setup 2.

Figure 6.9: Percentage of reproductions in which 9 experienced observers noticed visual distortions.

(a) (b) (c) (d)

(e) (f) (g) (h)

Figure 6.10: Example of artifacts. (a) Ground Truth. (b) Output of CE algorithm [Laird et al., 2009]. (c) Output of LCA algorithm [Laird et al., 2009]. (d) Output of our GEA-LAB2. (e)-(h) Zoomed-in view of the regions cropped from the top row. Note that these are wide-gamut images where out-of-sRGB pixels are masked green.

the results of the SDS and HCM methods. Note that the images depicted in Fig. 6.10 and Fig. 6.11 were in DCI-P3 format, originally. Since we are limited by

sRGB standard (that has the same primaries as of BT.709) for the paper, in order to present results we show only those colors that are inside the sRGB gamut and mask (in green) the rest of the colors.

In Fig. 6.8b we show the results of setup 2 where it can be seen that, when the difference between source and target gamut is smaller, the ranking order of the GEAs changes dramatically. The HCM algorithm that ranked as the second least accurate in the previous setup becomes the most accurate method. Our GEA shows comparable performance with the HCM algorithm. Similar to the results of setup 1, the LCA algorithm produces gamut extended images that are least faithful with the original content. It is also evident from Fig. 6.8b that the SDS method, unlike in setup 1, shows poor performance under setup 2. Fig. 6.9b shows that the SDS and HCM algorithms produce gamut extended images with strong color shifts for 13.6% and 7% of the input images, respectively. It can be seen in the same figure that none of the competing algorithms produces images

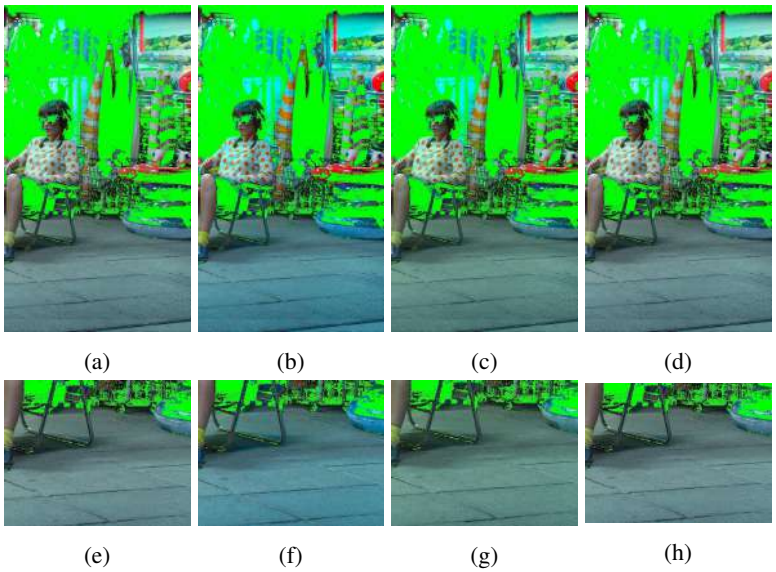


Figure 6.11: Example of hue shifts. (a) Ground Truth. (b) Output of SDS algorithm [Laird et al., 2009]. (c) Output of HCM algorithm [Laird et al., 2009]. (d) Output of our GEA-LAB2. (e)-(h) Zoomed-in view of the regions cropped from the top row. Note that these are wide-gamut images where out-of-sRGB pixels are masked green.

with distinct visual artifacts for setup 2, in which there are small color differences between source and target gamut.

Trends for both experimental setups show that our GEA-LAB2 is the most consistent and reliable method for both small and large color differences among the source-target gamut pair.

Temporal Consistency Test

In order to examine the temporal consistency of GEAs, we conducted a psychophysical study with 9 experienced observers and two colorful image sequences with different levels of motion. The representative frames for both image sequences are shown in Fig. 6.12. In this experiment, the gamut extended videos obtained using different GEAs were shown in isolation (without any ground truth) to each observer and they were asked to inspect the following attributes: temporal color consistency (objects should retain the same hue, chroma and brightness), global flickering, local region flickering, and excessive noise. None of the observers noticed any temporal artifacts, which supports our choice to apply all competing GEAs on each frame independently. Finally, we want to stress that the quality of the input video is of high importance; if it contains any spatial artifacts due to compression or noise they may become prominent in the reproduced video.

(a) Image sequence 1

(b) Image sequence 2

Figure 6.12: Representative frames of image sequences with toy gamut.

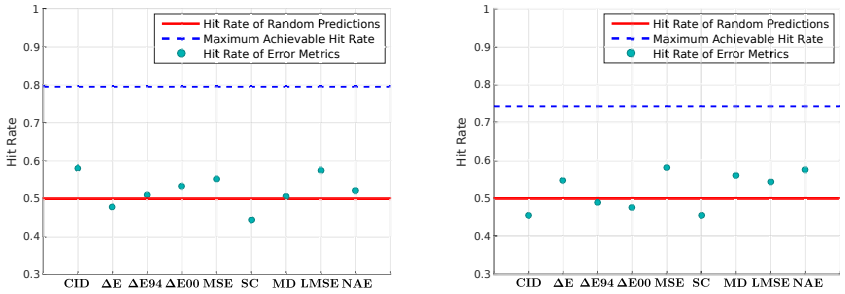
6.4 Identification of Error Metric Suitable for the GE Problem

We would like to find out whether or not there exists an error metric that is suitable for the gamut extension problem. First, we examine the prediction quality of nine error metrics: CIEΔE, CIEΔE94 [Berns, 1993], CIEΔE2000 [Luo et al., 2001], color image difference (CID) [Lissner et al., 2013], and the ones presented in [Eskicioglu and Fisher, 1995] such as mean square error (MSE), structural content (SC), maximum difference (MD), Laplacian mean square error (LMSE) and normalized absolute error (NAE). We test the efficiency of the GEAs for setup 1 and setup 2 using these error metrics, that find distortions between ground truth and reproductions. We list the overall error for all the images in Table 6.2. There it can be seen that for setup 1, only two out of nine error metrics selected the observers' preferred GEA (our proposed algorithm) as the best performing algorithm. Both the LCA and the HCM algorithms were chosen as the most accurate methods by three different error metrics, whereas none of the tested metrics picks SDS as a good performing method: all these results contradict the findings of the psychophysical experiment.

For setup 2, five out of the nine error metrics tested declare HCM as the most efficient algorithm, which is consistent with the results of the psychophysical experiment. However, the same image quality metrics rank other GEAs very

Table 6.2: Predictions of image quality measures: error across all images. Note that the values of ΔE, ΔE94 and ΔE00 are normalized in the range [0,1].

Error Metrics	Setup 1: Toy to DCI-P3 gamut					Setup 2: BT.709 to DCI-P3 gamut				
	SDS	HCM	CE	LCA	Our GEA	SDS	HCM	CE	LCA	Our GEA
CID	0.0903	**0.0858**	0.0862	0.1013	0.0987	0.0645	**0.0624**	0.0658	0.0657	0.0681
ΔE	0.1926	0.1766	0.1636	**0.1396**	0.1796	0.1603	**0.1401**	0.1679	0.1680	0.1896
ΔE94	0.2941	0.2684	0.2527	**0.2203**	0.2766	0.2203	**0.1927**	0.2141	0.2142	0.2194
ΔE00	0.3058	0.2802	0.2489	**0.2315**	0.2670	0.2302	**0.2037**	0.2317	0.2301	0.2326
MSE	0.0060	**0.0059**	0.0109	0.0144	0.0116	0.2034	0.2192	0.1318	0.1317	**0.1205**
SC	0.6157	0.6052	0.5449	0.5212	**0.5037**	**0.4885**	0.4933	0.5268	0.5269	0.5209
MD	0.3442	**0.3435**	0.5096	0.5853	0.4393	0.3698	0.3172	0.3644	0.3651	**0.2586**
LMSE	0.3133	0.3154	0.2361	0.3124	**0.2093**	0.2034	0.2192	0.1318	0.1317	**0.1205**
NAE	0.1005	0.0967	**0.0953**	0.0977	0.1031	0.0636	**0.0503**	0.0748	0.0747	0.0640



(a) Hit rate for setup 1

(b) Hit rate for setup 2

Figure 6.13: Hit rates obtained by image quality metrics for experimental setups.

differently as compared with the observers' choices. Another notable finding is that the CIEΔE metric and its variants (CIEΔE94 and CIEΔE2000) show a similar trend and pick LCA and HCM as the most efficient algorithms for setup 1 and setup 2, respectively.

Even tough some error metrics were able to make right predictions, we check the correlation between the choices made by our observers in the psychophysical experiments and the predictions of the error metrics to confirm the reliability of these image metrics. To achieve this, we use the hit rate (h) approach that is defined as

$$h = \frac{c}{N} \tag{6.5}$$

where N is the total number of choices for any experimental setup and c is the number of correctly predicted choices. We operate on the raw visual data, and as we know the choice of the observer for each paired comparison in the psychophysical experiment, we consider a choice is correctly predicted if an image quality metric gives less error for the image selected by the observer. Any quality measure that makes random predictions is expected to achieve a hit rate of $h = 0.5$. Since each image pair is compared by several observers and considering that not all of them have made the same choice, we computed the *maximum achievable hit rate* to which an error metric will approach if its predictions are in agreement with the majority of observers' choices for all image pairs.

Since our purpose is to find a metric that works for both small and large color

differences between source and target gamuts, in Fig. 6.13 we present the hit rates to validate the image quality metrics for both setup 1 and setup 2. It can be seen that the error metrics made, most of the time, random predictions and their hit rates are significantly lower than the maximum achievable hit rate. Note that if an error metric has a very low hit rate (i.e. less than 0.2), one can make it a good image quality measure by inverting its predictions for all image pairs. However, none of these metrics achieves either a very low hit rate or a very high hit rate and therefore they are unreliable to be used for evaluating results in the gamut extension problem.

Without an efficient error metric, gamut extension cannot be posed as an optimization scheme as in [Nakauchi et al., 1999], [Preiss et al., 2014] where, given a reference image and its gamut mapped version, the idea is to alter the gamut mapped image to minimize the distortions as computed by an objective metric. Consequently, it leaves us with only one option for the moment to validate GEAs, that is to perform subjective studies which are cumbersome and time consuming.

CHAPTER 7

Gamut Mapping Algorithms in HSV Color Space

The main limitations of the GEA-LAB2 [Zamir et al., 2017] presented in the previous chapter are, on one hand, its sensitivity to the correct choice of parameter values and, on the other hand, its significant computational cost: a non-optimized MATLAB implementation running on a machine with 8 cores 3.4-GHz CPU takes (on average) 4.5 minutes to process an image of resolution 656×1080 pixels.

In this chapter, we propose new GMAs (a GEA and a GRA) that are significantly faster than our previous methods. First, we are going to present a GEA published as a conference paper titled "Perceptually-based gamut extension algorithm for emerging wide color gamut display and projection technologies" [Zamir et al., 2016]. After this, we will provide details and results of our GRA.

7.1 Gamut Extension Using Kernel Based Retinex (GEA-KBR)

Since the human visual system is very sensitive to changes in hue, it is recommended in the gamut mapping literature to leave the image hues unmodified, whenever possible [CIE, 2004], [Morovič, 2008]. Therefore, we introduced in [Zamir et al., 2016] a new GEA that works in the HSV color space and modifies only the saturation component while keeping hue and value constant. Working in HSV color space allows us to add hue, chroma and saturation constraints within the main framework, rather than including them in separate stages as it was the case

with our GEA-LAB2 (presented in previous chapter). Our new GEA is based on the kernel-based Retinex (KBR) method [Bertalmío et al., 2009]. By using KBR we can pose the gamut extension problem as one of increasing saturation.

