Studies in ambient intelligent lighting

DOI:
10.6100/IR752369

Document status and date:
Published: 01/01/2013

Document Version:
Publisher's PDF, also known as Version of Record (includes final page, issue and volume numbers)

Studies in Ambient Intelligent Lighting

The work described in this thesis has been carried out at the Philips Research Laboratories in Eindhoven, the Netherlands, as part of the Philips Research programme.

Studies in Ambient Intelligent Lighting
by Dragan Sekulovski.
Eindhoven: Eindhoven University of Technology, 2013. Proefschrift.
A catalogue record is available from the Eindhoven University of Technology Library
ISBN: 978-90-386-3367-1

Cover design by Paul Verspaget

Studies in Ambient Intelligent Lighting

PROEFSCHRIFT

ter verkrijging van de graad van doctor aan de Technische Universiteit Eindhoven, op gezag van de rector magnificus, prof.dr.ir. C.J. van Duijn, voor een commissie aangewezen door het College voor Promoties in het openbaar te verdedigen op dinsdag 23 april 2013 om 16.00 uur

door

Dragan Sekulovski

geboren te Skopje, Macedonië

Dit proefschrift is goedgekeurd door de promotoren:

prof.dr. E.H.L. Aarts
en
prof.dr. I.E.J. Heynderickx

Copromotor:
dr.ir. R.A.W. Clout

Contents

1 Introduction **1**

1.1 Lighting Systems 4

1.2 Ambient Intelligence in Lighting Systems 5

1.3 Light as a new digital medium 7

1.4 Organization of the thesis 10

2 Digital Color **13**

2.1 Color and human vision 13

2.2 Digital color representation, color spaces 18

2.3 Digital content 22

I Light effect creation **25**

3 Selecting and rendering color **27**

3.1 Introduction 27

3.2 Related Work 30

3.3 Overview of the method and demonstrator 32

3.4 Using lyrics to create sets of images 35

3.5 Using the images to compute a set of representative colors 37

3.6 Experimental Results 44

3.7 Conclusions 50

4 Controlling spatial distribution of light **51**

4.1 Introduction 52

4.2 System overview . . . 53
4.3 Coded Light . . . 53
4.4 Interactive Light Design . . . 56
4.5 Prototypes and Usability Study . . . 63
4.6 Conclusions . . . 65

5 Spatio-temporal distributions of color 67
5.1 Introduction . . . 68
5.2 Light effect generation using stochastic models . . . 69
5.3 User study . . . 73
5.4 Extensions . . . 76
5.5 Conclusions . . . 79

II Temporal quality of light 81

6 Smoothness and flicker perception 83
6.1 Introduction . . . 84
6.2 Smoothness thresholds . . . 86
6.3 Speed of smooth transitions . . . 93
6.4 Flicker visibility thresholds . . . 96
6.5 Speed of transition for flicker . . . 99
6.6 Temporal control . . . 101
6.7 Conclusions . . . 102

7 Discrimination and preference of temporal color transitions 103
7.1 Introduction . . . 103
7.2 Discrimination of temporal color transitions . . . 105
7.3 Preference of temporal color transitions . . . 109
7.4 Discussion . . . 112
7.5 Conclusions . . . 113

8 Flicker perception in the periphery under mental load **115**

8.1 Introduction . . . 116

8.2 Flicker sensitivity thresholds at different eccentricities . . . 117

8.3 Effect of mental load on flicker perception . . . 122

8.4 Conclusions . . . 126

9 Conclusions and future work **127**

9.1 Light effect creation . . . 128

9.2 Temporal quality of light . . . 131

Acknowledgments **144**

Studies in Ambient Intelligent Lighting **146**

Curriculum Vitae **148**

1

Introduction

"The only objective way to measure a perceptual quantity is subjective"

Light is one of the main prerequisites of life. Sunlight sustains life, as the basic source of energy, that through photosynthesis goes into the food-chain. The omnipresence of electromagnetic radiation and its interaction with objects in the environment means that lifeforms that can detect these interactions have an evolutionary advantage. Thus light enabled the emergence of life's most advanced sense, vision.

Humans have extended the limited time that natural light is available by finding ways to produce it. The spark of the first man made fire liberated them from the dependence on nature's whim and marked the start of the modern society [68]. The development of artificial lighting, similar to the general development in sciences [61], followed a path of quick revolutions, bringing new ideas and technologies, followed by longer periods of steady evolution of the ideas and technologies [13]. The first revolution coincides with the first use of artificial lighting, self made fire, which was followed by a long period of development of combustion-based light sources. The second revolution was brought by the use of electricity in light generation, led by the venerable light bulb. As the technological understanding of electricity and the means of its transformation into light improved, so did the efficiency, the lifetime and the overall power of the light sources.

The lighting revolution we are, arguably, experiencing today is mainly supported by the new technology of high power light emitting diodes (LEDs) [50]. Even though LEDs have been in use since the early times of semiconductors, they were mainly used as indicators due to the low power of the emitted light. A series of recent developments have changed this fact. It is expected that LEDs will match the effectiveness of traditional light sources at comparable light levels and competitive prices by 2025 [77, 85, 31], while at the same time extending the lifetime of the light sources 3 to 10 fold. Organic LEDs (OLEDs) and polymer LEDs (PLEDs) complete the family of solid state lighting (SSL) sources.

The driving force behind the switch to solid state general lighting is the promise of better efficiency, longer lifetime, and environmentally friendly materials [31]. However, different application areas are adapting solid state light sources with different speed. The traditionally agile TV market has almost completely switched to LED back lighting within a very limited time span. Automotive lighting is another area where LEDs have seen fast growth, together with road signaling. However, general lighting, the largest part of the lighting market, has seen relatively slow adaptation of solid state light sources. While there are technological challenges behind the slow adaptation, there are also a number of important societal and economic reasons.

Most of the LED light sources on the market are developed as stand in replacements for older technologies, called retrofits. This direct correspondence comes with expectations on the price and quality of the new light sources. The initial price of solid state lighting on the market is between one and two orders of magnitude higher than traditional light sources. Even though the lifetime and energy efficiency savings outweigh the initial price difference, the initial cost is still a large hurdle for a more widespread market acceptance. The entrance of a large number of new companies, coming from the semiconductor business, to the lighting domain is driving the decrease of prices, but also introduces new problems. In some cases there is a large gap between the claimed and actual performance of solid state lighting on total light output, energy consumption, lifetime, and the quality of the produced light. The large variation between seemingly equivalent light sources, modules, and luminaires confuses customers and negative examples help the spread of the belief that solid state lighting is of inferior quality to the traditional light sources. To increase consumer confidence and facilitate quality control, a number of industry and government bodies have introduced quality testing programs. As one of the first, the US Department of Energy in 2006 started the CALiPER program [105]. More recently, the European Commission published a charter for LED Quality requirements [30] as a voluntary set of criteria limited to residential light sources.

Beyond the traditional advantages and retrofit applications discussed above, solid state lighting sources offer a number of new and unique differentiating features.

The first differentiator is the spectrum of the produced light. LEDs with a wide range of peak wavelengths having a relatively narrow spectral distribution can be produced by changing the semiconductor material and the dopants in the production process. This enables direct creation of highly saturated chromatic light without the need of filters, thus resulting in a much higher efficiency. By combining LEDs with different peak wavelengths and mixing the produced light or adding a fluorescent material, solid state lighting systems can produce a wide range of spectra. Even more, as the relative power of the individual spectral components can be changed, the resulting spectrum can be adapted, resulting in full color control.

The easy control of the intensity of the individual LEDs is an interesting side effect of the second advantage of LEDs, their fast response time. Most traditional light sources have a very limited range of intensity control, especially at shorter time intervals. Incandescent and halogen light sources are based on thermal processes, which are slow and usually result in reaction times on a seconds timescale. Furthermore, the spectrum of the produced light changes with the intensity of the light. Other traditional sources, including fluorescent, compact fluorescent and high intensity discharge sources, are even more limited and need times in the order of minutes to come to the nominal operating parameters or to be reactivated after being switched off. On the contrary, the response time of LEDs can be as low as $2ns$, due to the direct connection between the driving current and the amount of emitted light. This very fast switching enables the use of temporal modulation for intensity control, resulting in pulse width modulation (PWM [95]) and related control schemes.

The third unique advantage of LEDs is the large amount of light that can be produced from a small area. With the relatively low operational temperature and lack of radiative heat, a large number of individually controllable light sources can be tightly packed. As a consequence of the law of etendue conservation [16], smaller source area makes the design of optics that shape the light distribution to a required one much easier and cheaper. Even more, instead of only being part of a centralized source of light, i.e. the traditional luminaire, the light sources can be embedded into surfaces and objects in the environment whose primary purpose is not light production. The emergence of this so-called *embedded lighting* widens the definition of a luminaire, but also questions its existence as an entity separate from the rest of the environment. As artificial lighting freed humanity from the whim of nature, solid state lighting frees artificial lighting from the confines of the luminaire and truly integrates it back into the environment.

The combination of these advantages has brought the possibility to create spatially, spectrally and temporally complex light effects, with the sources of light embedded in the environment. The core of the revolution, arguably, is not in the basic technology, but in the control and interaction paradigm shifts it enables. With the newly found freedom the users can switch their focus from the confines of the technology to the expression of their needs, regardless of the technology. Identifying those needs, creating an effective language to communicate them to the system, and translating them to control signals that will fulfill them is the core of the revolution. Ensuring this will enable to go beyond retrofit applications and enable use of the full potential of solid state lighting. The change of the name on the basic enabling technology level to solid state lighting is not enough to encompass this change of focus; a new concept is needed. We will call this concept *Ambient intelligent lighting (AIL)*, which will be explained after a short introduction to lighting systems.

1.1 Lighting Systems

A lighting system consists of a set of light sources, a set of sensors, a set of user control devices and the interconnections between them. A simple example of a lighting system is a light switch with on-off functionality that is, via a power line, connected to an incandescent light source. Though simple, this example typifies most of the residential lighting systems in use today.

The history of the sources and their use has been discussed often, but other parts of the lighting system usually get comparatively lower attention. The interconnections used between the light sources and the user controls are very dependent on the application [95]. For most application areas, and especially in residential and industrial lighting systems, there has been relatively little progress from the basic form, the power line. On the other extreme are the stage lighting systems, like the ones in theaters and concerts, that typically have the most advanced controls and are the source of most innovations. One of these innovations was the introduction of digital communication channels, separate from the power line. This also mandates the move of the light source intensity control, from the site of the user control to the vicinity of the light source. In the case of some traditional light sources as well as solid state light sources, this only requires to extend the functionality of the already present driving electronics.

After a series of proprietary protocols, a widely used standard interconnect and communication protocol, DMX (Digital Multiplex, also called DMX512 [29]), emerged. Even though designed for stage lighting, DMX has been successfully used in other application areas, like architectural, retail and in rare cases residential lighting. Another standard digital interconnect, the DALI (Digital Addressable Lighting Interface [26]), emerged from office lighting applications, but has also been used in other application areas. Fueled by their popularity and wide availability, TCP or UDP over IP are increasingly being used as lighting interconnection, either by providing a direct interface on a luminaire level, or providing a bridge towards a local DALI or DMX based network. Following the same application path as the wired protocols, starting from entertainment and office lighting, a number of wireless interconnection protocols have been recently brought to the market. Most of the wireless protocols are proprietary, but a few candidates for a widely accepted standards, like Zigbee [33] and 6lowpan [47] are taking hold.