One fundamental mechanism of the KBR algorithm [Bertalmío et al., 2009] is that it is capable of increasing contrast while being monotonically increasing, i.e. it can increase the contrast without decreasing the image values. [Bertalmío et al., 2009] propose the following formula to convert the intensity values I of an image into perceived lightness values L:

$$\begin{aligned} L(x) = &\sum_y w(x,y) f\left(\frac{I(x)}{I(y)}\right) \operatorname{sign}^+(I(y) - I(x)) \\ &+ \sum_y w(x,y) \operatorname{sign}^-(I(y) - I(x)), \end{aligned} \tag{7.1}$$

where $I(x)$ and $I(y)$ represent image values at pixel locations x and y, respectively, and I is an image channel. Since Eq. (7.1) involves ratios, we normalize the dynamic range of I in $(0, 1]$ so as to avoid dividing by zero. This, for example, could be done on a 8-bit-per-channel image just by adding 1 to each pixel and dividing by 256. The scaling function f is monotonically increasing such that $f(r) \geq r, \forall r$, $w(x, y)$ is a normalized Gaussian kernel of standard deviation of σ, and the functions $\operatorname{sign}^+(\cdot)$ and $\operatorname{sign}^-(\cdot)$ are, respectively, defined as:

$$\operatorname{sign}^+(a) = \begin{cases} 1, & \text{if } a > 0, \\ \frac{1}{2}, & \text{if } a = 0, \\ 0, & \text{if } a < 0, \end{cases} \tag{7.2}$$

$$\operatorname{sign}^-(a) = 1 - \operatorname{sign}^+(a). \tag{7.3}$$

In [Bertalmío et al., 2009] it is shown that the KBR is not idempotent. Therefore, instead of manually choosing the number of times the KBR should be applied to obtain the best enhanced image, it is better to rewrite Eq. (7.1) in the form of a partial differential equation (PDE): $I_t(x) = L(x, t) - I(x, t)$, into which we can introduce an attachment to data term that leaves the image unmodified once it has departed too much from the original image. We use this framework in the GM context by applying this PDE to the saturation channel only, while keeping the hue and value channels fixed: we know that KBR will increase the saturation

values, which implies that the gamut will be extended. In the PDE we also attach a chroma term $C = SV$ (from [Ford and Roberts, 1998]) to treat low chromatic and high chromatic pixels differently. Taking Eq. (7.1) for the saturation channel S and adding also a chroma attachment we get:

$$\begin{aligned} S_t(x,t) =& \gamma \sum_y w(x,y) \left[f\left(\frac{S(x,t)}{S(y,t)}\right) \text{sign}^+(S(y,t) \right. \\ & \left. -S(x,t)) + \text{sign}^-(S(y,t) - S(x,t)) \right] \\ & - S(x,t) - \beta(S(x,t) - S_0(x)) \\ & - \tau(S_0(x))(S(x,t)V_0(x) - S_0(x)V_0(x)), \end{aligned} \tag{7.4}$$

where $\beta > 0$ controls the strength of the original data attachment term S_0 and γ is a positive constant. V_0 is the value component of the original image. In gamut extension, certain colors require a special treatment in order to look natural and pleasant such as skin tones, less saturated natural objects, neutral colors and some particular memory colors. To incorporate such a functionality into the proposed GEA-KBR, the function $\tau(\cdot)$, shown in Fig. 7.1, takes as an argument the saturation of the original image and associates weights with the pixels of the image. These weights indicate the amount of special treatment we provide to colors of the input image. For example, the low saturated colors of the input image that require little to no extension are given higher weights $\tau(\cdot)$, whereas low weights are attached to extend normally the high saturated (artificial objects) colors. With this we can extend the gamut of the input image in a controlled manner treating objects of low saturation and high saturation differently. To compute the weights for each pixel, we use the following adapted version of the generalised logistic function [Wikipedia, 2016]:

$$\tau(S_0(x)) = \tau_{max}\left(1 - \frac{1}{1 + 0.55e^{-1.74S_0(x)^2}}\right) \tag{7.5}$$

where τ_{max} is a positive constant. The values used in Eq. (7.5) have been chosen based on tests we performed on several images with different color characteristics. (Note that the function $\tau(\cdot)$ effects only the chroma of the image and this is a much smoother way of implementing chroma constraint than defining a crucial region on chromaticity diagram and treating pixels differently as we did for our GEA-LAB2 presented in section 6.2).

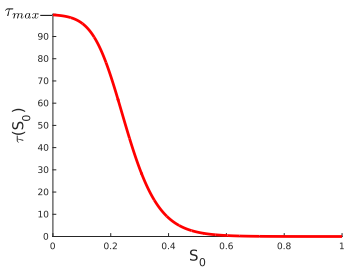


Figure 7.1: Logistic function.

Let us now discretize the derivative and apply a forward-time numerical scheme on Eq. (7.4) as

$$\begin{aligned} S^{k+1}(x) =& \gamma \Delta t R_{S^k}(x) + S^k(x)[1 - \Delta t(1+\beta)] \\ &+ \Delta t[\beta S_0(x) + \tau(S_0(x))V_0(x)(S_0(x) - S^k(x))], \end{aligned} \tag{7.6}$$

where Δt is the time step and $k \in \mathbb{N}$ denotes the iteration number. The initial condition is $S^{k=0}(x) = S_0(x)$, and the function $R_{S^k}(x)$ indicates the contrast modification function:

$$\begin{aligned} R_{S^k}(x) = \sum_{y \in \mathfrak{I}} w(x,y) \Bigg[f\left(\frac{S^k(x)}{S^k(y)}\right) \text{sign}^+(S^k(y) - S^k(x)) \\ + \text{sign}^-(S^k(y) - S^k(x)) \Bigg]. \end{aligned} \tag{7.7}$$

Following the work of [Bertalmío et al., 2009] we can rewrite Eq. (7.6) as

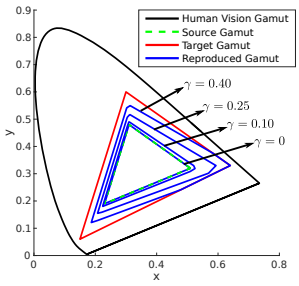
$$S^{k+1}(x) = \frac{S^k(x) + \Delta t \left(S_0(x)(\beta + \tau(S_0(x))V_0(x)^2) + \frac{\gamma}{2} R_{S^k}(x) \right)}{1 + \Delta t(\beta + \tau(S_0(x))V_0(x)^2)} \tag{7.8}$$

7.1.1 Action of the GEA-KBR on an Image

In order to perform gamut extension, we first convert the RGB input image to the HSV color space, and then apply Eq. (7.8) to the saturation component only. The

evolution Eq. (7.8) reaches steady state for each set of values for parameters β, Δt, and γ, and to show how it modifies the color gamut, an example with several different gamuts on a chromaticity diagram is shown in Fig. 7.2a where it can be seen that when $\beta = 1$, $\Delta t = 0.10$, and $\gamma = 0$ the steady state of the evolution equation produces the original image, and as we increase the value of γ the steady state of Eq. (7.8) yields an image with a larger gamut. Fig. 7.2a also shows that one can map the colors of an input image to the larger destination gamut just by using a large enough value for γ ($\gamma = 0.40$ in this case). For each γ value, the corresponding reproduced images are shown in Fig. 7.2b in which we can notice that the colors of sky, grass and skin have a limited (controlled) increase in saturation, whereas some objects are extended normally such as the bottles, the fruits, the cloths in the basket and the pink shirt. Since we are using a fixed value of γ for all the input images with different spatial and chromatic characteristics, our GEA-KBR may place a few colors outside the destination gamut that we can map back inside using a gamut reduction algorithm, e.g. our GRA-RGB [Zamir et al., 2014].

In order to compute the optimal value for γ that allows us to perform GE, we make use of the information provided by the color differences between the



(a)

(b)

Figure 7.2: Gamut extension approach. (a) Gamuts on chromaticity diagram. (b) Gamut extension results. Top left: input image ($\gamma = 0$). Top right: gamut extended image with $\gamma = 0.10$. Bottom left: gamut extended image with $\gamma = 0.25$. Bottom right: gamut extended image with $\gamma = 0.40$. Original image is from [CIE, 2004].

source and the destination gamuts (triangle-shaped due to the use of three color primaries):

$$\gamma = \sqrt[3]{|SG_{area} - DG_{area}|} \tag{7.9}$$

where SG_{area} and DG_{area} denote the area of the source gamut and the destination gamut, respectively.

7.1.2 Experiments and Results

In this section we assess the visual quality of the reproductions produced by the proposed GEA-KBR and other methods using the publicly available datasets [Andriani et al., 2013], [Froehlich et al., 2014]. To compute the results for our method, we work in HSV color space and in Eq. (7.8) the parameter values that we use are $\beta = 1$, $\Delta t = 0.10$. The non-linear scaling function is $f(r) = A\log(r)+1$, where $A = \frac{1}{\log(256)}$. The value for σ, the standard deviation for $w(x, y)$, is equal to one-third of the number of rows or columns of the input image (whichever is larger). Since we are limited by the sRGB standard (that has the same primaries as of BT.709) for the paper, all results presented in this section are mapped from 'Toy' gamut to sRGB gamut (the primaries of gamuts used in our experiments are mentioned in Table 7.1). To see the reproductions obtained using our GEA-KBR when mapping colors from 'Toy' to DCI-P3 gamut and from BT.709 to DCI-P3, visit the web page: `http://ip4ec.upf.edu/GamutExtension`.

While the capability of emerging wide-gamut displays in reproducing more vibrant colors helps generating a richer visual experience, it is highly important to render flesh tones in a very careful manner so as to avoid departing away

Table 7.1: Primaries of gamuts.

Gamuts	Red Primaries		Green Primaries		Blue Primaries	
	x	y	x	y	x	y
BT.2020	0.708	0.292	0.170	0.797	0.131	0.046
DCI-P3	0.680	0.320	0.265	0.690	0.150	0.060
BT.709/sRGB	0.640	0.330	0.300	0.600	0.150	0.060
Toy	0.570	0.320	0.300	0.530	0.190	0.130
Mock	0.510	0.320	0.310	0.480	0.230	0.190

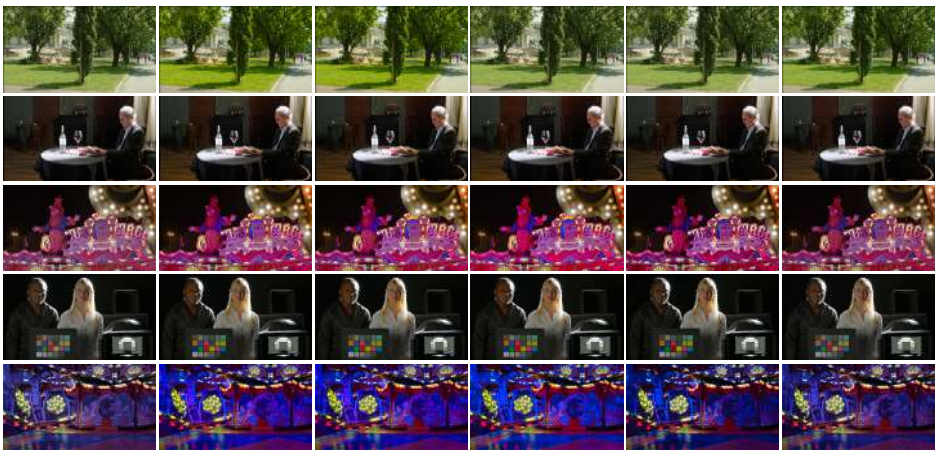


Figure 7.3: Results: mapping from Toy to sRGB gamut. Column 1: Input image. Column 2: HCM [Laird et al., 2009]. Column 3: SDS [Laird et al., 2009]. Column 4: Chroma extension [Laird et al., 2009]. Column 5: LCA [Laird et al., 2009]. Column 6: Our GEA-KBR. Original images are from [Andriani et al., 2013] and [Froehlich et al., 2014].

from the content creator's intent. In Fig. 7.3 we compare how different GEAs reproduce skin colors, which is always a key issue in movie postproduction. It can be seen in row 2 and row 4 that our GEA-KBR applies a limited extension to skin tones but the artificial objects undergo a normal color extension so that the final reproduction appears natural and pleasant, whereas in the same figure it is noticeable that other methods such as same-drive signal (SDS) [Laird et al., 2009] and chroma extension [Laird et al., 2009] have issues with skin tones and over-saturation. For better comparison, we show in Fig. 7.4 the zoomed-in view of regions cropped from Fig. 7.3. The lightness chroma adaptive (LCA) algorithm [Laird et al., 2009], due to its inherent behavior of modifying lightness and chroma, produces images with artifacts and the over-enhancement of contrast makes a few colors go towards black (loss of saturation and detail). For example in Fig. 7.4, see the napkin in row 2 for artifacts, and the picture of the elephant in the last row depicting some loss of detail. The chroma extension method [Laird et al., 2009] shows poor performance when there is a high chromatic object in the scene; one such example is shown in row 4 of Fig. 7.4 where it can be seen that the colors of the character are highly saturated.

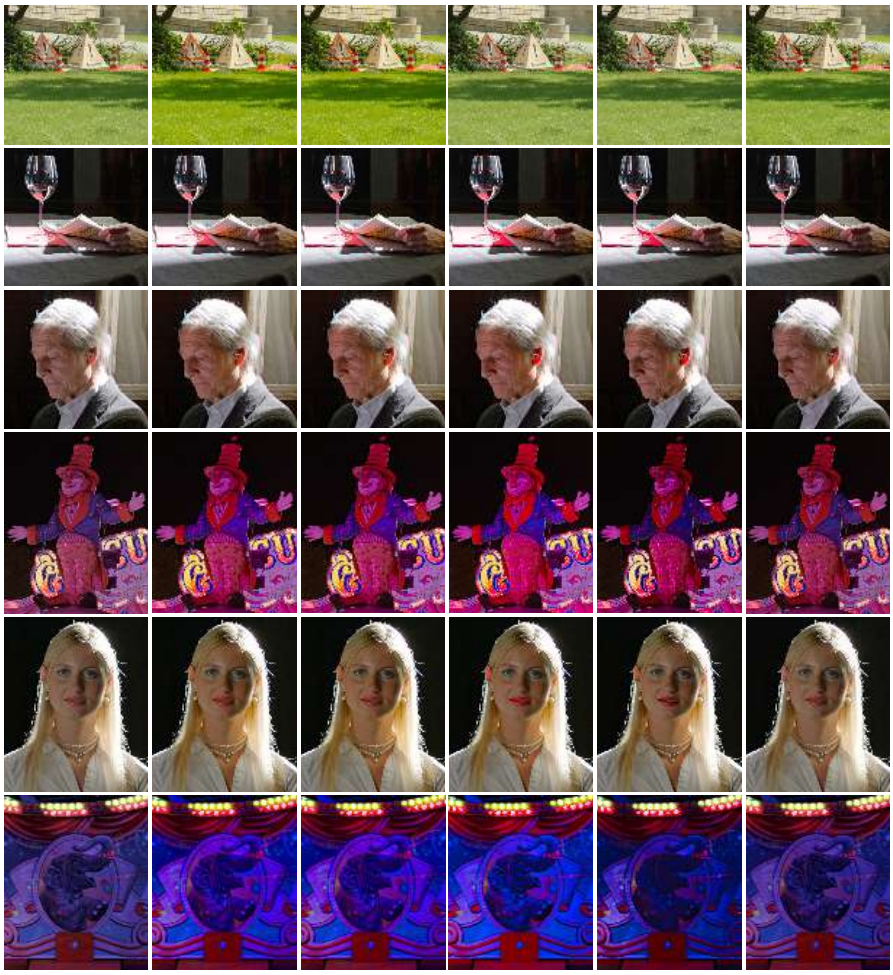


Figure 7.4: Zoomed-in view of the regions cropped from Fig. 7.3. Column 1: Input image. Column 2: HCM [Laird et al., 2009]. Column 3: SDS [Laird et al., 2009]. Column 4: Chroma extension [Laird et al., 2009]. Column 5: LCA [Laird et al., 2009]. Column 6: Our GEA-KBR.

In general, viewers tend to prefer more saturated colors [Fedorovskaya et al., 1997] when they are not aware of the original colors of either natural or man-made objects, but they are less welcoming to the change in colors of those objects for which they have memory such as shades of sky and grass, etc. Therefore, care

Figure 7.5: Representative frames of image sequences with sRGB gamut. Original images are from [Andriani et al., 2013] and [Froehlich et al., 2014].

should be taken while extending these colors. In Fig. 7.3, row 1, we present an example showing that our method performs controlled extension and reproduces the color of the grass accurately, whereas the hybrid color mapping (HCM), and SDS methods over-saturate the grass region, making it look artificial.

To test the temporal consistency of our method, we apply the GEA-KBR on a variety of image sequences with different levels of motion. Representative frames are shown in Fig. 7.5 and the complete resultant image sequences with extended gamut are added to the web page for which the link is given above. We confirm that the gamut extended videos obtained using our GEA are free from noticeable spatial and temporal artifacts, hue shifts, flickering and unexpected changes in chroma and brightness of objects.

7.1.3 Making the GEA-KBR Faster

The proposed GEA-KBR in a non-optimized MATLAB implementation running on a machine with 8 cores 3.4-GHz CPU takes (on average) 11 seconds to process an image of resolution 656×1080 pixels. However, the computational cost can be reduced drastically by applying a series of operations (only on the saturation component of an image represented in the HSV color space) that are shown in Fig. 7.6 and explained as follows:

- **Sub-sampling:** the first step is to take the full resolution image and sub-sample it by a scaling factor. The scaling factor of 0.40 provides a good trade-off between the speed and the quality of the reproduction.
- **Apply GEA:** next apply the proposed GEA-KBR on the downsampled

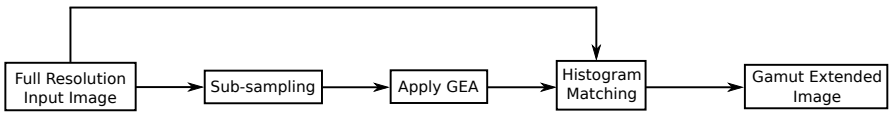


Figure 7.6: A schematic to reduce the computational cost of the proposed framework.

(a) (b) (c)

Figure 7.7: Example of reducing computational time. (a) Input image. (b) Result of GEA-KBR applied on a full resolution image. (c) Result of GEA-KBR applied on a sub-sampled image. Original image is from [Kodak, 1993].

image in order to obtain a reduced-size gamut extended image.

- **Histogram Matching:** finally we perform histogram matching of the full resolution input image and the subsampled image with extended gamut to obtain the final full resolution gamut extended image.

As shown in Fig. 7.7, using the above mentioned steps we can produce results having the same visual appearance but at a fraction (25%) of the time.

7.2 Gamut Reduction Using Kernel Based Retinex (GRA-KBR)

This section is devoted to present a key modification in the framework of GEA-KBR that will allow us to perform gamut reduction. Let us note that the chroma

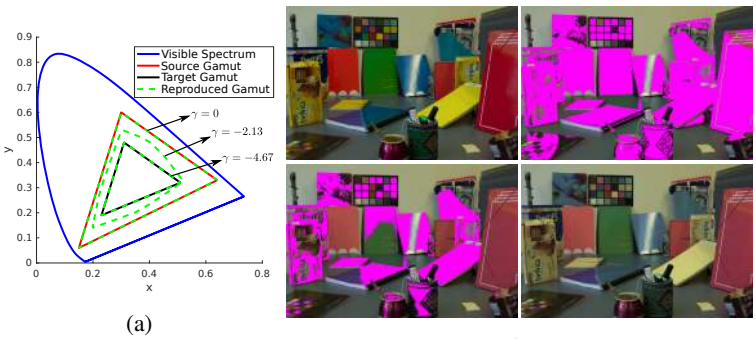
term in the PDE (Eq. 7.4) was exclusively used to deal with the gamut extension problem in which we need to treat low-saturated and high-saturated colors differently. So by removing the effect of the chroma term from Eq. (7.4), i.e. by setting $\tau(\cdot)$ to zero, we obtain the following corresponding evolution equation

$$S^{k+1}(x) = \frac{S^k(x) + \Delta t \left(\beta S_0(x) + \frac{\gamma}{2} R_{S^k}(x)\right)}{1 + \beta \Delta t}, \tag{7.10}$$

and by using $\gamma < 0$ in this equation we can decrease the saturation of the input image and subsequently reduce the color gamut.

7.2.1 Implementation

Let us note that it is a very common and recommended practice to leave those colors of the input image unchanged that are already inside the destination gamut and apply gamut reduction operation to the out-of-gamut colors only. To follow this approach, the proposed GRA-KBR works iteratively. The general structure of our iterative GRA-KBR, though similar to GRA-RGB (section 5.1.1), is as follows: at each iteration, we run Eq. (7.10) for some particular values of β and γ until we reach the steady state (in the first iteration the values are $\beta = 1$, and $\gamma = 0$). The steady state of each iteration will provide us with some pixels that we mark as "done" and we would not modify their values in subsequent iterations (i.e. these pixels are now part of the final output). We move to the next iteration where we make a small decrement in the γ value (for example, setting $\gamma = -0.05$) and run again Eq. (7.10) until steady state, and then check whether any of the pixels that were outside the gamut at the previous iteration are now inside the destination gamut: we select those pixels for the final image and leave them untouched for the following iterations. We keep repeating this process until all the out-of-gamut colors are mapped inside the destination gamut. An example of this iterative procedure is shown in Fig. 7.8b, where pixels in color magenta represent the out-of-gamut pixels remaining in that iteration. The corresponding evolution of the gamut is illustrated in Fig. 7.8a, showing that as γ decreases the gamut is gradually reduced.



(a) (b)

Figure 7.8: Gamut reduction procedure. (a) Gamuts on chromaticity diagram. (b) Results of gamut reduction. Top left: input image. Top right: reduced gamut image (same as of input image) when $\gamma = 0$. Bottom left: reduced-gamut image with $\gamma = -2.13$. Bottom right: reduced-gamut image with $\gamma = -4.67$. Out-of-gamut colors are masked with magenta color; as we reduced γ value the number of out-of-gamut colors is also decreased.

7.2.2 Experiments and Results

In this section we compare, visually and by using a perceptual error metric [Lissner et al., 2013], the performance of our GRA with other methods such as LCLIP [Sara, 1984], HPMINDE [Murch and Taylor, 1989] and the algorithms of [Schweiger et al., 2016] and [Alsam and Farup, 2012]. In order to map colors from a larger source gamut to a smaller destination gamut, we keep iterating the evolution equation (7.10) for each γ value until the difference between two consecutive steps falls below 0.5%.