Similar to the interconnect, the user controls of lighting systems are most advanced in stage lighting systems. A variety of DMX controllers are available on the market, enabling the control of all the lamps and all the color channels in those lamps, as well as a multitude of other devices that use DMX controls. Similarly, a number of standalone user controls for DALI systems are available. In addition, a number of PC based solutions are available. By grouping light sources, defining scripted transitions and synchronizing the light output to an audio signal or output from a sensor, complex light effects can be created. In modern office applications the controls usually depend on the output of sensors, like the ones detecting occupancy or measuring daylight intensity. Most contemporary retail and residential lighting systems have simple on/off controls that are connected to a number of lights in a room or in an area. A smaller number of these systems use dimming controls, enabling the creation of a wider range of light intensities. A very limited number of residential systems use digital light controls, usually as a part of a wider home automation system.

One common factor in all the above mentioned control devices is their technological focus. In all cases, a clear link between the individual light sources in the environment and the lighting controls is maintained. Both the integration of the sources of light into the environment and their extended capabilities make these traditional lighting controls unsuitable for solid state lighting systems. Fortunately, parallel to the technological developments in lighting systems, there have been a number of revolutionary developments in the interaction with those technologies.

1.2 Ambient Intelligence in Lighting Systems

Artificial lighting, as one of the most fundamental technologies, has disappeared into the background of our lives, so much integrated that we simply take it for granted and forget how life was without it. Its basic use, extending the hours of the day in which we can function normally, has been supplemented by a number of related, but unique uses. Among them is the ability to transmit information, the most valuable good of the 21st century. The developments in imaging and video systems enable instantaneous communication of large quantities of data through our most natural communication means, vision.

Still, the way we interact with lighting at its basic use has been technology centered and explicit. The main interaction paradigms are traditionally based on the particularities of the technology, and not the needs of the users. In the development of human computer communication, Weiser [113] predicted that the combination of ongoing miniaturization in the semiconductor industry and device inter-connectivity would not only lead to better, faster computers but also to a new way of interacting with computers that could make computing as effortless and ubiquitous as writing or flipping a wall switch. Ironically, controlling the lighting using a wall switch is a good example of a traditional interaction paradigm that is based on the technology and not on the user needs. In this Ubiquitous Computing future, people would interact with many computers, in a wide variety of different shapes and sizes, each suited for a particular task, at the same time. Ambient intelligence [1, 71] builds further on this Ubiquitous Computing vision and accentuates the sensitivity and reactiveness of the environment to the presence of people. The goal of Ambient Intelligent environments is to support people in their everyday life activities through technology and media. By positioning the user and his needs in the center of attention and technology in a supporting role, Ambient Intelligence essentially adopts a user-centric design philosophy on Ubiquitous Computing [106]. This user-centric approach is central to the development of the ideas in Ambient Intelligent Lighting (AIL) and this thesis.

The largest differentiator of AIL is the position that the user takes in the system. Traditional systems require the user to have an extensive understanding of the system to be able to interact with it. The interaction is based on the language of the system, which in the area of lighting, usually amounted to flipping a switch. This requires knowledge not only of the capabilities of the system, but also of the connections inside the system. This traditional interaction paradigm is called *control driven*. The increased complexity of solid state systems, together with the limited ability of humans to effectively traverse a multidimensional space, make this type of interaction hard and cumbersome.

Contrary to this, AIL systems have a model of the user and the context, and this can bring the interaction at a level that is closer to the concepts of the user. This enables the users to imagine their target experience of the lighting in the environment and communicate it to the system using a language that is natural to them. In this case, instead of controlling the switch, the user describes the desired effect the lighting system should have on the environment, which is then translated into controls for the system by a set of intelligent, context aware algorithms. This interaction paradigm is called *effect driven*. The advantages to this approach are stressed even more in the case of embedded lighting, where a control driven design would be awkward and forced after loosing the conceptual connection between the wall switch and the luminaire it controls. Furthermore, in an ideal case, the details of any change in the hardware or connections in the system should concern only the system. The user should only be aware of the change of the capabilities of the system. For example, after an upgrade, the user should only notice an improvement in the execution of the desired effect, providing transparent extensibility and scalability.

The set of intelligent context aware algorithms, embedded in distributed devices, providing a personalized and adaptive control of everyday devices is one of the central ideas of Ambient Intelligence. In the context of lighting systems, its application enables, through effect driven control, easier use of the advanced capabilities of the modern light sources. The idea itself, however, is much broader and is applicable to any system capable of interactions with a user.

1.3 Light as a new digital medium

Ambient intelligent lighting is a broad emerging research area. To limit the set of challenges to tackle in this thesis, a few topics with the highest potential impact were selected. To be able to select these challenges, it is useful to look past the functional use of artificial lighting. The ease of inter-connectivity and control of large digital systems enables the creation of very complex lighting effects, changing at a high rate. This, in turn, enriches the possibilities of use of lighting as a communication device, capable of transferring information, as well as emotion. As a result, lighting is emerging as a new digital communication and artistic medium.

Imaging has recently also transformed into a digital medium. The digitization of the whole imaging pipeline has not only introduced a new level of flexibility, but also contributed to price decrease. This contributed to an explosion of available digital images and videos, democratizing digital imaging. Similar to the imaging case, to use the full potential of digital lighting, a few basic challenges have to be overcome. These can be classified into :

<table>
<tr><td>1. Creation or capture</td><td>2. Storage and transmission</td><td>3. Reproduction</td></tr>
<tr><td colspan="3">4. Measuring and controlling quality</td></tr>
</table>

Due to the fact that AIL is a broad and emerging area, the two basic challenges of new content creation and quality control are selected for further exploration in this thesis.

1.3.1 Content creation

The ease with which new content can be created or captured is the basis of any medium. In imaging, the transition from film to digital sensors enabled a more affordable and easier capture of reality from a wider array of users, exponentially growing the amount of produced content. At the same time, the use of digital image creation and rendering software added the possibility of creating new worlds and expressions.

The capture and creation of new content in the world of digital lighting, however, has seen relatively little progress. The distribution of light in the environment can be captured by measuring the radiant power incident on all the points in the environment, the so-called light field [76]. The capture of the light field, however, requires specialized hardware with an extended spatial, temporal, and intensity range. Even more, the captured light field is hard to reproduce on a different lighting system. Both these reasons have resulted in a very limited use of light field capture, mainly concentrated in the computer graphics domain.

A similar problem arises in the creation of a lighting effect. The easy interconnection between light sources (Dali, DMX), brought to life hardware and software that assists the creation of fully dynamic light scenes, but the created scenes are still linked to a specific hardware and hardware arrangement and cannot be generalized.

To trace the core of the problem better we can use a model similar to the Image Quality Circle Model of Engeldrum [28]. The Image Quality Circle is a process for managing the image quality of imaging products. It consists of four elements: *Technology Variables*, *Physical Image Parameters*, *Customer Perceptions*, and the *Customer Image Quality Rating*. The *Technology Variables* are the almost endless list of parameters or variables, hardware, firmware and software, that an imaging technologist trades-off to produce a known level of image quality. The *Physical Image Parameters* are the quantitative functions and parameters that are ascribed to images; generally, measured quantities that physically characterize an image or imaging system (image physics). *Customer Perceptions* are the components, or perceived attributes of images. Often the names of these perceptions have the suffix "ness", such as darkness, sharpness, colorfulness, etc. A shorthand notation that is used in the Image Quality Circle is "nesses." *Customer Image Quality Rating* is the overall quality measure. This number is determined by having customers, or customer surrogates, scale sample images and express an image quality judgment.

A similar model to the Image Quality Circle can be used to describe levels of abstraction a light effect can be defined at. This Ambient Intelligent Lighting model differentiates the following levels :

- *Hardware level*: all the hardware device properties and device settings.
- *Effect level*: the effect that the hardware produces on the environment. This level represents the physically measurable quantities.
- *Perceived effect level*: the perceived properties of the environment.
- *Perceived atmosphere*: the overall perceived atmosphere of the environment.

Moving from the hardware level to the perceived atmosphere level, the description of the lighting effect is transformed from a technology based one to a user based one. Unfortunately, as discussed above, most contemporary lighting systems are controlled using descriptions on the hardware level. This is the source of both the generalization and reproduction problems as well as the complexity challenges introduced with solid state lighting when compared to traditional lighting systems.

To be able to define the light effect in one of the more abstract levels, a suitable language, as well as a link between the primitives in the language and the physical level are needed. The exploration of this new language is the topic of the first part of this thesis. The studies are presented in an increasing order of complexity of the created light effects, from the selection of colors, through their spatial distribution in the environment, towards a full spatio-temporal distribution.

There are a limited number of examples of existing systems that use a higher level of abstraction. The amBX system [5] uses a extensible virtual representation of the environment based on a simplified compass spatial model. The devices in the system can be pre-assigned to different positions on the compass (N, W, SE, etc...) as well as to user defined locations. In the reproduction phase, a script that depends on time and external triggers can control the effect using the compass abstraction, providing a definition on the *effect level* of the AIL model.

Recently, a new and exciting way of describing the overall impression of an environment has been developed by Vogels [108]. In the so-called Atmosphere Perception model, the overall perception of an environment is defined through the affective evaluation of the environment. Vogels developed a method to measure the *affective evaluation* of a room, also called *atmosphere*. Factor analysis on a large space of possible descriptors revealed that atmosphere can be described by four dimensions: *coziness* (defined by words like cozy, pleasant, and intimate), *liveliness* (lively, exciting, inspiring), *tenseness* (tense, terrifying) and *detachment* (formal, business-like). The effect of several light characteristics of white light (light level, color temperature and spatial distribution) on atmosphere perception was studied and an effect on each of the four atmosphere dimensions was found.

Other systems use translation of information from a different medium to a light effect. Kater [52] gives an overview of systems that use features computed from audio to produce a light effect that follows music. Other systems use digital images or digital video as the input medium. The Philips Ambilight TV [82] computes features from the video stream that are then used to create lighting effects that enrich the experience of watching the video as well as lower eye strain [14, 91].

1.3.2 Light quality

Another important aspect that is essential both to the acceptance of the retrofit LED solutions as well as AIL systems is the perceived quality of the produced light effects. Light quality can also be defined on different levels. On the hardware level, parameters that influence light quality are the spectral content, the spatial distribution of intensity and chromaticity and the temporal modulations of the light output. On a higher abstraction level, these influence:

- The appearance of the light emitting surfaces;
- The appearance of space;
- The appearance of objects, plants, animals, and humans in the environment;

- The biological and chemical processes of objects, plants, animals, and humans;

This amounts to a large number of parameters of light quality, with relative importance that depends on the application. Boyce [9, 10] gives an overview of general parameters and their specific importance in different applications. To select the aspects of light quality that are the topic of the thesis, we turn to the peculiarities of solid state lighting and AIL.

As mentioned before, the temporal response of LEDs to changes in driving current is much faster than the response of classical light sources. This fact has two interesting consequences. First, LEDs are more susceptible to temporal artifacts like flicker and visibility of digital steps. Second, the study of temporal light changes, and especially chromatic light changes, has been very limited due to the inability of traditional sources to produce them. Both these consequences nominated temporal quality of light as the most interesting aspect of light quality to study in this thesis.