Visual Quality Assessment

To evaluate the image reproduction quality, we map colors of sRGB images (Kodak dataset [Kodak, 1993] and images of Fig. 7.9) to a challenging smaller 'Mock' gamut using the competing GRAs. The color primaries of sRGB and 'Mock' gamuts are given in Table 7.1. The results presented in Fig. 7.10 show that the proposed GRA-KBR produces images that are perceptually more faithful to the original images than the other methods, preserving hues and retaining

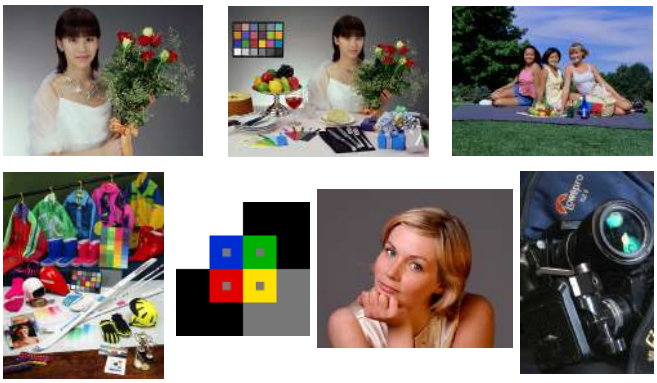


Figure 7.9: sRGB test images. From left to right, top to bottom: first 2 images are from [ISO, 2004], image 3 and 4 are from [CIE, 2004], image 5 is from [Farup et al., 2007] and the last two images are from the photographer Ole Jakob Bøe Skattum.

texture and color gradients. In the close-ups shown in Fig. 7.11 we can see that HPMINDE may introduce noticeable artifacts in the reproductions because it may project two nearby out-of-gamut colors to far-away points on the destination gamut. Also, the HPMINDE algorithm can produce images with loss of spatial detail, as it can be observed in rows 1 and 4 of Fig. 7.11. The methods of LCLIP and Schweiger et al. [Schweiger et al., 2016] may produce results with excessive desaturation in bright regions, as shown on the close-ups of the helmet and on the neck of the yellow parrot. The algorithm of Alsam et al. [Alsam and Farup, 2012] can over-compensate the contrast, as shown in the example of the first row of 7.11. In the example of the second row of Fig. 7.11, all tested GRAs except the proposed one produce tonal discontinuities on the face of the woman.

Quantitative Assessment

This section is devoted to examining the quality of the compared GRAs using the CID metric [Lissner et al., 2013], which is particularly tailored to assess the quality of gamut reduction results. We apply the competing GRAs on the Kodak dataset [Kodak, 1993] and on seven other images (that are shown in Fig. 7.9). In Table 7.2 we summarize the results obtained using the CID measure. It can be seen that our GRA outperforms the other methods in 24 out of 31 test images.

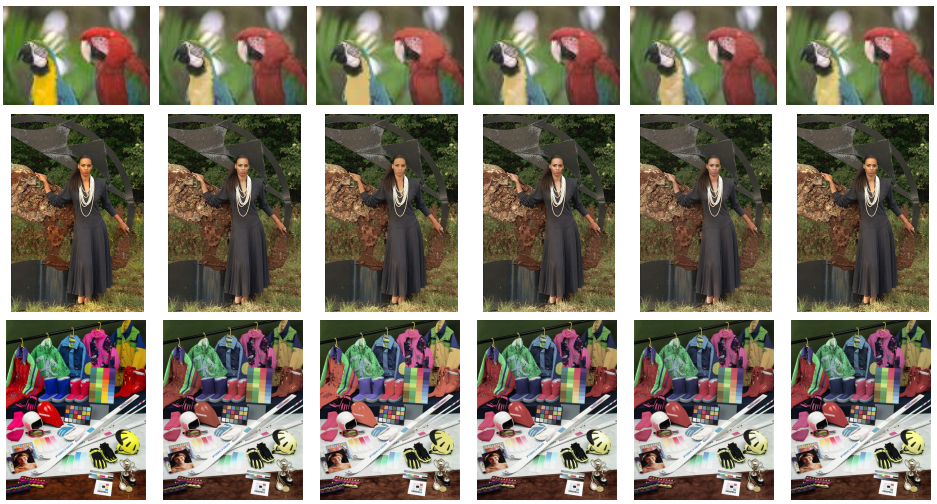


Figure 7.10: Reproductions of GRAs. Column 1: input images. Column 2: LCLIP [Sara, 1984]. Column 3: HPMINDE [Murch and Taylor, 1989]. Column 4: Schweiger et al. [Schweiger et al., 2016]. Column 5: Alsam et al. [Alsam and Farup, 2012]. Column 6: Our GRA. The original image in the last row is from [CIE, 2004], while rest of the input images are from Kodak dataset [Kodak, 1993].

Furthermore, the statistical data (mean, median and root mean square) presented in Table 7.3 also underlines the good performance of the proposed algorithm over the other GRAs.

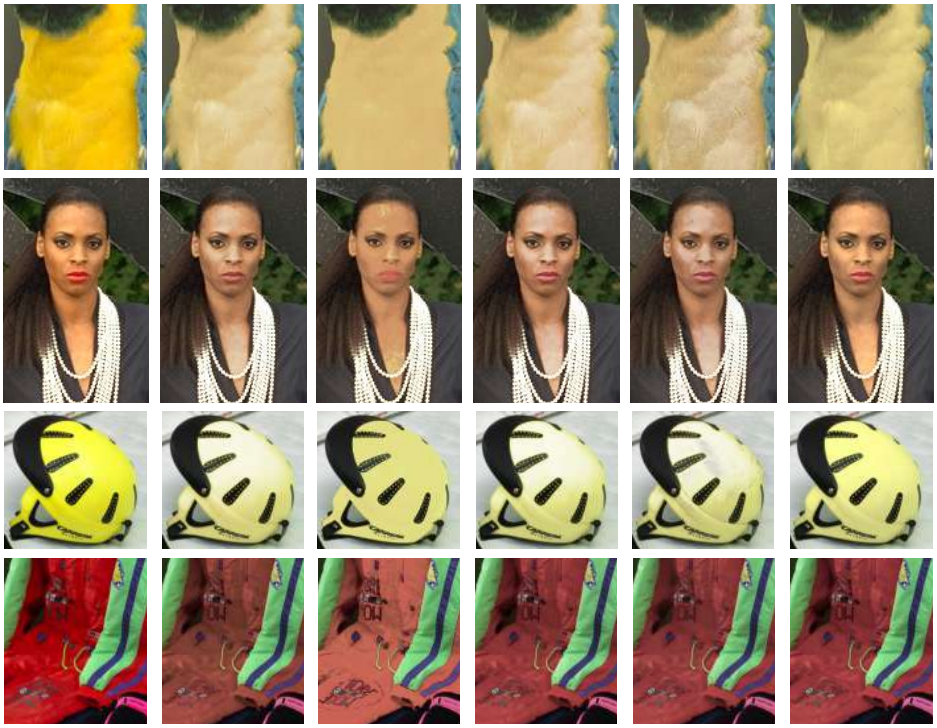


Figure 7.11: Comparison of GRAs: crops are from Fig. 7.10. Column 1: original cropped regions. Column 2: LCLIP [Sara, 1984]. Column 3: HPMINDE [Murch and Taylor, 1989]. Column 4: Schweiger et al. [Schweiger et al., 2016]. Column 5: Alsam et al. [Alsam and Farup, 2012]. Column 6: Our GRA.

Table 7.2: Quantitative assessment using CID metric [Lissner et al., 2013]. The first 24 images are from Kodak dataset [Kodak, 1993] and the rest of the images are sequentially shown from left to right in Fig. 7.9.

Image #	LCLIP [Sara, 1984]	HPMINDE	[Schweiger et al., 2016]	[Alsam and Farup, 2012]	Our GRA-KBR
1	0.0046	0.0242	0.0059	0.0255	**0.0051**
2	**0.1098**	0.3238	0.1195	0.1668	0.1135
3	0.0397	0.0831	0.0405	0.0631	**0.0291**
4	0.0114	0.0766	0.0131	0.0398	**0.0101**
5	0.0158	0.0252	0.0188	0.0330	**0.0061**
6	0.0275	0.0155	0.0324	0.0792	**0.0002**
7	0.0123	0.0369	0.0121	0.0504	**0.0109**
8	0.0067	0.0061	0.0157	0.0326	**0.0001**
9	**0.0037**	0.0097	0.0047	0.0525	0.0047
10	0.0010	0.0019	0.0012	0.0216	**0.0004**
11	0.0014	0.0055	0.0022	0.0277	**0.0010**
12	0.0077	0.0114	0.0082	0.0453	**0.0008**
13	0.0038	0.0060	0.0049	0.0231	**0.0012**
14	0.0238	0.0635	0.0252	0.0473	**0.0224**
15	0.0278	0.0452	0.0275	0.0716	**0.0157**
16	0.0014	0.0032	0.0018	0.0214	**0.0003**
17	0.0018	0.0012	0.0030	0.0116	**0.0003**
18	0.0082	0.0175	0.0134	0.0325	**0.0056**
19	0.0042	0.0096	0.0086	0.0428	**0.0013**
20	0.0478	0.0358	0.0517	0.0945	**0.0041**
21	0.0071	0.0037	0.0095	0.0293	**0.0001**
22	0.0080	0.0157	0.0121	0.0327	**0.0009**
23	0.0309	0.1480	0.0336	0.0668	**0.0308**
24	0.0028	0.0017	0.0057	0.0175	**0.0001**
25	0.0040	0.0175	**0.0038**	0.0284	0.0045
26	**0.0142**	0.0413	0.0145	0.0467	0.0156
27	**0.0087**	0.0443	0.0094	0.0321	0.0100
28	0.0752	0.1503	**0.0741**	0.1067	0.0832
29	0.1476	0.2081	0.1333	0.1324	**0.1148**
30	0.0299	0.0769	0.0372	0.0830	**0.0289**
31	**0.0083**	0.0130	0.0102	0.0230	0.0098

Table 7.3: Quantitative assessment: statistical data

	LCLIP [Sara, 1984]	HPMINDE	[Schweiger et al., 2016]	[Alsam and Farup, 2012]	Our GRA-KBR
Mean	0.0225	0.0491	0.0243	0.0510	**0.0171**
Median	0.0083	0.0175	0.0121	0.0398	**0.0051**
RMS	0.0396	0.0855	0.0397	0.0618	**0.0346**

CHAPTER 8

Conclusion and Future Work

The aim of this thesis was to develop gamut mapping algorithms that take into account some basic perceptual properties of the human visual system such that the content reproduced by these algorithms appears natural, pleasant and, most importantly, perceptually faithful to the original material. With this goal, we have presented several gamut reduction and gamut extension algorithms that are capable of mapping colors of an input image (or video) to a challenging destination gamut.

In chapter 5, we presented spatial gamut mapping algorithms (a GRA and a GEA) that rely on a perceptually-based variational framework. Our algorithms adapt an existing image energy functional [Bertalmío et al., 2007] whose minimization leads to image enhancement and contrast modification. We show how by varying the importance of the contrast term in the image functional we were able to perform gamut reduction and gamut extension. We proposed an iterative scheme that allows our algorithms to successfully map the colors from the gamut of the original image to a given destination gamut while keeping the perceived colors close to the original image. Both the visual quality assessment and the objective evaluation presented in chapter 5 validate the good results achieved using our proposed algorithms. Later in the same chapter, we presented another spatial GRA that is capable of better preserving the saliency of the original image in the gamut mapped image, but with a minor loss in overall image quality. All three of these GMAs work in RGB color space.

Our GEA of chapter 5, due to its inherent behavior to expand colors by increas-

ing the contrast of the image, may produce results with over-enhanced contrast, which in turn makes a few colors to go towards black (loss of saturation). To overcome this problem, we proposed a new spatial GEA (in chapter 6) that performs gamut extension by modifying only the chromatic components in the CIELAB color space while keeping the lightness constant, and therefore preserving the contrast of the image. Even though our GEA did provide a functionality through a threshold parameter to set the amount of color extension that we want to perform, the reproductions were sometimes off (over-saturated, or under-saturated) from the ground truth for a few images. Furthermore, we noticed that the proposed GEA introduced hue shifts in some images. To deal with these issues, we presented another GEA in chapter 6 that adapts itself according to the content of the input image. This novel GEA was implemented as a PDE-based optimization procedure related to visual perception models that performs gamut extension by taking into account the analysis of distortions in hue, chroma and saturation. User studies performed using a digital cinema projector under cinema-like environment showed that the proposed algorithm outperforms the state of the art, producing gamut extended images that are perceptually more faithful to the wide-gamut ground truth, as well as free from noticeable color artifacts and hue shifts. We also concluded how currently available image quality metrics, when applied to the gamut extension problem, provide results that do not correlate with users' choices.

Although the proposed GEA (chapter 6) showed promising and consistent performance under different settings of source and destination gamuts, there were some limitations associated with it. We presented an example showing that a poor choice of parameters for the constraints used in the GEA to make binary decisions may produce results with noticeable hue shifts, false edges, and undesirable chroma extension of memory colors. Additionally, the computational cost was significantly high, taking upward of seven minutes to process a full HD image using a non-optimized MATLAB implementation running on a desktop PC. To cope with these challenges, we presented a new local GEA in chapter 7 that works in the HSV color space and operates only on the saturation channel of the image, while keeping hue and value constant. This GEA extends gamut by analyzing the color characteristics of the input image. On one side, the proposed GEA modifies less those colors that require a special treatment such as skin tones, less saturated natural objects, neutral colors and some particular memory colors in a controlled manner. On the other side, it extends normally colors of high chromatic (natural

and artificial) objects. Our GEA produces gamut extended images and videos that look natural and pleasant in appearance. Though this method is able to extend the gamut of a full HD image in nine seconds, we also presented a procedure that enables our GEA to produce results having the same visual appearance but at a fraction (25%) of the time. In the second half of the chapter 7 we introduced a spatially-variant iterative GRA that comes from inverting the sign of a parameter of our GEA framework presented in the same chapter. This method also operates in the HSV color space and iteratively reduces the saturation component of the input image in order to perform gamut reduction. The results show that our GRA outperforms the state of the art both according to a quantitative metric and in terms of subjective visual appearance.

Future Work:

The following are some possible directions in gamut mapping that we would like to investigate as future works.

- Our user study (chapter 6) showed that the current image quality metrics, when applied to the gamut extension problem, provide results that do not correlate well with users' choices [Zamir et al., 2017]. Therefore, there is the need to develop an error metric specifically for gamut extension. Without a suitable image quality metric, the gamut extension problem cannot be posed as an optimization procedure; moreover, we are forced to conduct psychophysical tests for evaluating GEAs but these subjective tests are cumbersome, time-consuming and expensive.

- Although the spatial GMAs that we have proposed are adaptive and flexible, at the same time they are computationally expensive. These local GMAs need to be fast so that they can yield color reproductions on the fly in live broadcasts. The development of fast spatial GMAs is also important if we want to implement them in the image processing pipeline of a camera.

- The industry has been treating high dynamic range (HDR) and wide color gamut (WCG) as independent image attributes (so did we in this thesis), for example a picture may have a high dynamic range but a reduced color gamut such as BT.709, or the other way around [François et al., 2016]. The standardization of these technologies occurred in two separate routes [ITU-R, 2012], [SMPTE, 2014], mainly because the HDR displays were although

bright enough, they were not able to reproduce very wide color gamut. However, it is now foreseeable that the laser displays (which are capable of generating vivid colors and greater contrast) will become mainstream in near future. Therefore, we should look into studying HDR and WCG in combination.

- While performing gamut mapping, the proposed algorithms do not take into account the effect of ambient lighting. It would be interesting to develop GMAs that can generate content according to intended viewing conditions, for example reproducing a footage differently for home viewing (that has moderate light levels) and for cinema viewing (which has a dark ambiance).
- GMAs should preserve the artistic intent of the content's creator in the reproduced image if they are to be adopted by the movie industry. In this thesis we developed and psychophysically evaluated GMAs considering this accurate reproduction intent. However in the television industry, TV manufacturers prefer distorting image attributes such as tones and colors in ways that they think consumers may find visually pleasant [Poynton, 2010a]. This may imply departing from the original content as much as needed, and incorporating steps that may modify the aesthetic choices of the content's creator. As a future work we would like to develop and investigate GMAs that aim at generating content with pleasant reproduction intent.

Bibliography

[Adams et al., 1998] Adams, J., Parulski, K., and Spaulding, K. (1998). Color processing in digital cameras. *Micro, IEEE*, 18(6):20–30.

[Ajito et al., 2000] Ajito, T., Obi, T., Yamaguchi, M., and Ohyama, N. (2000). Expanded color gamut reproduced by six-primary projection display. In *Proc. SPIE Projection Displays*, pages 130–137.

[Alsam and Farup, 2012] Alsam, A. and Farup, I. (2012). Spatial colour gamut mapping by orthogonal projection of gradients onto constant hue lines. In *Proc. of 8th International Symposium on Visual Computing*, pages 556–565.

[Ancuti et al., 2011] Ancuti, C. O., Ancuti, C., and Bekaert, P. (2011). Enhancing by saliency-guided decolorization. In *Proc. of IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 257–264.

[Anderson et al., 2007] Anderson, H., Garcia, E., and Gupta, M. (2007). Gamut expansion for video and image sets. In *Proc. of the 14th International Conference of Image Analysis and Processing - Workshops*, pages 188–191.

[Andriani et al., 2013] Andriani, S., Brendel, H., Seybold, T., and Goldstone, J. (2013). Beyond the kodak image set: A new reference set of color image sequences. In *IEEE International Conference on Image Processing*, pages 2289–2293.

[Anikeeva et al., 2009] Anikeeva, P. O., Halpert, J. E., Bawendi, M. G., and Bulovic, V. (2009). Quantum dot light-emitting devices with electroluminescence tunable over the entire visible spectrum. *Nano letters*, 9(7):2532–2536.

[Bala et al., 2001] Bala, R., Dequeiroz, R., Eschbach, R., and Wu, W. (2001). Gamut mapping to preserve spatial luminance variations. *Journal of Imaging Science and Technology*, 45:122–128.

[Bankston, 2005] Bankston, D. (2005). The color-space conundrum, part one. *American Cinematographer*, page 6.

[Barańczuk et al., 2009] Barańczuk, Z., Zolliker, P., and Giesen, J. (2009). Image quality measures for evaluating gamut mapping. In *Color and Imaging Conference*, pages 21–26.

[Barco, 2007] Barco (2007). `http://www.barco.com/en/Products-Solutions/Projectors/Digital-cinema-projectors/2K-Digital-cinema-projector-for-screens-up-to-12m-40ft.aspx`.

[Beck, 2014] Beck, B. (2014). Lasers: Coming to a theater near you. `http://spectrum.ieee.org/consumer-electronics/audiovideo/lasers-coming-to-a-theater-near-you`.

[Belle, 2016] Belle, K. V. (2016). Laser-phosphor illumination in projectors. Barco Whitepaper.

[Berns, 1993] Berns, R. S. (1993). The mathematical development of CIE TC 1-29 proposed colour difference equation: CIELCH. In *Proc. of the Seventh Congress of International Colour Association, B, C19.1C19.4*.

[Berns, 2000] Berns, R. S. (2000). *Billmeyer and Saltzmans Principles of Color Technology*. John Wiley & Sons.

[Bertalmío, 2014] Bertalmío, M. (2014). *Image Processing for Cinema*, volume 4. CRC Press, Taylor & Francis.

[Bertalmío et al., 2009] Bertalmío, M., Caselles, V., and Provenzi, E. (2009). Issues about retinex theory and contrast enhancement. *International Journal of Computer Vision*, 83(1):101–119.

[Bertalmío et al., 2007] Bertalmío, M., Caselles, V., Provenzi, E., and Rizzi, A. (2007). Perceptual color correction through variational techniques. *IEEE Transactions on Image Processing*, 16(4):1058–1072.

[Bonnier et al., 2006] Bonnier, N., Schmitt, F., Brettel, H., and Berche, S. (2006). Evaluation of spatial gamut mapping algorithms. In *Proc. of IS&T/SID 14th Color Imaging Conference*, pages 56–61.

[Borer, 2014] Borer, T. (2014). Non-linear opto-electrical transfer functions for high dynamic range television. BBC Whitepaper WHP 283.

[Borer and Cotton, 2016] Borer, T. and Cotton, A. (2016). A display-independent high dynamic range television system. *SMPTE Motion Imaging Journal*, 125(4):50–56.

[Braun, 1999] Braun, G. J. (1999). *A Paradigm for Color Gamut Mapping of Pictorial Images*. PhD thesis, Rochester Institute of Technology, USA.

[Braun and Fairchild, 1999] Braun, G. J. and Fairchild, M. D. (1999). Image lightness rescaling using sigmoidal contrast enhancement functions. *Journal of Electronic Imaging*, 8(4):380–393.

[Brown, 2016] Brown, M. S. (2016). Understanding the in-camera image processing pipeline for computer vision. In *IEEE Computer Vision and Pattern Recognition - Tutorial*.

[Buades et al., 2005] Buades, A., Coll, B., and Morel, J.-M. (2005). A non-local algorithm for image denoising. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 60–65.

[Buades et al., 2009] Buades, A., Coll, B., Morel, J. M., and Sbert, C. (2009). Self-similarity driven color demosaicking. *IEEE Transactions on Image Processing*, 18(6):1192–1202.

[Buchsbaum, 1980] Buchsbaum, G. (1980). A spatial processor model for object colour perception. *Journal of the Franklin Institute*, 310(1):1–26.

[Cao et al., 2010] Cao, G., Pedersen, M., and Baranczuk, Z. (2010). Saliency models as gamut-mapping artifact detectors. In *5th European Conference on Colour in Graphics, Imaging, and Vision (CGIV)*, pages 437–443.

[Cardei and Funt, 1999] Cardei, V. C. and Funt, B. (1999). Committee-based color constancy. In *IS&T/SIDs Color Imaging Conference*, pages 311–313.

[Casella et al., 2008] Casella, S. E., Heckaman, R. L., and Fairchild, M. D. (2008). Mapping standard image content to wide-gamut displays. In *Color and Imaging Conference*, pages 106–111.

[CBS, 2013] CBS (2013). `http://www.cbs.com/shows/big_bang_theory/`.

[Chen et al., 2006] Chen, H., Sung, J., Ha, T., Park, Y., and Hong, C. (2006). Backlight local dimming algorithm for high contrast LCD-TV. In *Proc. of ASID*, volume 6, pages 168–171.

[Chen and Beghdadi, 2011] Chen, S. and Beghdadi, A. (2011). Improve gamut mapping of color management system by perceptual quality-oriented analysis. In *IEEE European Workshop on Visual Information Processing*, pages 12–17.

[Chen, 2002] Chen, X. (2002). Investigation of gamut extension algorithms. Master's thesis, University of Derby, UK.

[Cheng et al., 2010] Cheng, H.-C., Ben-David, I., and Wu, S.-T. (2010). Five-primary-color LCDs. *Journal of Display Technology*, 6(1):3–7.