Contrary to spatial models of human vision, existing temporal models are designed to predict visibility of one specific temporal artifact, flicker, and are mainly luminance based, with very limited experiments on chromatic flicker. Improving and extending the models of temporal human color vision can be used to control the temporal quality of the produced light effects, from simple transitions to the effects produced by an Ambilight TV. The second part of the thesis explores the quality of the produced dynamic light effects through modeling the perception of smoothness, flicker visibility, and temporal path preference of dynamic light.

1.4 Organization of the thesis

Chapter 2 provides a short introduction of digital color and gives an overview of the definitions needed in the rest of the thesis. The rest of the chapters consist of original work that has been previously published and can be read separately.

The work in the thesis comprises of two main parts. The first Part (Chapters 3, 4, and 5) covers the creation of complex light effects using abstractions that simplify the process. The solutions are presented in an increasing complexity progression, from the selection of a set of colors using keywords in Chapter 3, through the creation of complex spatial distributions of light using a simple interaction paradigm in Chapter 4, to the creation and rendering of a full nondeterministic spatio-temporal distribution based on examples from nature in Chapter 5.

The second Part, through the exploration of the temporal properties of the human visual system, ensures the creation of dynamic light effects that are perceived as pleasant and devoid of digital artifacts. Chapter 6 explores the sensitivity of the human visual system to unsmooth light transitions and gives a contribution towards building a model of their visibility and building a boundary on the speed of smooth transitions based on the hardware capabilities. Chapter 7 deals with the preferred temporal path between two colors in a color space. Chapter 8 studies the difference between the temporal sensitivity of central and peripheral vision.

Chapter 9 provides an overall conclusion and gives a vision of an ambient intelligent lighting system that uses the results of the other chapters.

2

Digital Color

"For the rays, to speak properly, are not colored. In them there is nothing else than a certain power and disposition to stir up a sensation of this or that color."

Isaac Newton

Throughout this thesis, a few concepts and definitions are used repeatedly. This chapter serves as a central place for the definition of these concepts. The basics of human vision, color perception and mathematical models of color are presented. Building on the basics, digital color and image models are introduced and concepts used in the rest of the thesis are defined. This short introduction to digital color and imaging is by no means meant to be extensive and serves more as a starting point and a glossary of concepts needed to follow the rest of the thesis. The content of the chapter is based on [22, 32, 80, 94, 114]

2.1 Color and human vision

Color is easy to define intuitively. During the earliest years of life, humans experience colors and are taught the spoken language representations that correspond to those experiences. In most communication between humans, both spoken and written, this intuitive definition is sufficient as the need for high precision and discrimination is seldom present. Similarly, in classical reproductions of colors and images of the environment, like paintings, a human matches the reproduction to the reproduced environment or the memory of it.

Mass production, with its need for automated quality control, and recently the development of digital imaging have introduced more stringent requirements on the reproduction of colors. To be able to faithfully capture, store and then reproduce the visual representation of an environment with high precision the intuitive definition is not sufficient. Instead, a mathematical representation, a model, of color is needed.

When talking about properties of objects or light, it is common to say that they posses color. However, strictly speaking, because color is a subjective psychophysical phenomenon, there is no color without an observer. The attribution of color to objects/light is a particular instance of what psychologists refer to as a *stimulus error*, wherein a sensation experienced by an observer is identified with the stimulus causing that sensation. In our everyday activities, the usage of color as a property of objects and light is widely accepted and does not cause ambiguity. In technical applications, on the other hand, this might lead to significant conceptual errors.

The perception of color is a result of interaction between a physical stimulus, receptors in the human eye that sense that stimulus and the brain which is responsible for communicating and interpreting the signals sensed by the eye. This clearly involves several physical, neural and cognitive phenomena, which must be understood to comprehend color vision completely.

2.1.1 Light and color, physical basis

The physical stimulus for color is electromagnetic radiation with wavelengths between $\lambda_{min} = 360$nm and $\lambda_{max} = 730$nm, which is commonly referred to as *light* or *visible light*. The wavelength range of visible light is commonly referred as *visible spectrum*. Light, when it falls onto the inner back part of the eye, stimulates the visual receptors, ultimately causing the phenomenon of vision and perception of color.

Light, as electromagnetic radiation has its defining properties. Newton demonstrated that sunlight can, with the help of a prism, be decomposed into a spectrum of *monochromatic* components which cannot be decomposed further. Accordingly, light can be characterized by the amounts of energy in the constituent wavelengths, commonly referred to as *spectral power distribution (SPD)* of the light. Other properties, like the direction of polarization, are not differentiated by the human visual system. Thus, the SPD of the light falling onto the retina is the only property that influences the perceived color for a constant, isolated stimulus. The light that falls onto the eye may come from different sources. When viewing light sources or self-luminous objects, the light directly originates from the object being viewed. More commonly, the object being viewed is illuminated by an external light source. In this case, the SPD of the light entering the eye can be estimated from the product of the SPD of the light source and the spectral reflectance/transmittance of the object. Though this only holds for ideally flat surfaces, it is reasonably accurate if the object does not change geometry and the measurements are done under similar conditions. Absolute spectral power distribution for emitted or reflected light is typically specified in radiometric units (e.g. Watts per steradian per square meter) but the absolute SPD is rarely used in practice. Usually the SPD is represented on a relative scale with arbitrary units.

2.1.2 Human visual system

The sensory part of the human visual system consists of two different types of receptors, *rods* and *cones*. Rods are very sensitive to light and are primarily used for vision under low light conditions. This type of vision is called *scotopic vision*. In scotopic vision, only shades of gray can be perceived, and no chromaticity can be distinguished. An example of scotopic vision condition is starlight. Under typical light levels the rods get saturated and don't have a large contribution to vision. Instead, the less sensitive cones are active. The term *photopic vision* is used to describe this domain. Vision on intensity levels between scotopic and photopic is referred to as *mesopic vision*, and under these conditions both the rods and the cones are active. In the remainder of this thesis, photopic light levels and photopic vision models are used. A third type of photosensitive receptors, the photosensitive ganglion cells, have been discovered around the turn of the century [36, 44]. These receptors don't contribute to conscious vision beyond playing a role in the regulation of the sustained pupil size [41], but instead play a main role in the regulation of the circadian rhythm.

Observers with normal vision have three different types of cones which are responsible for color vision. These cones have photosensitive pigments with different spectral absorption characteristics. They are commonly called S, M and L cones, which is abbreviated from short, medium and long wavelength sensitive cones. Assuming that the human visual system can be represented by a vector space, an assumption supported by Grassmann's law, the response of the cones can be modeled by a system defined by the spectral sensitivities of the cones. Because in this case the system has a base that consists of three vectors, the resulting space of all the colors which can be represented by the model is tridimensional. In such a model, every perceivable color can be represented as a linear combination of the cone sensitivities. The resulting weight vector in the linear combination is known as a *tristimulus vector*, and the phenomenon is known as *trichromacy*. Some authors discuss the possibility of the perceivable color space having more than three dimensions [34].

Due to overlap of the sensitivities of the three types of cones in the retina, a difference signal between them carries more information. This fact supports a theory of color vision, called *opponent color theory*, which states that the color signals from the cones are transformed into three opponent color channels, a red-green, a yellow-blue and a dark-light achromatic channel. For further reading, Coren, Ward, and Enns [22] give a gentle introduction to vision and perception of color, while Wyszecky and Styles [114] give a broad overview of color science.

Because the space of all possible SPDs is more dimensional than the space of perceivable colors the mapping between them is not injective, i.e., light with different SPD can cause the same color sensation. SPDs that are different but cause the same color sensation are called *metamers*. Usage of metamers is central in digital color imaging. To reproduce a certain color sensation, we only need to reproduce some metamer of the light that originally caused the sensation. This is the basis of color science and of digital color imaging.

2.1.3 Basic color terminology

On the road to building a model for the perception of color, the basic quantities need to be named and their connection to the physical properties of the light changing these quantities established.

The terminology used in the field of color and light is broad and not always clearly defined. In literature, authors seem to have their personal preferred set of terms, each coming with their own definition. This is further complicated by the fact that in different applications, different aspects of color perception are dominant. This section gives a short overview of some basic color terms and their use, mainly following the definitions of International Commission on Illumination - CIE [19, 32].

The simplest color stimulus, named *unrelated color*, is an abstract stimulus consisting of a constant uniform color surface in isolation. An approximation of an unrelated color is a single light source uniformly illuminating a part of a dark environment. The basic perceptual quantities that can describe the experience of such a stimulus are defined as:

Hue: *Attribute of a visual sensation according to which an area appears be similar to one of the perceived colors: red, yellow, green and blue, or to a combination of two of them.*

Colorfulness: *Attribute of a visual sensation according to which the perceived color of an area appears to be more or less chromatic.*

Brightness: *Attribute of a visual sensation according to which an area appears to emit more or less light.*

As noted before, these attributes define the subjective experience of a stimulus that can be described by its SPD. A number of quantities, correlated with these attributes can be derived from the SPD of the light. A basic quantity related to brightness is intensity defined as :

Intensity: *A measure over some interval of the electromagnetic spectrum of the flow of power that is radiated from, or incident on, a surface. It is also called optical power and is expressed in Watts per steradian per square meter.*

Equal intensity light of different wavelength results in different levels of perceived brightness. A quantity that takes this difference in sensitivity to different wavelengths into account is luminance defined as :

Luminance: *Radiant power weighted by a spectral sensitivity function that is characteristic of vision. The magnitude of luminance is proportional to physical power but the spectral composition of luminance is related to the brightness sensitivity of human vision. Absolute luminance should be expressed in candelas per meter squared, but in practice it is often normalized to* 1 *or* 100 *units with respect to a specified or implied reference.*

Most color stimuli appear *in contrast* with other color stimuli next to them, or further in the visual surround. Furthermore, the composition of the visual field changes over time. Due to these dependencies, the human visual system constantly adapts to these changes in order to be able to estimate the reflectance or transmittance of an object, independently of its illumination condition. This results in a large influence of context, defined by both the spatial distribution of colors and the history of changes, on the perceived color. This complex stimulus, dependent on spatial and temporal context, is called *related color*.

As most colors in nature result from reflections, the most important part of the environmental context is the color of the light source, or the color of a diffuse white surface that reflects all incoming light, called a *white point*. Classical light sources, including daylight and black body radiators, have a limited range of spectral power distributions. To ensure an easier and more accurate exchange of color information, a standard set of these light source SPDs, or *standard illuminants*, were defined by the CIE. Examples of such illuminants are the CIE Illuminant E where white is defined as a uniform SPD across the visual light range of wavelengths. The A illuminant approximates an incandescent light source, while the D series, out of which D65 is most often used, represents different phases of daylight.

The definition of a white point changes in the case of light sources with a variable spectrum, among which we could count displays. More details are discussed later, after the introduction of device dependent color spaces.

A number of additional quantities can be defined for related colors :

Lightness: *The brightness of an area judged relative to the brightness of a similarly illuminated area that appears to be white or highly transmitting.*

$$Lightness = \frac{Brightness}{Brightness(White)}$$

Lightness can only be defined for related colors and should not be confused with relative luminance as described above.

Chroma: *Colorfulness of an area judged as a proportion of the brightness of a similarly illuminated area that appears white or highly transmitting.*

$$Chroma = \frac{Colorfulness}{Brightness(White)}$$

Saturation: *Colorfulness of an area judged in the proportion to its brightness.*

$$Saturation = \frac{Colorfulness}{Brightness} = \frac{Chroma}{Lightness}$$

2.2 Digital color representation, color spaces

To be able to represent color in digital form, a mathematical representation of the stimuli is needed. The earliest such systems use example patches of color to classify the perceptual space of colors. For easier navigation, these systems are ordered according to some of the perceptual attributes discussed above, hue, chroma and lightness for example. Such a system is called a *color order system*. Kuehni [59] gives a historical account of the evolution of color order systems. An example of such a color system still in use is the Munsell Color Order System [75]. Even though using color order systems a subset of colors can be represented in a digital form, any operations between representations are nonsensical as they assume no mathematical structure past partial linear or cyclic ordering.

A next step in modeling color is to associate a color sensation to a point in a vector space. In such a vector space, called a *color space*, each color is represented by a point with coordinates on a set of base vectors, called *primaries*, *color channels* or color axes. Unlike in color order systems, operations between color representations, like for example differences, are meaningful in a color space. The base vectors of the color space can be defined based on properties of a color rendering device, resulting in a *device dependent color space*, or based on the properties of human vision, resulting in a *device independent color space*.

2.2.1 Device independent color spaces

One of the earliest color spaces was the ***CIE XYZ*** color space, also known as CIE 1931 color space, defined by the International Commission on Illumination (Commission Internationale de L'clairage, CIE) in 1931. Many color spaces have been defined since for different applications, but the ***CIE XYZ*** color space has a special place, because it serves as the basis from which many others are defined. It was developed from a series of experiments done in the late 1920's by W. David Wright and John Guild . Their experimental results were combined into the specification of the ***CIE RGB*** color space, from which the ***CIE XYZ*** color space was derived. The axes in the ***CIE XYZ*** color space represent three virtual stimuli with SPDs that are not physically possible, but are selected such that any visible color can be represented as a positive linear combination of these stimuli. The value on the Y axis is the luminance of the stimulus, but the other two axes have no direct perceptual correlates. To make the chromaticity information easier to depict, ***CIE XYZ*** is usually transfered to ***CIE xyY***, another CIE color space with a luminance component $\boldsymbol{Y}$ and two chromaticity components x and y. The chromaticity coordinates x and y represent the projection of the vectors $\boldsymbol{X}$ and $\boldsymbol{Y}$ to the $\boldsymbol{X}+\boldsymbol{Y}+\boldsymbol{Z}=1$ plane respectively. The two dimensional graphical representation of the xy plane is called a *chromaticity diagram*. The chromaticity diagrams of the ***CIE xyY*** and the ***CIE Luv*** (discussed below) color spaces are the most often found representation of color. Figure 2.1 depicts the CIE 1931 chromaticity diagram.

CIE XYZ was designed as a standard way of representing and communicating absolute color, but the question of difference between colors was still open. MacAdam, based a series of experiments that tested the precision of color matching under different conditions, proposed the so-called macadam ellipses. The ellipses represent the set of colors around a center color that are indistinguishable from that center color. Based on the work of MacAdam, the ***CIE 1960 UCS*** color space was developed in 1960, designed such that an equal distance in the color space corresponds to an equal perceived difference in color, which implies that the space was designed to be *perceptually uniform*. ***CIE 1960 UCS*** was later succeeded by the ***CIE*** $\boldsymbol{L^*a^*b^*}$ and ***CIE Luv*** color spaces, also designed to be perceptually uniform. A cylindrical representation of ***CIE*** $\boldsymbol{L^*a^*b^*}$, namely ***CIE*** $\boldsymbol{LCh_{a^*b^*}}$, is often used as the axes are perceptual correlates of lightness, chroma and hue. Although ***CIE Luv*** and ***CIE*** $\boldsymbol{L^*a^*b^*}$ are a distinct improvement over the CIE 1931 system, they are not completely perceptually uniform, do not provide correlates to all the perceptual qualities defined above or account for all the effects of the surround on color appearance. To overcome the later issues, color spaces were succeeded by color appearance models.

CIE $\boldsymbol{L^*a^*b^*}$ is the simplest example of a *color appearance model*, as it takes into account the chromaticity of the illuminant and has correlates for lightness, chroma and hue. These properties are the minimum needed for a color space to predict the appearance of related colors. ***CIECAM02*** is the most recent in the series of color appearance models designed to encompass the effects and perceptual correlates missing

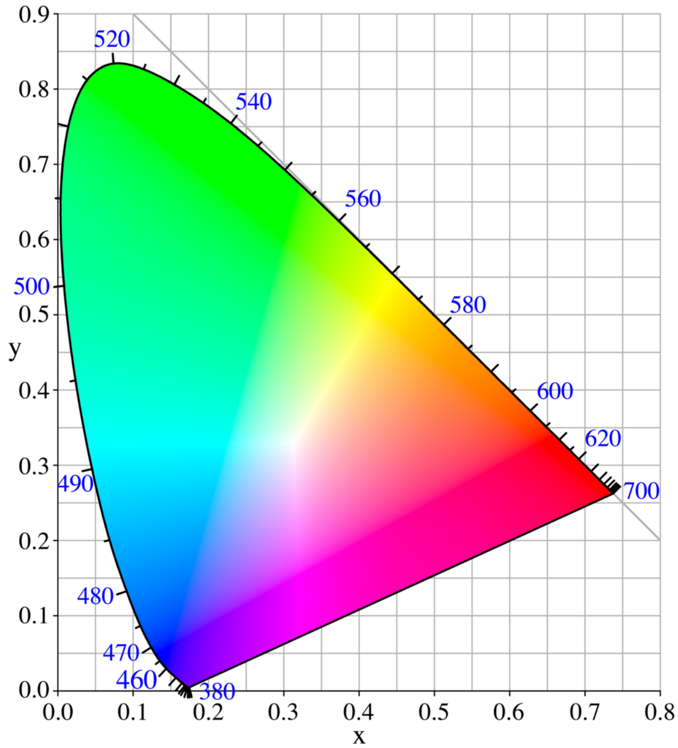

Figure 2.1: The ***CIE xyY*** chromaticity diagram

in ***CIE*** $\boldsymbol{L^*a^*b^*}$. It provides forward and inverse transforms from the ***CIE XYZ*** color space to the color appearance attributes taking into account the effects of the surrounding. This enables the reproduction of related colors in an environment that is different from the one where colors were recorded in. Fairchild [32] provides an overview and recommendations for the use of different color appearance models.

2.2.2 Color matching objectives

The dependence of the appearance of a color on the context makes the appropriate reproduction of a color dependent on the application and the environment in which it is reproduced. For example, a suitable reproduction for a single isolated light source used in illumination is to render the same unrelated color. In contrast, the appearance of a single pixel on a computer monitor, also a light source, depends on the surrounding pixels and the illumination in the environment and clearly needs to be treated as a related color.

Spectral Color Reproduction *represents the reproduction of the same reflectance or emittance spectrum of an object or an image. Historically important, this type of* *reproduction has been superseded by the metameric match based reproductions. In recent years, however, with the emergence of multi-ink printing and multi-primary lighting, the spectral capture and reproduction is regaining popularity.*

Colorimetric Color Reproduction *is defined as the metameric match between the original and the reproduction. This is done by matching the relative* ***CIE XYZ*** *values of the original and the reproduction, if taken under illuminants with the same relative spectral power distribution and equal surround. It should be noted that a colorimetric match only entails relative luminance matching, meaning that the luminance relations in an image are kept, but not the absolute luminance.*

Exact Color Reproduction *is defined as an absolute metameric match between the original and the reproduction. All* ***CIE XYZ*** *values under exact color matching are equal.*

Equivalent Color Reproduction *enables matching of the original and the reproduction under different illuminant chromaticities and intensities. An exact color reproduction under this condition would produce clearly incorrect results. To account for this, the notions of a color appearance model and chromatic adaptation were introduced. Thus, the equivalent match provides the same appearance of colors in the reproduction as in the original, independent of the illumination and surround conditions.*

Corresponding Color Reproduction *relaxes the requirements of the equivalent reproduction on the brightness / lightness matching. As it is not always possible to reproduce the same brightness / colorfulness of the original in the reproduction due to technology limitations, the corresponding reproduction only requires lightness / chroma matching.*

Preferred Color Reproduction *is the final and most subjective reproduction type. The correct reproduction of the original is not always the most preferred one. Artists may change parameters of the reproduction to produce a certain artistic effect. Most often, though, the preferred reproduction is made on a target medium and it is further transfered using equivalent or corresponding color reproduction.*

In the rest of the thesis, the *Corresponding Color Reproduction* using ***CIE*** $\boldsymbol{L^*a^*b^*}$ lightness and chroma matches is used.

2.2.3 Device dependent color spaces

CIE xyY, ***CIE Luv*** and ***CIE*** $\boldsymbol{L^*a^*b^*}$ spaces are models of human color perception and as such are device independent. In color capture and reproduction devices, on the other hand, different kind of color spaces, called device dependent color spaces, are used. A device dependent color space is connected to the way a device captures or reproduces color and it is limited to the colors the device can capture or reproduce.

Based on the phenomenon of trichromacy, most color capture and reproduction devices use three different types of sensors or variable light sources, with different wavelength sensitivity or SPD. The coordinates of the maximum input of the sensors or the maximum output of the variable light sources in ***CIE XYZ*** are called *color primaries* of the device and the device color space. The most often used color space for emissive reproduction and digital capture devices has three primaries denoted by ***R***, ***G***, and ***B***. The device dependent space is denoted by ***RGB***[1]. Any color that can be represented in the device color space is a linear combination of the primaries with constants in the range $[0,1]$. Knowing the primaries, the color represented in a device dependent color space can be transfered to ***CIE XYZ*** and vice versa. For a variable light source, the *whitepoint* is defined to be the chromaticity, described by the (x,y) point in a ***CIE xyY*** color space, which corresponds to the point in the device dependent color space having coordinates $(1,1,1)$ and it is a function of the ratio of power between the primaries. Sharma [94], Poynton [80], and Lindbloom [63] give an overview of the methods of calibration, modeling and transfer of color from one color space to another.

Gamma

Most traditional reproduction devices, like CRT monitors, have a nonlinear light intensity response to the variations in the control parameters, for example the driving current. Usually the nonlinearity can be approximated by a power curve. The power in the correction formula is called the *gamma* of the device and the power transform gamma correction [80]. For legacy reasons and coding efficiency, PC monitors use a standard gamma of 2.2. The components of the nonlinear device dependent color space are denoted by $\boldsymbol{R'}$, $\boldsymbol{G'}$ and $\boldsymbol{B'}$.

To ease the exchange of image information over the internet, a standard device depended color space, the sRGB color space [96] was introduced. The standard defines the chromaticities of the primaries and the nonlinear transfer curve.

2.3 Digital content

Digital representations of color, color sequences, images and videos are used in the subsequent chapters of the thesis repeatedly. This section introduces their definitions in the context of this thesis and the notation used.

The simplest building block for digital color content is a single color in a given color space.

[1]Different device dependent color spaces such as ***CMYK*** exist for devices that are based on reflected light, such as printers, but for this work only color spaces of emissive devices will be used.

*A **digital color** $\boldsymbol{c} \in \boldsymbol{C}$ is the digital representation in the color space $\boldsymbol{C}$ of a physical stimulus, that in a given context gives rise to a visual perceptual response. A color vector is represented with its coordinates on the set of primaries $\boldsymbol{c_i}, i \in \{1, \ldots, n\}$ of the color space $\boldsymbol{C}$.*

Colors can be grouped in a simple unordered set or arranged according to different index sets. A one dimensional index set produces a color collection, while a two dimensional one an image.