[Chino et al., 2006] Chino, E., Tajiri, K., Kawakami, H., Ohira, H., Kamijo, K., Kaneko, H., Kato, S., Ozawa, Y., Kurumisawa, T., Inoue, K., Endo, K., Moriya, H., Aragaki, T., and Murai, K. (2006). Development of wide-color-gamut mobile displays with four-primary-color LCDs. *SID Symposium Digest of Technical Papers*, 37(1):1221–1224.

[CIE, 2001] CIE (2001). Improvement to industrial colour-difference evaluation. Technical report, CIE 142.

[CIE, 2004] CIE (2004). Guidelines for the evaluation of gamut mapping algorithms. Technical report, CIE 156.

[Cok, 1987] Cok, D. (1987). Signal processing method and apparatus for producing interpolated chrominance values in a sampled color image signal. US Patent 4,642,678.

[Daly et al., 2013a] Daly, S., Kunkel, T., Sun, X., Farrell, S., and Crum, P. (2013a). Preference limits of the visual dynamic range for ultra high quality and aesthetic conveyance. In *Proc. of IS&T/SPIE Electronic Imaging*, pages 1–11.

[Daly et al., 2013b] Daly, S., Kunkel, T., Sun, X., Farrell, S., and Crum, P. (2013b). Viewer preferences for shadow, diffuse, specular, and emissive luminance limits of high dynamic range displays. *SID Symposium Digest of Technical Papers*, 44(1):563–566.

[Debevec and Malik, 1997] Debevec, P. and Malik, J. (1997). Recovering high dynamic range radiance maps from photographs. In *Proc. of the 24th annual conf. on Computer graphics*, pages 369–378.

[Dudley et al., 2003] Dudley, D., Duncan, W. M., and Slaughter, J. (2003). Emerging digital micromirror device (DMD) applications. In *Micromachining and Microfabrication*, pages 14–25.

[Dugay et al., 2008] Dugay, F., Farup, I., and Hardeberg, J. Y. (2008). Perceptual evaluation of color gamut mapping algorithms. *Color Research & Application*, 33(6):470–476.

[Eskicioglu and Fisher, 1995] Eskicioglu, A. M. and Fisher, P. S. (1995). Image quality measures and their performance. *IEEE Transactions on Communications*, 43(12):2959–2965.

[Fairchild, 2004] Fairchild, M. (2004). Color appearance models:CIECAM02 and beyond. In *In tutorial notes, IS&T/SID 22nd Color Imaging Conference*.

[Fairchild, 2013] Fairchild, M. (2013). *Color appearance models*. John Wiley & sons.

[Fairman et al., 1997] Fairman, H. S., Brill, M. H., and Hemmendinger, H. (1997). How the CIE 1931 color-matching functions were derived from wright-guild data. *Color Research & Application*, 22(1):11–23.

[Farid, 2001] Farid, H. (2001). Blind inverse gamma correction. *IEEE Transactions on Image Processing*, 10(10):1428–1433.

[Farup et al., 2007] Farup, I., Gatta, C., and Rizzi, A. (2007). A multiscale framework for spatial gamut mapping. *IEEE Transactions on Image Processing*, 16(10):2423–2435.

[Fedorovskaya et al., 1997] Fedorovskaya, E. A., de Ridder, H., and Blommaert, F. J. (1997). Chroma variations and perceived quality of color images of natural scenes. *Color research & application*, 22(2):96–110.

[Ferwerda, 2001] Ferwerda, J. A. (2001). Elements of early vision for computer graphics. *IEEE Computer Graphics and Applications*, 21(5):22–33.

[Ford and Roberts, 1998] Ford, A. and Roberts, A. (1998). Colour space conversions. `http://www.poynton.com/PDFs/coloureq.pdf`.

[François et al., 2016] François, E., Fogg, C., He, Y., Li, X., Luthra, A., and Segall, C. (2016). High dynamic range and wide color gamut video coding in HEVC: Status and potential future enhancements. *IEEE Transactions on Circuits and Systems for Video Technology*, 26(1):63–75.

[Froehlich et al., 2014] Froehlich, J., Grandinetti, S., Eberhardt, B., Walter, S., Schilling, A., and Brendel, H. (2014). Creating cinematic wide gamut HDR-video for the evaluation of tone mapping operators and HDR-displays. In *Proc. of IS&T/SPIE Electronic Imaging*.

[Froehlich et al., 2013] Froehlich, J., Schilling, A., and Eberhardt, B. (2013). Gamut mapping for digital cinema. In *SMPTE Annual Technical Conference & Exhibition*, volume 2013, pages 1–11.

[Gatta and Farup, 2017] Gatta, C. and Farup, I. (2017). Gamut mapping in RGB colour spaces with the iterative ratios diffusion algorithm. In *Proc. of IS&T/SPIE Electronic Imaging*, pages 12–20.

[Gentile et al., 1990] Gentile, R. S., Walowitt, E., and Allebach, J. P. (1990). A comparison of techniques for color gamut mismatch compensation. *Journal of Imaging Technology*, 16:176–181.

[Gonzalez and Woods, 2006] Gonzalez, R. and Woods, R. (2006). *Digital Image Processing (3rd Edition)*. Prentice-Hall, Inc.

[Gooch et al., 2005] Gooch, A. A., Olsen, S. C., Tumblin, J., and Gooch, B. (2005). Color2gray: salience-preserving color removal. *ACM Trans. Graph.*, 24:634–639.

[Grossberg and Nayar, 2003] Grossberg, M. D. and Nayar, S. K. (2003). Determining the camera response from images: what is knowable? *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25(11):1455–1467.

[Grossberg and Nayar, 2004] Grossberg, M. D. and Nayar, S. K. (2004). Modeling the space of camera response functions. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 26(10):1272–1282.

[Guild, 1932] Guild, J. (1932). The colorimetric properties of the spectrum. *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 230(681-693):149–187.

[Hamilton and Adams, 1997] Hamilton, J. J. and Adams, J. J. (1997). Adaptive color plan interpolation in single sensor color electronic camera. US Patent 5,629,734.

[Hardeberg et al., 2008] Hardeberg, J. Y., Bando, E., and Pedersen, M. (2008). Evaluating colour image difference metrics for gamut-mapped images. *Coloration Technology*, 124(4):243–253.

[Hartline, 1940] Hartline, H. K. (1940). The receptive fields of optic nerve fibers. *American Journal of Physiology – Legacy Content*, 130(4):690–699.

[Heckaman and Sullivan, 2011] Heckaman, R. L. and Sullivan, J. (2011). Rendering digital cinema and broadcast TV content to wide gamut display media. *SID Symposium Digest of Technical Papers*, 42(1):225–228.

[Herzog and Müller, 1997] Herzog, P. G. and Müller, M. (1997). Gamut mapping using an analytical color gamut representation. In *Proc. of Color Imaging: Device-Independent Color, Color Hard Copy, and Graphic Arts*, pages 117–128.

[Hibbard, 1995] Hibbard, R. (1995). Apparatus and method for adaptively interpolating a full color image utilizing luminance gradients. US Patent 5,382,976.

[Hoffman et al., 2014] Hoffman, D. M., Johnson, P. V., Kim, J. S., Vargas, A. D., and Banks, M. S. (2014). 240Hz OLED technology properties that can enable improved image quality. *Journal of the Society for Information Display*, 22(7):346–356.

[Hornbeck, 1997] Hornbeck, L. J. (1997). Digital light processing for high-brightness high-resolution applications. In *Proc. SPIE Electronic Imaging*, pages 27–40.

[Hoshino, 1991] Hoshino, T. (1991). A preferred color reproduction method for the HDTV digital still image system. In *Proc. of IS&T Symposium on Electronic Photography*, pages 27–32.

[Hoshino, 1994] Hoshino, T. (1994). Color estimation method for expanding a color image for reproduction in a different color gamut. US Patent 5,317,426.

[Hubel, 1995] Hubel, D. (1995). *Eye, Brain, and Vision*. Scientific American Library Series. Henry Holt and Company.

[Hubel et al., 1997] Hubel, P., Holm, J., Finlayson, G., Drew, M., et al. (1997). Matrix calculations for digital photography. In *The Fifth Color Imaging Conference: Color, Science, Systems and Applications*.

[Hurvich and Jameson, 1957] Hurvich, L. and Jameson, D. (1957). An opponent-process theory of color vision. *Psychol Rev.*, 16(3):384–404.

[ISO, 2004] ISO (2004). 12640-2: Graphic technology – Prepress digital data exchange – Part 2: XYZ/sRGB encoded standard colour image data (XYZ/SCID).

[ITU-R, 2002] ITU-R (2002). Recommendation BT.709-5: parameter values for the HDTV standards for production and international programme exchange.

[ITU-R, 2011] ITU-R (2011). Recommendation BT.1886: reference electro-optical transfer function for flat panel displays used in hdtv studio production.

[ITU-R, 2012] ITU-R (2012). Recommendation BT.2020: parameter values for ultra high definition television systems for production and international programme exchange.

[ITU-R, 2013] ITU-R (2013). Recommendation BT.2035: a reference viewing environment for evaluation of HDTV program material or completed programmes.

[ITU-R, 2016] ITU-R (2016). Recommendation BT.2100: image parameter values for high dynamic range television for use in production and international programme exchange.

[Jacobson and Nguyen, 2011] Jacobson, N. and Nguyen, T. Q. (2011). Video processing with scale-aware saliency: Application to frame rate up-conversion.

In *IEEE International Conference on Accoustic, Speech and Signal Processing*, pages 1313–1316.

[Jian et al., 2014] Jian, C., Hardev, V., and Yurek, J. (2014). Quantum dot displays: Giving LCDs a competitive edge through color. *Nanotechnology Law & Business*, 11:4–13.

[Johnson, 1979] Johnson, A. J. (1979). Perceptual requirements of digital picture processing. *Paper presented at IARAIGAI symposium and printed in part in Printing World*.

[Judd, 1979] Judd, D. (1979). *Contributions to color science*, volume 545, chapter Estimation of chromaticity differences and nearest color temperature on the standard 1931 ICI colorimetric coordinate system, pages 207–212. NBS.

[Kalloniatis and Luu, 2007] Kalloniatis, M. and Luu, C. (2007). Light and dark adaptation. `http://webvision.med.utah.edu/`.

[Kang et al., 2003] Kang, B. H., Morovič, J., Luo, M. R., and Cho, M. S. (2003). Gamut compression and extension algorithms based on observer experimental data. *ETRI journal*, 25(3):156–170.

[Katoh and Ito, 1996] Katoh, N. and Ito, M. (1996). Gamut mapping for computer generated images (ii). In *Proc. of 4th IS&T/SID Color Imaging Conference*, pages 126–129.

[Kefalov et al., 2003] Kefalov, V., Fu, Y., Marsh-Armstrong, N., and Yau, K.-W. (2003). Role of visual pigment properties in rod and cone phototransduction. *Nature*, 425(6957):526–531.

[Kennel, 2007] Kennel, G. (2007). *Color and mastering for digital cinema: digital cinema industry handbook series*. Taylor & Francis US.

[Kim et al., 2004] Kim, M. C., Shin, Y. C., Song, Y. R., Lee, S. J., and Kim, I. D. (2004). Wide gamut multi-primary display for HDTV. In *Proc. of 2nd European Conference on color Graphics, Imaging and Vision*, pages 248–253.

[Kim and Pollefeys, 2008] Kim, S. J. and Pollefeys, M. (2008). Robust radiometric calibration and vignetting correction. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 30(4):562–576.

[Kimmel et al., 2005] Kimmel, R., Shaked, D., Elad, M., and Sobel, I. (2005). Space-dependent color gamut mapping: A variational approach. *IEEE Transactions on Image Processing*, 14:796–803.

[Kodak, 1993] Kodak (1993). `http://r0k.us/graphics/kodak/`.

[König and Dieterich, 1892] König, A. and Dieterich, C. (1892). *Die Grundempfindungen in normalen und anomalen Farbensystemen und ihre Intensitätsverteilung im Spektrum*. L. Voss.

[Kuffler, 1953] Kuffler, S. W. (1953). Discharge patterns and functional organization of mammalian retina. *Journal of Neurophysiology*, 16(1):37–68.

[Kunkel and Reinhard, 2010] Kunkel, T. and Reinhard, E. (2010). A reassessment of the simultaneous dynamic range of the human visual system. In *Proceedings of the 7th Symposium on Applied Perception in Graphics and Visualization*, pages 17–24. ACM.

[Kupaev, 2012] Kupaev, T. (2012). `http://www.youtube.com/watch?v=ezyrSKgcyJw`.

[Kusakabe et al., 2013] Kusakabe, Y., Iwasaki, Y., and Nishida, Y. (2013). Wide-color-gamut super hi–vision projector. In *Proc. ITE annual convention (in Japanese)*.

[Laird et al., 2009] Laird, J., Muijs, R., and Kuang, J. (2009). Development and evaluation of gamut extension algorithms. *Color Research & Application*, 34(6):443–451.

[Lam and Fung, 2008] Lam, E. and Fung, G. (2008). *Single-sensor imaging: Methods and applications for digital cameras, chapter Automatic White Balancing in Digital Photography*. CRC press.

[Land, 1977] Land, E. H. (1977). The Retinex theory of color vision. *Scientific American*, 237:108–128.

[Land and McCann, 1971] Land, E. H. and McCann, J. J. (1971). Lightness and retinex theory. *Journal of the Optical Society of America*, pages 1–11.

[Lau et al., 2011] Lau, C., Heidrich, W., and Mantiuk, R. (2011). Cluster-based color space optimizations. In *Proc. of IEEE International Conference on Computer Vision*, ICCV '11, pages 1172–1179.

[Li et al., 2011] Li, Y., Song, G., and Li, H. (2011). A multilevel gamut extension method for wide gamut displays. In *Proc. of International Conference on Electric Information and Control Engineering (ICEICE),*, pages 1035–1038.

[Lindner et al., 2011] Lindner, A., Bonnier, N., and Susstrunk, S. (2011). Saliency driven black point compensation. In *Proc. of IS&T/SPIE Electronic Imaging*.

[Ling, 2001] Ling, Y. (2001). Investigation of a gamut extension algorithm. Master's thesis, University of Derby, UK.

[LIPA, 2014] LIPA (2014). `http://lipainfo.org/laser-projection/laser-projectors/`.

[Lissner et al., 2013] Lissner, I., Preiss, J., Urban, P., Lichtenauer, M. S., and Zolliker, P. (2013). Image-difference prediction: From grayscale to color. *IEEE Transactions on Image Processing*, 22(2):435–446.

[Liu et al., 2010] Liu, Y., Song, G., and Li, H. (2010). A hue-preserving gamut expansion algorithm in CIELUV color space for wide gamut displays. In *Proc. of the 3rd International Congress on Image and Signal Processing (CISP)*, pages 2401–2404.

[Lueder, 2010] Lueder, E. (2010). *Liquid Crystal Displays: Addressing Schemes and Electro-Optical Effects*. John Wiley & Sons, Ltd.

[Luo et al., 2001] Luo, M. R., Cui, G., and Rigg, B. (2001). The development of the CIE 2000 colour-difference formula: CIEDE2000. *Color Research & Application*, 26(5):340–350.

[Luo et al., 2014] Luo, Z., Xu, D., and Wu, S. T. (2014). Emerging quantum-dots-enhanced LCDs. *Journal of Display Technology*, 10(7):526–539.

[MacEvoy, 2015] MacEvoy, B. (2015). `http://www.handprint.com/HP/WCL/color2.html`.

[Mandle, 2010] Mandle, G. (2010). Oled; what is it and how does it work? In *SMPTE Annual Technical Conference & Exhibition*, pages 1–17.

[Mantiuk, 2007] Mantiuk, R. (2007). *High-fidelity imaging: the computational models of the human visual system in high dynamic range video compression, visible difference prediction and image processing*. PhD thesis, Saarland University, Germany.

[Marcu and Abe, 1996] Marcu, G. and Abe, S. (1996). Gamut mapping for color simulation on CRT devices. In *Proc. of Color Imaging: Device-Independent Color, Color Hard Copy, and Graphic Arts*.

[Masaoka et al., 2016] Masaoka, K., Kusakabe, Y., Yamashita, T., Nishida, Y., Ikeda, T., and Sugawara, M. (2016). Algorithm design for gamut mapping from UHDTV to HDTV. *Journal of Display Technology*, 12(7):760–769.

[McCann, 1999] McCann, J. J. (1999). Lessons learned from mondrians applied to real images and color gamuts. In *Proc. of Color Imaging Conference*, pages 1–8.

[McCann, 2002] McCann, J. J. (2002). A spatial colour gamut calculation to optimize colour appearance. In *Colour Image Science: Exploiting Digital Media*, pages 213–233.

[McKnight and Curtis, 2014] McKnight, D. and Curtis, K. (2014). System and method for vibrating screens to reduce speckle. US Patent App. 14/298,633.

[Meng et al., 2013] Meng, X., Song, G., and Li, H. (2013). A human skin-color-preserving extension algorithm for wide gamut displays. In *Proc. of International Conference on Information Technology and Software Engineering*, Lecture Notes in Electrical Engineering, pages 705–713.

[Meyer and Barth, 1989] Meyer, J. and Barth, B. (1989). Color gamut matching for hard copy. In *Proc. of SID Digest*, pages 86–89.

[Michelson, 1927] Michelson, A. A. (1927). *Studies in Optics*. University of Chicago Press.

[Miller et al., 2013] Miller, S., Nezamabadi, M., and Daly, S. (2013). Perceptual signal coding for more efficient usage of bit codes. *SMPTE Motion Imaging Journal*, 122(4):52–59.

[Mollon and Cavonius, 1987] Mollon, J. D. and Cavonius, C. R. (1987). *The Chromatic Antagonisms of Opponent Process Theory are not the Same as Those Revealed in Studies of Detection and Discrimination*, pages 473–483. Springer Netherlands, Dordrecht.

[Monnier, 2008] Monnier, P. (2008). Standard definitions of chromatic induction fail to describe induction with S-cone patterned backgrounds. *Vision Research*, 48(27):2708 – 2714.

[Montag and Fairchild, 1997] Montag, E. D. and Fairchild, M. D. (1997). Psychophysical evaluation of gamut mapping techniques using simple rendered images and artificial gamut boundaries. *IEEE Transactions on Image Processing*, 6(7):977–989.

[Montag and Fairchild, 1998] Montag, E. D. and Fairchild, M. D. (1998). Gamut mapping: Evaluation of chroma clipping techniques for three destination gamuts. In *Color and Imaging Conference*, pages 57–61.

[Morovič, 1998] Morovič, J. (1998). *To Develop a Universal Gamut Mapping Algorithm*. PhD thesis, University of Derby, UK.

[Morovič, 2008] Morovič, J. (2008). *Color gamut mapping*, volume 10. Wiley.

[Morovič and Wang, 2003] Morovič, J. and Wang, Y. (2003). A multi-resolution, full-colour spatial gamut mapping algorithm. In *Proc. of Color Imaging Conference*, pages 282–287.

[Muijs et al., 2006] Muijs, R., Laird, J., Kuang, J., and Swinkels, S. (2006). Subjective evaluation of gamut extension methods for wide-gamut displays. In *Proc. of the 13th International Display Workshop*, pages 1429–1432.

[Murch and Taylor, 1989] Murch, G. M. and Taylor, J. M. (1989). Color in computer graphics: Manipulating and matching color. *Eurographics Seminar: Advances in Computer Graphics V*, pages 41–47.

[Murray et al., 2013] Murray, N., Vanrell, M., Otazu, X., and Prraga, C. A. (2013). Low-level spatiochromatic grouping for saliency estimation. *IEEE Trans. Pattern Anal. Mach. Intell.*, 35(11):2810–2816.

[Myers, 2003] Myers, R. (2003). *Display interfaces: fundamentals and standards*. John Wiley & Sons.

[Nakamura, 2005] Nakamura, J. (2005). *Image sensors and signal processing for digital still cameras [Chapter 3]*. CRC.

[Nakauchi et al., 1999] Nakauchi, S., Hatanaka, S., and Usui, S. (1999). Color gamut mapping based on a perceptual image difference measure. *Color Research & Application*, 24(4):280–291.

[Nayar and Mitsunaga, 2000] Nayar, S. K. and Mitsunaga, T. (2000). High dynamic range imaging: spatially varying pixel exposures. In *Proc. of IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 472–479.

[Normann and Perlman, 1979] Normann, R. and Perlman, I. (1979). The effects of background illumination on the photoresponses of red and green cones. *The Journal of physiology*, 286:491.

[Otsu, 1979] Otsu, N. (1979). A threshold selection method from gray-level histograms. *IEEE Transactions on Systems, Man and Cybernetics*, pages 62–66.

[Packer and Williams, 2003] Packer, O. and Williams, D. R. (2003). Light, the retinal image, and photoreceptors. *The Science of Color, 2nd ed SK. Shevall, ed. Optical Society of America, Elsevier (Oxford, UK)*, pages 41–102.

[Palma-Amestoy et al., 2009] Palma-Amestoy, R., Provenzi, E., Bertalmío, M., and Caselles, V. (2009). A perceptually inspired variational framework for color enhancement. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 31(3):458–474.

[Pan and Daly, 2008] Pan, H. and Daly, S. (2008). A gamut-mapping algorithm with separate skin and non-skin color preference controls for wide-color-gamut TV. *SID Symposium Digest of Technical Papers*, 39(1):1363–1366.

[Pedersen et al., 2013] Pedersen, M., Liu, X., and Farup, I. (2013). Improved simulation of image detail visibility using the non-subsampled contourlet transform. In *Proc. of IS&T/SID Color and Imaging Conference (CIC)*, pages 191–196.

[Pictures, 2013] Pictures, P. (2013). `http://www.painandgainmovie.com`.

[Pointer, 1980] Pointer, M. R. (1980). The gamut of real surface colours. *Color Research & Application*, 5(3):145–155.

[Poynton, 2007] Poynton, C. (2007). Wide-gamut displays. *Information Display*, 23(7):10–15.

[Poynton, 2010a] Poynton, C. (2010a). Contrast, brightness, and the naming of things. Poynton's Vector 1.

[Poynton, 2010b] Poynton, C. (2010b). Wide-gamut image capture. In *IS&T CGIV, Fourth European Conf. on Colour in Graphics and Imaging*, pages 471–482.

[Poynton, 2012] Poynton, C. (2012). *Digital Video and HDTV Algorithms and Interfaces*. Morgan Kaufmann Publishers Inc., 2 edition.

[Preiss et al., 2014] Preiss, J., Fernandes, F., and Urban, P. (2014). Color-image quality assessment: from prediction to optimization. *IEEE Transactions on Image Processing*, 23(3):1366–1378.

[Preiss and Urban, 2012] Preiss, J. and Urban, P. (2012). Image-difference measure optimized gamut mapping. In *Proc. of IS&T/SID 20th Color Imaging Conference*, pages 230–235.

[Pridmore, 1999] Pridmore, R. W. (1999). Bezold-brucke hue-shift as functions of luminance level, luminance ratio, interstimulus interval and adapting white for aperture and object colors. *Vision Research*, 39(23):3873 – 3891.