*A **color collection** $c = \{\boldsymbol{c}_i\}$ is a function $c : I \subset \mathbb{N} \to \boldsymbol{C}$, where $\boldsymbol{C}$ is a color space. We call the color $c(i) = \boldsymbol{c}_i$ an element of the collection. By $|c| = |\{\boldsymbol{c}_i\}|$ we denote the size of the collection, i.e. the number of elements.*

*A **digital color image**, or **image**, i is a function $i : P \subset \mathbb{N}^2 \to \boldsymbol{C}$, where $\boldsymbol{C}$ is a color space and $P \subset \mathbb{N}^2$ is the position space. We denote the space of images $\boldsymbol{I}$.*

We are often not interested in the whole image, but only a part of it that has a particular property or the distribution of colors.

*An **image region**, $i|_R$ is a restriction of the function $i : P \subset \mathbb{N}^2 \to \boldsymbol{C}$, to $R \subset P$.*

*A **color histogram**, or an **image histogram** $h^{\boldsymbol{C},\boldsymbol{m}}$ is a function $h : \boldsymbol{C} \to \mathbb{R}$, where $\boldsymbol{C}$ is a d dimensional color space and $\boldsymbol{m} \in \mathbb{N}^d$ is a discretization parameter. The histogram is a discretized representation of the distribution of colors $\mathcal{P}$ in a set of colors $\boldsymbol{C}$.*

In Chapter 5 and in the second part of the thesis (Chapters 6, 7, and 8) dynamically changing light effects are generated. The basis for their digital representation is the digital representation of time from which temporal sequences of colors and images can be defined.

*A **digital timestamp** or **timestamp**, t is a quantized representation of time used in digital media. It is an integer multiplier of a well defined base time t_{base}. We denote the space of timestamps $\boldsymbol{TS} = t_{base} \cdot \mathbb{Z}$*

*A **color sequence** $c = \{\boldsymbol{c}_t\}$ is a function $c : \boldsymbol{TS} \to \boldsymbol{C}$, from a space of timestamps $\boldsymbol{TS}$ to a color space $\boldsymbol{C}$. By $|c| = |\{\boldsymbol{c}_t\}|$ we denote the size of the sequence, i.e. the number of elements.*

*A **digital video**, or **video** $v = \{v_t\}$ is a function $v : \boldsymbol{TS} \to \boldsymbol{I}$, from a space of timestamps $\boldsymbol{TS}$ to the space of images $\boldsymbol{I}$. By $|v| = |\{v_t\}|$ we denote the number of frames in the video.*

*An **video region**, $v|_R$ is a restriction of the function v, mapping every frame v_t to its restriction $v_t|_R$, $R \subset \mathbb{N}^2$.*

Part I

Light effect creation

3

Selecting and rendering color

"I want meadows red in tone and trees painted in blue. Nature has no imagination."

Charles Baudelaire

Colors are the basic building blocks of an effect-driven lighting atmosphere design. In this chapter, the creation of a color palette based on words is explored. Images acquired through a web search, and subsequently processed, are used as an intermediary step between terms and colors and as a mean to bridge the semantic gap. Two algorithms for extracting representative colors from the acquired images are presented and are evaluated using color names and terms in English and Finnish. A measure of suitability of a word for color extraction is proposed. A prototype system based on this method is presented. The words are extracted from song lyrics synchronized to the audio signal of the song. Rendering the extracted colors and a selection of images that best represent the palette of colors together with the audio results in a full multimedia experience. This chapter is based on [88]

3.1 Introduction

Colors are the basic building blocks of atmosphere creation using lighting. A user of a lighting system might for example want to create a candle light atmosphere, but may not know the exact system settings to create it. A natural way to select these settings is to describe the atmosphere using other media in which the user can better express herself. This chapter introduces the idea of using human written language as the natural expression medium. In the example, the term "candle light" is the natural representation of the idea the user has. As the number of light sources becomes larger, the sources themselves become smaller and capable of producing a wide set of colors, the challenge presented to the user increases. Specially in the home environment, easy and natural controls are required in order to ensure the realization of the user's desires and expectations based on an effective use of the system capabilities. A set of algorithms is presented that enables unsupervised atmosphere creation based on terms. To illustrate the algorithms in a system, the application of generating lighting atmosphere based on music was selected as an example.

Light and music

Light has been used in conjunction with other media to enrich the overall experience. For example, it has been used in conjunction with music in disco and concert lighting. Recently, in the Philips Ambilight television concept, light effects were used to to enhance the experience of watching video. In the examples, the effect is most pronounced when the light effects are directly connected and time synchronized to the other medium.

The features that are traditionally used in atmosphere creation for music are low or mid-level, for example volume, pitch, energy in specific frequency bands, harmonic progression, etc. Using such features provides a good temporal correlation between the light effects and the music, but often fails to establish a semantic connection. Kater in [52] presented a method to enrich music with automatically computed light effects. In his work, the ***CIE*** $\boldsymbol{LCh_{a^*b^*}}$ [20, 114] Lightness, Chroma and hue of the light was computed based on the volume, the complexity of the sound and the pitch of the audio played. The results from the user study in [52] indicated that the users found the changes in Lightness and Chroma well connected to the music. However, the hue of the lights, when automatically produced, was seen as arbitrary. The users preferred to change the hue manually to account for semantics in the music. It is not unusual to expect a semantic connection between color names and the music through the lyrics. For example, when the Police sing *Roxanne put on the red light* one would also expect red light effects. When Wham's joyful *Club Tropicana* is played, different colors are appropriate than when R.E.M.'s *Everybody Hurts* is played.

Color representation

From the above, it is clear that there is a potential in using words or multi-word expressions from lyrics as a cue in the production of light effects. However, the computation of a color or a set of colors that correspond to a word or term is often not trivial. This is the case in even the simplest examples, when the word directly refers to a name of a color. One of the main difficulties is the difference in the way humans and machines represent and communicate color.

In every day activities colors are usually represented by their name. The color names, learned by example, are also used to communicate color between people. Contrary to that, machines use a digital representation of the light spectrum that produces the sensation of color. For a successful digitization of color and light, a connection between the two representations is needed. The first step towards this goal, i.e. a consistent machine representation of color, has been well addressed over the last decades by the CIE [20, 114] and led to a general digital definition of color in the form of color spaces like ***CIE XYZ*** and its derivatives.

The next step is the establishment of the actual connection between the machine representation of color and the terms humans use for the colors. The existence and definition of basic color categories (for example: black, red, green) and their most representative examples has been a matter of a lively discussion. The most popular view in recent years has been the one derived from the research of Berlin and Kay [7]. Using color terms from 20 languages and a review from 78 other languages, they discovered that the basic color vocabulary carries remarkable regularity among unrelated languages, giving the basis for an innate color categorization system. They identified 11 basic color categories (black, white, red, green, yellow, blue, brown, pink, orange, purple, and gray) and established color regions that correspond to those categories as well as the most representative examples for the categories. Even though there is a regularity among the languages studied, the variance between languages in the choice for both the boundaries and the most representative examples is also high. This led to a development of color categorization using fuzzy sets. Among others, Kay and McDaniel [53] built a model based on fuzzy set theory which included four basic colors (red, green, yellow and blue.). The same set of colors was used by Kuehni [60] to build a model in ***CIE*** $\boldsymbol{LCh_{a^*b^*}}$ connecting hue ranges to the selected colors categories. The last model is used in this work as part of the evaluation.

The World Wide Web

The Web nowadays is a place where people share thoughts, stories, experiences by texts and audio-visual content. This leads to billions of freely accessible images. The accompanying texts and tags provide annotations that are used by search engines. Very precise queries nowadays can produce images of the exact object in mind. Earlier work by Shamma et al. [93] shows that the enrichment of music by *Googled* images is possible and can be an attractive feature.

Main contributions

In this work we add the step of automatic selection of colors that are associated with the lyrics. We do so by selecting words or terms within the lyrics and by sending these as queries to an image search engine. The images retrieved are then used to compute a set of the most representative colors (i.e. the colors that are perceived to be the most present) among the images. We use the retrieved images and computed colors in a prototype system that renders them synchronized to the music.

The main contributions of this work are:

- An algorithm to extract the terms from lyrics which are the best candidates to retrieve relevant images that lead to computation of meaningful colors.
- Two alternative algorithms to extract the most representative colors in a set of images.
- A method to correct for the bias in the color distribution over images in large image repositories.
- A measure of suitability of a term for representative color extraction.
- A demonstrator which uses our results to add synchronized images and colored lights to music.

This chapter is organized as follows. In the following section the closely related work is discussed. Section 3.3 gives an high level overview of the overall method and the demonstrator. Section 3.4 handles the selection of the terms within the lyrics and the creation of a set of images per term. In Section 3.5 two alternative methods to identify a representative color for a set of images are presented. Section 3.6 handles the experimental results, and conclusions are given in Section 3.7.

3.2 Related Work

As most interesting for this work, we discuss related work on the connection between colors and language, extraction of representative colors from images, analysis of music and lyrics, and linking music to other media. To the best of our knowledge, there is no previous work that deals with the same subject in its entirety.

Most work relating colors to language is done using color names or memory colors and using human observers. Bartleson [6] used 50 observers to define 10 memory colors in the Munsell Color System. Moroney [74] used crowd sourcing to build a link between color names and ***RGB*** values. Recently, in a series of papers Lindner [65, 64] proposed a framework similar to the one proposed in this work. In the proposed framework, a set of images labeled with a specific keyword are retrieved from the Flickr online image sharing community. Subsequently, the image set is used to build a color histogram of colors present in the images and the mode of the color distribution is computed as the color associated with the keyword. Additionally, the significance of the result is computed using a z test. The work presented in this chapter differs from the framework in a few points. First, the algorithms presented in this chapter compute a set of colors instead of a single color and the relative contribution of all neighboring color bins are taken into account, diminishing the effect of noise. Second, the testing of the performance in the framework of Lindner is done using color names and memory colors while we use color names and arbitrary words. Finally, instead of only using a histogram based method that can produce quantization errors, a second algorithm based on mean shift is proposed in this chapter.

Representative color extraction is used extensively in content based image retrieval as one of the features used in the similarity computation between images. The terms *representative* , *salient* [18] and *dominant* [70] colors are used in prior work. Usually the representative colors are computed with a coarse quantization of a color space and computing representatives for the bins produced in the quantization. MPEG-7 Visual [70] defines dominant color as one of the low level visual descriptors for use in image and video indexing and retrieval. As part of the MPEG-7 experimentation platform [115] a Generalized Lloyd Algorithm, an adaptive k-means variant, in the ***CIE Luv*** space is used for extraction of the dominant color descriptor. Even though the notion of representative or dominant color is used, a clear definition is not given in literature.

In Kater [52] the Lightness, Chroma and hue of the lights are computed using the volume, complexity of the sound and pitch of the audio signal. The selection of the hue was studied and two different approaches were proposed. The user study however showed that most users picked their own hue and the automatic hue selection was seen as arbitrary.