[Provenzi et al., 2005] Provenzi, E., De Carli, L., Rizzi, A., and Marini, D. (2005). Mathematical definition and analysis of the Retinex algorithm. *Journal of the Optical Society of America A*, 22(12):2613–2621.

[Purdy, 1931] Purdy, D. M. (1931). Spectral hue as a function of intensity. *The American Journal of Psychology*, 43(4):541–559.

[Purves and Lotto, 2003] Purves, D. and Lotto, R. (2003). *Why we see what we do: An empirical theory of vision.* Sinauer Associates.

[Pytlarz et al., 2016] Pytlarz, J., Thurston, K., Brooks, D., Boon, P., and Atkins, R. (2016). Real time cross-mapping of high dynamic range images. *IET Conference Proceedings*.

[Raja and Pedersen, 2013] Raja, K. B. and Pedersen, M. (2013). Artifact detection in gamut mapped images using saliency. In *IEEE Colour and Visual Computing Symposium (CVCS)*.

[Ramanath et al., 2005] Ramanath, R., Snyder, W., Yoo, Y., and Drew, M. (2005). Color image processing pipeline. *Signal Processing Magazine, IEEE*, 22(1):34–43.

[Ramanath et al., 2002] Ramanath, R., Snyder, W. E., Bilbro, G. L., and Sander, W. A. (2002). Demosaicking methods for bayer color arrays. *Journal of Electronic imaging*, 11(3):306–315.

[Rizzi et al., 2003] Rizzi, A., Gatta, C., and Marini, D. (2003). A new algorithm for unsupervised global and local color correction. *Pattern Recognition Letters*, 24:1663–1677.

[Roth et al., 2003] Roth, S., Ben-David, I., Ben-Chorin, M., Eliav, D., and Ben-David, O. (2003). Wide gamut, high brightness multiple primaries single panel projection displays. *SID Symposium Digest of Technical Papers*, 34(1):118–121.

[Samsung, 2016] Samsung (2016). `https://news.samsung.com/global/why-are-quantum-dot-displays-so-good`.

[Sapiro and Caselles, 1997] Sapiro, G. and Caselles, V. (1997). Histogram modification via partial differential equations. *Journal of Differential Equations*, 135(2):238–266.

[Sara, 1984] Sara, J. J. (1984). *The automated reproduction of pictures with nonreproducible colors*. PhD thesis, Massachusetts Institute of Technology (MIT).

[Schweiger et al., 2016] Schweiger, F., Borer, T., and Pindoria, M. (2016). Luminance-preserving colour conversion. In *SMPTE Annual Technical Conference and Exhibition*, pages 1–9.

[Seo and Milanfar, 2009] Seo, H. J. and Milanfar, P. (2009). Static and space-time visual saliency detection by self-resemblance. *Journal of Vision*.

[Sharma, 2002] Sharma, G. (2002). *Digital Color Imaging Handbook*. CRC Press, Inc., Boca Raton, FL, USA.

[Sharma et al., 2005] Sharma, G., Wu, W., and Dalal, E. N. (2005). The CIEDE2000 color-difference formula: Implementation notes, supplementary test data, and mathematical observations. *Color Research & Application*, 30(1):21–30.

[Silverstein et al., 2011] Silverstein, B. D., Kurtz, A. F., Bietry, J. R., and Nothhard, G. E. (2011). A laser-based digital cinema projector. *SID Symposium Digest of Technical Papers*, 42(1):326–329.

[SMPTE, 2011] SMPTE (2011). D-cinema quality – reference projector and environment (report 431-2:2011).

[SMPTE, 2014] SMPTE (2014). ST2084: high dynamic range electro-optical transfer function of mastering reference displays.

[Soneira, 2004] Soneira, R. M. (2004). Display technology shoot-out: Comparing CRT, LCD, plasma and DLP displays. `http://www.displaymate.com/ShootOut_Part_1.htm`.

[Song et al., 2014a] Song, G., Cao, H., and Huang, H. (2014a). Hue preserving multi-level expansion method based on saturation for wide gamut displays. *Journal of Information & Computational Science*, 11(2):461–472.

[Song et al., 2014b] Song, G., Meng, X., Li, H., and Han, Y. (2014b). Skin color region protect algorithm for color gamut extension. *Journal of Information & Computational Science*, 11(6):1909–1916.

[Stojmenovik, 2014] Stojmenovik, G. (2014). RGB laser projection for premium cinema screens. NAB Show.

[Stokes et al., 1996] Stokes, M., Anderson, M., Chandrasekar, S., and Motta, R. (1996). A standard default color space for the internet-srgb. *Available online:* *`https://www.w3.org/Graphics/Color/sRGB.html`*.

[Sugiura et al., 2006a] Sugiura, H., Kaneko, H., Kagawa, S., Ozawa, M., Someya, J., Tanizoe, H., Ueno, H., and Kimura, T. (2006a). Improved six-primary-color 23-in. WXGA LCD using six-color LEDs. *SID Symposium Digest of Technical Papers*, 37(1):1126–1129.

[Sugiura et al., 2006b] Sugiura, H., Kaneko, H., Kagawa, S., Someya, J., and Tanizoe, H. (2006b). Six-primary-color LCD monitor using six-color LEDs with an accurate calibration system. In *Proc. SPIE Color Imaging XI: Processing, Hardcopy, and Applications*.

[Susstrunk et al., 2000] Susstrunk, S. E., Holm, J. M., and Finlayson, G. D. (2000). Chromatic adaptation performance of different RGB sensors. In *Proc. SPIE Electronic Imaging*, pages 172–183.

[Thurstone, 1927] Thurstone, L. L. (1927). A law of comparative judgment. *Psychological Review*, 34(4):273–286.

[Torgerson, 1954] Torgerson, W. (1954). A law of categorical judgment, consumer behaviour. *Consumer Behaviour*, pages 92–93.

[Trisnadi, 2002] Trisnadi, J. I. (2002). Speckle contrast reduction in laser projection displays. In *Proc. SPIE Electronic Imaging*, pages 131–137.

[Ueki et al., 2009] Ueki, S., Nakamura, K., Yoshida, Y., Mori, T., Tomizawa, K., Narutaki, Y., Itoh, Y., and Okamoto, K. (2009). Five-primary-color 60-inch LCD with novel wide color gamut and wide viewing angle. *SID Symposium Digest of Technical Papers*, 40(1):927–930.

[UGRA, 1995] UGRA (17 July 1995). UGRA GAMCOM version 1.1: Program for the color gamut compression and for the comparison of calculated and measured values. Technical report, UGRA, St. Gallen.

[Van Hurkman, 2013] Van Hurkman, A. (2013). *Color correction handbook: professional techniques for video and cinema*. Pearson Education.

[Vazquez-Corral and Bertalmío, 2015] Vazquez-Corral, J. and Bertalmío, M. (2015). Simultaneous blind gamma estimation. *IEEE Signal Processing Letters*, 22(9):1316–1320.

[Vazquez-Corral et al., 2012] Vazquez-Corral, J., Vanrell, M., Baldrich, R., and Tous, F. (2012). Color constancy by category correlation. *IEEE Transactions on image processing*, 21(4):1997–2007.

[Vazquez-Corral et al., 2014] Vazquez-Corral, J., Zamir, S. W., and Bertalmío, M. (2014). Considering saliency in a perception inspired gamut reduction algorithm. In *Proc. of IS&T/SID 22nd Color Imaging Conference*.

[Wandell, 1995] Wandell, B. (1995). *Foundations of Vision*. Sinauer Associates.

[Wandell and Silverstein, 2003] Wandell, B. A. and Silverstein, L. D. (2003). 8 - digital color reproduction. In Shevell, S. K., editor, *The Science of Color (Second Edition)*, pages 281 – 316. Optical Society of America Oxford.

[Wang et al., 2004] Wang, Z., Bovik, A. C., Sheikh, H. R., and Simoncelli, E. P. (2004). Image quality assessment: From error visibility to structural similarity. *IEEE Transactions on Image Processing*, 13(4):600–612.

[Whelan and DeLair, 2012] Whelan, M. and DeLair, R. (2012). `http://www.edisontechcenter.org/ArcLamps.html`.

[Wikipedia, 2006] Wikipedia (2006). `https://en.wikipedia.org/wiki/Cathode_ray_tube`.

[Wikipedia, 2007] Wikipedia (2007). `https://en.wikipedia.org/wiki/CIE_1931_color_space`.

[Wikipedia, 2016] Wikipedia (2016). `https://en.wikipedia.org/wiki/Generalised_logistic_function`.

[Wikipedia, 2017a] Wikipedia (2017a). `https://en.wikipedia.org/wiki/HelmholtzKohlrausch_effect`.

[Wikipedia, 2017b] Wikipedia (2017b). `http://en.wikipedia.org/wiki/Standard_illuminant`.

[Wright, 1929] Wright, W. (1929). A re-determination of the trichromatic coefficients of the spectral colours. *Transactions of the Optical Society*, 30(4):141.

[Wyszecki and Stiles, 1982] Wyszecki, G. and Stiles, W. S. (1982). *Color science: Concepts and methods, quantitative data and formulas*. John Wiley & Sons.

[Yang et al., 2005] Yang, Y.-C., Song, K., Rho, S., Rho, N.-S., Hong, S., Deul, K. B., Hong, M., Chung, K., Choe, W., Lee, S., Kim, C. Y., Lee, S.-H., and Kim, H.-R. (2005). Development of six primary-color LCD. *SID Symposium Digest of Technical Papers*, 36(1):1210–1213.

[Yeo et al., 2008] Yeo, D.-M., Kwon, Y.-h., Kang, E.-J., Park, S.-K., Yang, B., Kim, G., and Jang, T. (2008). 64.4: Smart algorithms for local dimming LED backlight. *SID Symposium Digest of Technical Papers*, 39(1):986–989.

[Yuan et al., 2015] Yuan, J., Hardeberg, J. Y., and Chen, G. (2015). Development and evaluation of a hybrid point-wise gamut mapping framework. In *Colour and Visual Computing Symposium*, pages 1–4.

[Zamir et al., 2014] Zamir, S. W., Vazquez-Corral, J., and Bertalmío, M. (2014). Gamut mapping in cinematography through perceptually-based contrast modification. *IEEE Journal of Selected Topics in Signal Processing*, 8(3):490–503.

[Zamir et al., 2015] Zamir, S. W., Vazquez-Corral, J., and Bertalmío, M. (2015). Gamut extension for cinema: Psychophysical evaluation of the state of the art, and a new algorithm. In *Proc. of IS&T/SPIE Electronic Imaging*.

[Zamir et al., 2016] Zamir, S. W., Vazquez-Corral, J., and Bertalmío, M. (2016). Perceptually-based gamut extension algorithm for emerging wide color gamut display and projection technologies. In *SMPTE Annual Technical Conference & Exhibition*.

[Zamir et al., 2017] Zamir, S. W., Vazquez-Corral, J., and Bertalmío, M. (2017). Gamut extension for cinema. *IEEE Transactions on Image Processing*, 26(4):1595–1606.

[Zhu et al., 2016] Zhu, M., Hardeberg, J. Y., Wang, N., and Sun, B. (2016). Spatial gamut mapping based on guided filter. In *Proc. of IS&T/SPIE Electronic Imaging*, pages 1–4.

[Zhu et al., 2015] Zhu, R., Luo, Z., Chen, H., Dong, Y., and Wu, S.-T. (2015). Realizing Rec. 2020 color gamut with quantum dot displays. *Opt. Express*, 23(18):23680–23693.

[Zolliker and Simon, 2007] Zolliker, P. and Simon, K. (2007). Retaining local image information in gamut mapping algorithms. *IEEE Transactions on Image Processing*, 16(3):664–672.