Shamma et al. [93] present MusicStory, where they introduce the concept of automatic music slide show creation using lyrics. Our work differs from MusicStory in a number of aspects. We scan the lyrics for terms rather than single words. In MusicStory the beats are used to display images, i.e. contrary to our approach no synchronization of lyrics with images and music is applied. Next to the images, we extract and render colors synchronized to the audio and the images. As part of the demonstrator we use lyrics synchronization, a currently well-addressed topic. LyricAlly [112] is a system where music and lyrics are aligned. However, the song structure of *verse-chorus-verse-chorus-bridge-outro* is pre-assumed. This simplifies the task of categorizing both the audio and the lyrics. We present a more generalized lyrics synchronization method, where we discover both the structure of the audio and the lyrics before aligning the two. The same approach can be further developed, as shown by Iskander et al. in [49], where a synchronization on the syllable level is achieved using a dynamic programming table of the lyrics and the texts recognized by a speech recognizer. Chen et al.

[17] use vocal/non-vocal detection to segment music. The vocals within the audio are processed by a speech recognizer. Forced alignment is then applied on the output to match the lyrics. Although the algorithm is only evaluated on six (Chinese) songs, the method is promising.

In earlier work on the retrieval of images relevant to music [84], album covers are retrieved given an artist and album title. We do not focus specifically on album covers, although album covers typically appear among the images when the title is one of the identified terms (e.g. *Yellow Submarine*).

3.3 Overview of the method and demonstrator

One of the possible ways of translating terms to colors is to have a predefined database of terms and their corresponding sets of colors. This approach, however, limits the amount of terms that can be used and the possibility to use the method for a wide set of languages. To overcome these shortcomings of the fixed database method, we use images as an intermediate step between terms and colors. The advantage of using images is that we have an easy interface to a large, albeit noisy, database of images connected to terms.

One disadvantage of this approach is the large time latency, mainly due to the amount of time needed to retrieve the set of images for a given term. To overcome this, a hybrid approach is used in the demonstrator. A database of terms and corresponding colors is used, but the database is build from the terms extracted from the songs. Furthermore, the experience is enhanced by the addition of a set of images corresponding to the term and shown at the same time.

Using this approach, the system consists of two computational parts. First, an off-line part, having a long computation time, producing a database of terms and their corresponding images and colors. Second, an on-line part, that renders the colors and images at the appropriate time, synchronous to the audio. The on-line part of the demonstrator consists of a music player that enables user interaction with the set of songs in the library, an atmosphere creation agent and a set of lighting and imaging devices. An overview of the system is given in Figure 3.1, and overviews of the two parts are given in Algorithm 3.1 and Algorithm 3.2.

The off-line part of the system is executed on the set of songs S in the collection and stores the output in a database for subsequent use in the on-line part. The off-line part needs to be executed for all new songs in the collection.

The unsynchronized lyrics $LU(s)$ of a song s are retrieved using the World Wide Web. In this work, we do not elaborate on the retrieval of the lyrics, but we refer to earlier work on automatic lyrics extraction from the web [58, 39].

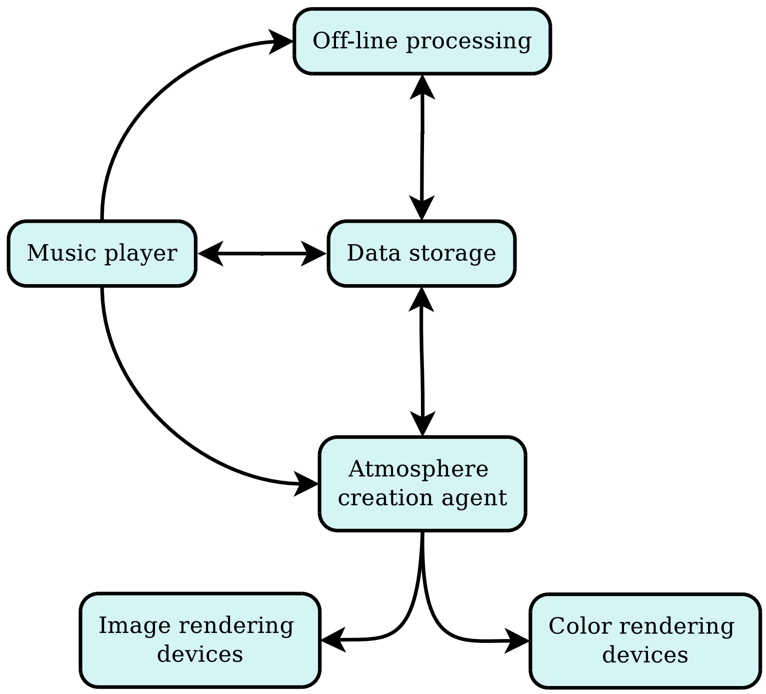

Figure 3.1: Overview of the system parts

To synchronize the lyrics and audio, first a method of synchronization on the stanza level is used [40]. The audio is segmented into stanzas, where segments with similar harmonic progression are assigned with the same label. The lyrics are scanned for similar fragments (i.e. the choruses) and fragments with a similar textual structure (i.e. the verses). We produce a labeling of the lyrics stanzas based on these similarities. Now, we have independently computed two sequences of labels, one for the audio and one for the lyrics. By expressing the length of the lyrics stanzas in syllables, we can compute a most likely mapping between the two given the durations of the audio segments.

Given the stanza synchronization, we now can approximately synchronize the terms with the audio. For a given stanza, its length is now expressed both in seconds and in the number of syllables. These parameters are used to estimate the moment the terms identified are being sung in the music. We assume that the syllables sung in the lyrics are of equal length in seconds. Hence, suppose a stanza has a length of l seconds and contains n syllables. Then the term starting at syllable m is assumed to be sung at $\frac{l \cdot m}{n}$ seconds after the start of the stanza.

The rest of the off-line processing functions represent the core of this work and are elaborated on in subsequent sections.

Even though the off-line processing is represented as a centralized process, this is not

Algorithm 3.1 Overview of the off-line part of the system

Input

set of songs S , digital audio $A(s)$, unsynchronized lyrics $LU(s)$, $s \in S$

Output

set of terms T, $\forall s \in S$, term timestamps $TS(s)$,
synchronized terms $T(s) = \{(ts,t)|ts \in TS(s), t \in T\}$,
term representative colors $\boldsymbol{C}(t)$, term representative images $R(t)$

---

for all $s \in S$ **do**

(TS, L) = synchronize_song$(A(s), LU(s))$

$(T(s), TS(s))$ = select_terms(TS, L)

$T = T \cup T(s)$

for all $t \in T(s)$ **do**

I = retrieve_image_set(t)

$\boldsymbol{C}(t)$ = extract_representative_colors(I)

$P(t)$ = select_images$(I, C(t))$

---

a requirement. In an ambient intelligent environment, with a large number of devices, processing agents can run on any of the devices with the appropriate capabilities and carry out subtasks. Due to the independent nature of the *for all* statements, the subtasks can be delegated at those points and computed in parallel.

The on-line part of the demonstrator consists of a music player that enables user interaction with the set of songs in the library, an atmosphere creation agent and a set of lighting and imaging devices. With the music player, users can play and control the position in a song, can create, edit and schedule play-lists. The music player has an interface to the atmosphere creation agent to which the current song and the position in the song can be communicated. Furthermore, the music player can communicate any special events like the beginning, pause or the end of a song. Using this information and the precomputed output of the off-line part, the atmosphere creation agent renders the light and the images at the appropriate time instances. The lights and images are rendered on the connected devices through a predefined interface.

The atmosphere creation agent serves as the point of synchronization between the music player and the output devices. It is aware of the number of, capabilities of, and the relationships between the devices. The set_neutral_colors(), set_neutral_info(), set_song_info(s), render_colors$(\boldsymbol{C}(t))$, and render_images$(R(t))$ methods serve as central dispatching points for the corresponding methods of the devices.

Algorithm 3.2 Overview of the atmosphere creation agent

Input
S, $A(s)$, $TS(s)$, $T(s)$, $\boldsymbol{C}(t)$, $R(t)$

Output
light and images, synchronized with the music

On start_song($s \in S$) **do**

set_neutral_colors()
set_song_info(s)

- - -

On stop_song($s \in S$) **do**

set_neutral_colors()
set_neutral_info()

- - -

On timestamp_reached($ts^*, s \in S$) **do**

if $ts^* \in TS(s)$ **do**
$t = T(s)(ts^*)$
render_colors($\boldsymbol{C}(t)$)
render_images($R(t)$)

3.4 Using lyrics to create sets of images

We are interested in the words in the lyrics that are the best candidates to generate relevant images. As objects depicted in images can in general be described by noun phrases (e.g. *little red corvette*, *crazy little thing called love*, *brown eyed girl*) and proper nouns (e.g. *Eindhoven*, *Prince Charles*, *Johnnie Walker*), we focus on these groups of words within the lyrics. We use these noun phrases and proper nouns to query large image repositories (`flickr.com`, `images.google.com`). The sizes of these image repositories enable querying these precise terms. One benefit of querying noun phrases instead of single words is that it leads to disambiguation. For example the term *wiley, windy Moors* (taken from *Wuthering Heights* by Kate Bush) reflects the moors in Yorkshire rather than the medieval Muslims in Iberia.

We identify terms within a text using a part of speech (PoS) tagger [11]. In this case we used QTag by Oliver Mason, since this package is provided with training data in English (`www.english.bham.ac.uk/staff/omason/software/qtag.html`). Since PoS taggers are typically trained on newspaper corpora, we first transform lyrics into somewhat more prosaic English. We consider each lyrics line as a sentence, so we capitalize the first letter and add punctuation marks where necessary. Then, we rewrite common slang and abbreviations. Hence, abbreviations like *gonna* and *kinda* are rewritten to *going to* and *kind of*. Words ending with *in'* are rewritten to *-ing*, thus *kiddin'* becomes *kidding*.

We tag the adapted lyrics using QTag. The tags are used to identify noun phrases and proper nouns. We select the longest sequences of nouns possibly preceded by adjectives. We send each term within the lyrics as a query to an image search engine. In our experiments we used both *Flickr* and *Google Images*, where we selected the option to only download full-color images. For each term we retrieve up to *max_images* images and store them in the order as they are presented by the search engine. Hence, we make use of the ranking provided by the search engines [12]. However, if the number of images found using a multi-word term does not surpass a threshold *min_images*, we broaden the query by removing the first word. For example, *little red corvette* can be broadened by querying first *red corvette* and consequently *corvette*. We remove duplicates by comparing the URLs of the images.

This phase resolves in a annotated lyrics file, where each term is annotated with a list of at most *max_images* images. Figure 3.2 depicts an example query from one of the image search engines.

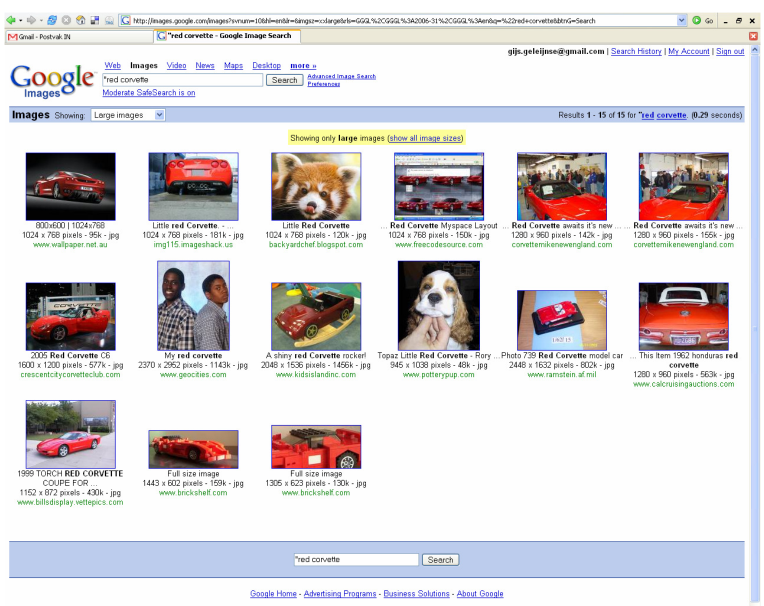

Figure 3.2: Example Query for the term "Red Corvette"

3.5 Using the images to compute a set of representative colors

After the set I of images for each term t has been identified, we use the color information from the set to compute a set $\boldsymbol{C}(t)$ of representative colors for t. For each color $\boldsymbol{c} \in \boldsymbol{C}(t)$, we also compute a measure of relevance $r(\boldsymbol{c}, t)$ of that color for the term t.

Two algorithms for representative color extraction are described, one using a histogram in Section 3.5.2 and one using mean shift in Section 3.5.4. As noted before in [51] the color distribution in images that can be found with image search engines is highly biased towards non-chromatic, or less saturated, colors. The bias is strongest towards white and black. We present a method to estimate and correct for the bias in Section 3.5.3.

3.5.1 Computing representative colors

The definition of a representative color is highly dependent on the context in which the extraction is done. The importance of an exact definition is diminished in our work because we extract a set of representative colors rather than one representative color. In this work, we define a color to be in the set of representative colors if it is a local maximum in the density, i.e. the local mode, of the distribution of colors found in an image or set of images and has a corrected frequency above a certain threshold. The frequency of appearance is corrected for the color bias as described in Section 3.5.3 Both methods, described in Sections 3.5.2 and 3.5.4 are mode finding methods.

The extraction is done in two stages. First, for each of the images in $I(t)$, the representative colors are extracted. This can be the set of colors which have a count larger than a threshold (Section 3.5.2) or the set of local maxima with a small granularity (Section 3.5.4). In the case of extracting local maxima, we keep only the maxima that have a large enough contribution for the region they represent. This first step is done for two reasons: (1) to take into account only the presence of a color, and not the relative contribution, something that in our experience would bias the distribution in the color set for all images even more and (2) to account for the colors that are present only as noise.

In the second step the overall set distribution $\mathcal{P}(t)$ is computed for all images in $I(t)$, using the color sets built previously. The distribution is then corrected to account for the bias in the distribution of colors. After the correction, the set of representative colors is extracted by finding the local maxima of the distribution with a given granularity. The exact details depend on the algorithm that is selected. Beside extraction of the set, we also compute a relative contribution measure for every color in the set and filter the colors that have a significantly lower contribution than the one with the highest contribution in the set.

3.5.2 Histogram-based method

The basic object this method uses is the color distribution $\mathcal{P}(t)$ over all the images in the set $I(t)$ for a term t. The color histogram $h^{\boldsymbol{C},m}(t)$ over all images in the set described in a color space $\boldsymbol{C}$ and with m bins per color channel is an estimate of that distribution. We assume $\boldsymbol{C}$, t, and m to be fixed in the rest of the section and denote the color histogram by h and the image set with $\boldsymbol{I}$. As described in the overall method overview, in order to filter out the colors that appear in the image only as noise and to reduce the bias we do the extraction in two steps.

First, we quantize the representation of every image $i \in \boldsymbol{I}$ in the color space $\boldsymbol{C}$. This is done with uniform quantization on every color channel and with the same number m of bins per channel. Next we build a joint histogram $h_i(x,y,z)$, for $1 \leq x,y,z \leq m$ being color channels in a three dimensional color space. Next, the histogram is filtered to remove the bins with a relative frequency lower than a computed threshold t_1. To compute the threshold t_1 we use the fact that an image that has only noise will have a uniform distribution over the bins in the histogram. The outliers from the uniform distribution will have bin counts larger than twice the mean bin count. Taken this into account we compute t_1 for an image i as $t_1 = \frac{2 \cdot |i|}{m^3}$. Thus we get a set of colors which are mainly present in the image in the nonzero count bins in the histogram $h_i(x,y,z)$.

Given all the image histograms $h_i(x,y,z)$, we compute the histogram h over the whole set of images as

$$h = \sum_{i=1}^{|I(t)|} \operatorname{sgn}(h_i(x,y,z)), \tag{3.1}$$

for $1 \leq x,y,z \leq m$. The extended sign function $\operatorname{sgn}(u)$, defined as

$$\operatorname{sgn}(u) = \begin{cases} -1 & u < 0 \\ 0 & u = 0 \\ 1 & u > 0 \end{cases} \tag{3.2}$$

is used to calculate the number of appearances of the colors, rather than the area which the colors cover in the images.

Next the set of locally representative colors is identified by searching for colors that are local maxima in a square neighborhood $\mathcal{N}^{ns}$ with size ns around them. The neighborhood is defined on the histogram bins. A color $\boldsymbol{c}^* = (x^*, y^*, z^*)$ is added to the set $\boldsymbol{C}$ if

$$h(\boldsymbol{c}^*) = \max_{\boldsymbol{c}' \in \mathcal{N}^{ns}(\boldsymbol{c}^*)} h(\boldsymbol{c}'). \tag{3.3}$$

For all colors $\boldsymbol{c}^*$ in the set $\boldsymbol{C}$, their relative contribution $r(\boldsymbol{c}^*)$ is computed, taking into account neighboring colors they represent. The computation of $r(\boldsymbol{c}^*)$ for a representative color $\boldsymbol{c}^*$ is done by summing the relative frequencies of all colors $\boldsymbol{c}$ that have a path of steepest ascent on the relative frequencies to $\boldsymbol{c}^*$. A given color $\boldsymbol{c}$ is on the path of steepest ascent to $\boldsymbol{c}^*$ if $\boldsymbol{c}^*$ is the local maximum in the neighborhood $\mathcal{N}^{ns}(\boldsymbol{c})$. The set of representative colors $\boldsymbol{C}$ is then ordered according to those contributions $r(\boldsymbol{c}^*)$. To filter the noise in the set all representative colors which have a relative contribution smaller than a threshold t_2 are removed from the set. The threshold can be a fixed number, for example all the colors with a relative contribution of less than 5% can be removed. The threshold can also be computed from the properties of the set, for example all colors with a contribution less than the mean contribution over all colors in the set can be removed.Algorithm 3.3 depicts the histogram method.

The size of the neighborhood ns and the number of bins m are parameters that can be varied to produce the desired size and granularity of the resulting set of colors. Using a larger number of bins also has an impact on the estimate of the distribution. A small value for the number of bins m will give a rougher estimate of the distribution in the images, but at the same time will allow for better noise removal and for a smaller impact of the variance of the appearance in color over the images. The desired number of colors in the resulting set depends on the application.

The choice of a color space is also a variable in the method. The results presented in the evaluation are computed using sRGB (we assume the sRGB color space for all images) without gamma correction. A test with $\boldsymbol{CIE\ L^*a^*b^*}$ showed only a slight improvement, with the biggest hurdle being the harder quantization of the $\boldsymbol{CIE\ L^*a^*b^*}$ space, due to its complex shape. Other studies in related subjects like skin color detection noted a small impact of the selection of the color space on the result. For an overview of related work with the same conclusion we refer to [107]. A much larger impact on the results comes from the observed bias in the distribution of colors over a large set of unrelated terms. Due to the importance, we give a more detailed discussion of this topic in the following Subsection.

3.5.3 Color bias correction

The method described above assumes that the distribution of colors over a very large randomly selected image set would be uniform. However, during our experiments with a large set of terms we observed that the distribution of colors is highly non uniform. This fact is also shown in previous work [51] where a non-uniformity correction is used for improvement of a skin color detection algorithm.

When assuming a uniform distributing of colors, the most prevalent colors over all terms were black and white. Moreover there was a general color bias towards desaturated colors. Hence, the colors that would be perceived to be the representative were not identified as such. Figure 3.3 shows a view on the marginal distribution of the colors estimated as representative using a large image database on the ab plane of

Algorithm 3.3 (histogram) method : $\boldsymbol{C}$ = extract_representative_colors(I, m, ns)

Input

set of images I, number of bins m, neighborhood size ns

Output

set of representative colors for the set of images $\boldsymbol{C}$

$h(\{1,\ldots,m\},\{1,\ldots,m\},\{1,\ldots,m\}) = 0$

for all $i \in I$ **do**

$\quad h_i(\{1,\ldots,m\},\{1,\ldots,m\},\{1,\ldots,m\}) = 0$

$\quad$ **for all** $(x,y) \in \{1,\ldots,\text{width}(i)\} \times \{1,\ldots,\text{height}(i)\}$ **do**

$\quad\quad h_i(\lfloor \frac{i(x,y)}{256/m} \rfloor) = h_i(\lfloor \frac{i(x,y)}{256/m} \rfloor) + 1$

$\quad t_1 = \frac{2|i|}{m^3}$

$\quad$ **for all** $\boldsymbol{c} \in \{1,\ldots,m\}^3$ **do**

$\quad\quad$ **if** $h_i(\boldsymbol{c}) > t_1$ **do**

$\quad\quad\quad h(\boldsymbol{c}) = h(\boldsymbol{c}) + 1$

for all $\boldsymbol{c} \in \{1,\ldots,m\}^3$ **do**

$\quad \boldsymbol{c}^* = \boldsymbol{c}$

$\quad$ **while** $h(\boldsymbol{c}) < \max\limits_{\boldsymbol{c}' \in \mathcal{N}^{ns}(\boldsymbol{c}).} h(\boldsymbol{c}')$ **do**

$\quad\quad \boldsymbol{c} = \arg\max\limits_{\boldsymbol{c}' \in \mathcal{N}^{ns}(\boldsymbol{c})} h(\boldsymbol{c}')$

$\quad h(\boldsymbol{c}) = h(\boldsymbol{c}) + h(\boldsymbol{c}^*)$

$t_2 = 0.05 \max\limits_{\boldsymbol{c} \in \{1,\ldots,m\}^3} h(\boldsymbol{c})$

for all $\boldsymbol{c} \in \{1,\ldots,m\}^3$ **do**

$\quad$ **if** $(h(\boldsymbol{c}) = \max\limits_{\boldsymbol{c}' \in \mathcal{N}^{ns}(\boldsymbol{c}).} h(\boldsymbol{c}'))$ *and* $(h(\boldsymbol{c}) > t_2)$ **do**

$\quad\quad \boldsymbol{C} = \boldsymbol{C} \cup \{\frac{(\boldsymbol{c}+0.5)}{m/256}\}$

CIE L*a*b*. Figure 3.4 shows the marginal distribution of the same color distribution on the $\boldsymbol{L}$ axis of ***CIE L*a*b****. The bias towards desaturated colors and skin tones (reds and yellows) is clearly visible in Figure 3.3 and confirms the findings in [51]. The model used in [51] is not directly usable in our method because we count only the appearance of a color in an image, while Jones et al. take into account the area that the color covers in the images.

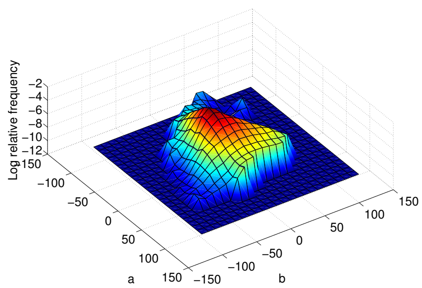

Figure 3.3: Histogram of the estimated color distribution for a random image set in the *ab* plane.

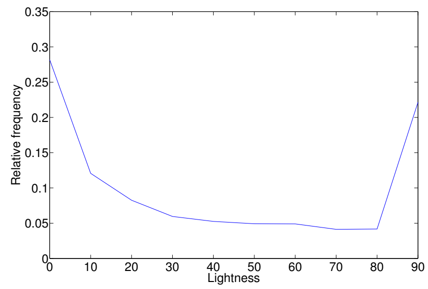

Figure 3.4: Histogram of the estimated color distribution for a random image set on the ***L*** axis.

To correct for the non-uniformity, the frequency in the set histogram $h(\boldsymbol{c})$ is corrected depending on the properties of the color $\boldsymbol{c}$. The simplest correction takes into account the saturation of the color and gives a weight which is proportional to the saturation, thus correcting for the bias towards the desaturated colors in the overall distribution. However, a function of this type is only applicable when we are interested in fully chromatic colors, as it tends to over express the saturation of the resulting colors. Another problem with this approach lies in the fact that different, even fully chromatic colors, are represented at a different saturation level in the set of images. Red and yellow, for example, are represented at near to full saturation, while specially green tends to be present at lower saturation levels. A simple correction would thus result in under representation of green colors in the resulting representative color set.

To take into consideration the full non-uniformity in the distribution, we need to build an estimate for the distribution of representative colors for an ideal random set of images. To estimate this distribution, we used the union of the color histograms of a collection of terms, part of which are the terms given in Table 3.3.

Having computed the histograms for all the terms in a set $t \in T$, we estimate the distribution $\hat{h}$ as

$$\hat{h}(\boldsymbol{c}) = \frac{1}{|T|} \sum_{t=1}^{|T|} h_t(\boldsymbol{c}), \tag{3.4}$$

where all the histograms $h_t(\boldsymbol{c})$ are first normalized to a sum of unity.

Color bias corrected histogram method

After the estimate $\hat{h}(\boldsymbol{c})$ is computed, the histograms for the terms are normalized using $\hat{h}$ as

$$h'_t(\boldsymbol{c}) = \frac{h_t(\boldsymbol{c})}{\hat{h}(\boldsymbol{c}) + const}, \tag{3.5}$$

where $h_t(\boldsymbol{c})$ is the non-modified term histogram, and *const* is a parameter introduced to control the influence of the bins with a small probability of appearance. The value of *const* is typically in the order of the average probability of appearance for all the bins in the distribution $\hat{h}(\boldsymbol{c})$. The value of *const* approaches zero as the histogram becomes more fine-grained. Hence we conclude that the need for an additional term can be explained by the quantization errors due to the use of a histogram.

3.5.4 Mean shift

The histogram method has limitations in the precision with which the representative colors can be extracted. The precision can be improved by using smaller bin sizes, but smaller bin sizes also render the method more sensitive to noise and to small changes in color appearance due to natural object color variation, illumination changes or camera properties. Using the histogram method in perceptually uniform color spaces is hard due to the properties of those spaces which make the quantization a much harder problem than in the case of ***RGB***. To overcome these limitations of the histogram method, we developed a method based on the mean shift algorithm.

Mean shift is a nonparametric method first introduced by Fukunaga [37] and recently made popular again in computer vision applications by Commaniciu and Meer [21]. It is a robust mode finding and data clustering method. The extraction is done as previously in two steps: first per image and then for the whole set of images.

Given n data points $\boldsymbol{x_i} \in R^d, 1 \leq i \leq n$, a radially symmetrical kernel $K = c \cdot k(\|\boldsymbol{x}\|^2)$ and having $g(x) = -k'(x)$, the mean shift for a starting point $\boldsymbol{x^0}$ is an iteration of the form

$$\boldsymbol{x^{k+1}} = \frac{\sum_{i=1}^{n} \boldsymbol{x_i} g\left(\left\|\frac{\boldsymbol{x^k} - \boldsymbol{x_i}}{hs}\right\|\right)}{\sum_{i=1}^{n} g\left(\left\|\frac{\boldsymbol{x^k} - \boldsymbol{x_i}}{hs}\right\|\right)}, \tag{3.6}$$

where hs is the unit window or kernel size. The size of the kernel is used to tune the number of segments and so, the number of resulting representative colors. The iteration is done as long as $\|\boldsymbol{x^{k+1}} - \boldsymbol{x^k}\| > \varepsilon$. We used an Epanechnikov kernel with a profile $k(x)$

$$k(x) = \begin{cases} 1 - x & x \leq 1 \\ 0 & x > 1 \end{cases} \tag{3.7}$$

The Epanechnikov kernel was used due to the speed of convergence and the simplicity of its derivative and thus the function $g(x)$.

The first step is done along the same line with Comaniciu[21], to find a set of colors which are a good low resolution representation of the image (i.e. in a color sense). The image is transferred to $\boldsymbol{CIE\ L^*a^*b^*}$ taking into account only the color information, treating the image pixels as a set of data points in the color space. For the transformation to $\boldsymbol{CIE\ L^*a^*b^*}$, sRGB is assumed for all the images and the D65 illuminant is assumed as the viewing illuminant. The image representation is segmented and for each segment the representative mode added to the set of representative colors for the image.

The extracted colors for the images corresponding to a given term are added to a pool of representative colors. A second mean shift procedure is applied to extract the final representative color set for the term. For the extraction we again use the Epanechnikov kernel with a size that produces a small number of segments. The size of the kernel in the second step can be varied to produce representative color sets with different granularity. Algorithm 3.4 depicts the mean shift method.

Algorithm 3.4 (mean shift) method : $\boldsymbol{C}$ = extract_representative_colors(I, hs, hs^*)

Input

set of images I, kernel size hs, image kernel size hs^*

Output

set of representative colors for the set of images $\boldsymbol{C}$

$t_1 = \frac{2|i|}{m^3}$

for all $i \in I$ **do**

$(\boldsymbol{C_i}, r(\boldsymbol{c}))$ = mean_shift_cluster(i, hs^*)

for all $\boldsymbol{c} \in \boldsymbol{C_i}, r(\boldsymbol{c}) > t_1$ **do**

$\boldsymbol{C}^* = \boldsymbol{C}^* \cup \boldsymbol{c}$

$(\boldsymbol{C}^*, r(\boldsymbol{c}))$ = mean_shift_cluster$(\boldsymbol{C}^*, hs)$

for all $\boldsymbol{c} \in \boldsymbol{C}^*, r(\boldsymbol{c}) > 0.05 \max_{\boldsymbol{c} \in \boldsymbol{C}^*} r(\boldsymbol{c})$ **do**

$\boldsymbol{C} = \boldsymbol{C} \cup \boldsymbol{c}$

Color bias corrected mean shift

To add the correction for the color bias in the images, we extend the mean shift iteration given in Equation 3.6 to the form

$$\boldsymbol{x}^{\boldsymbol{k+1}} = \frac{\sum_{i=1}^{n} w_i \boldsymbol{x_i} g\left(\left\|\frac{\boldsymbol{x}^k - \boldsymbol{x_i}}{hs}\right\|\right)}{\sum_{i=1}^{n} w_i g\left(\left\|\frac{\boldsymbol{x}^k - \boldsymbol{x_i}}{hs}\right\|\right)}, \tag{3.8}$$

where w_i are weights computed for all the data points $\boldsymbol{x_i}$ from the estimated set distribution $\hat{h}$. The set distribution has values only for the bin centroids and in order to compute the value for an arbitrary point we use trilinear interpolation.

3.6 Experimental Results

Having no known prior work that deals with the problem addressed, to evaluate our method we used results from related topics. The first part of the evaluation compares the representative colors resulting from using color names as terms to acceptable hue ranges from [60]. The second part of the evaluation uses three sets of terms in English and their translations to Finnish and computes color differences between the resulting representative color sets.

Color names

To subjectively assess the performance of the method, we present the colors which were computed as most relevant for a number of color names in English and Finnish in Table 3.1 and Figure 3.5. The presented results use the histogram method with 16 bins per channel and images from both Google and Flickr with a set of around 200 images for each image search engine.

The colors with high expected chroma (blue, yellow, red, green, brown and orange) have also a high computed chroma value. Furthermore, the hue difference between the colors for the English and the Finnish terms is small. Visual inspection also shows that the results are acceptable as representatives of the colors which names are used in the method. The colors with lower expected chroma (black, white and gray) have a chroma value in all cases smaller than the chroma values for the first group. The difference between the colors for the color names in English and Finnish is also bigger in this case. Note that for very low chroma values, the hue value is undefined. For each of the colors in the low chroma group, at least one of the colors is not acceptable as a representative for the color with the name used in the method. It has to be noted that in each of the cases, one of the other colors in the representative set is acceptable as a representative color.

	Lightness	Chroma	Hue
blue	27	59	280
sininen	38	38	282
yellow	74	63	94
keltainen	83	73	83
red	35	55	31
punainen	41	68	33
black	16	0.	315
musta	68	0.	315
white	63	17	286
valkoinen	87	6	19
green	53	28	130
vihreä	45	39	132
brown	48	25	43
ruskea	65	23	67
gray	75	11	69
harmaa	74	0.	315
orange	68	76	59
oranssi	57	69	52

Table 3.1: Most representative colors for the color names in English and Finnish. The colors are given in the $\boldsymbol{CIE\ LCh_{a^*b^*}}$ color space.

To give a more objective assessment on the performance of the method and to give a basis for an automatic evaluation we use the results of Kuehni [60], which give the acceptable range in $\boldsymbol{CIE\ LCh_{a^*b^*}}$ hue for the colors blue, red, green and yellow. Table 3.2 gives an overview of the selected colors with the range of hue angles that correspond to each color. The hues for the most representative colors shown in Table 3.1 fall into the boundaries defined in Table 3.2. Furthermore, applying either the histogram method with 8, 16 and 32 bins per channel or the mean shift algorithm produced most representative colors with hues that fall into the boundaries given in Table 3.2. Taking only the images which were in the set of search results of one of the image search engines, either Google or Flickr, also produced results with hues inside the given boundaries.

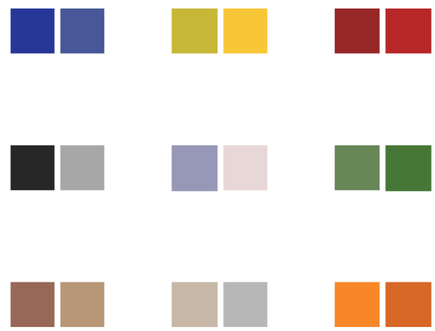

Figure 3.5: Most representative colors for nine color names in Table 3.1 in English (right patches) and Finnish (left patches).

	Minimum h	Focal h	Maximum h
red	350	24	41
yellow	75	87	101
green	112	175	222
blue	223	246	287

Table 3.2: Hue angles for color regions.

Relative comparison

The accuracy with which the representative colors are extracted is also evaluated with a relative comparison between terms in two languages. This is done on three sets of terms, being color names, names of real objects and colorless terms. Two terms with the same meaning are expected to be associated with the same color. We therefore select a set of terms in English and their translations into an unrelated other language, namely Finnish. The translation of the terms from English to Finnish was done by a native Finnish speaker and given the simplicity of the terms, a large variation between translations is not expected. Since the terms are different, the queries in English and Finnish result in different sets of images. However, when querying for the terms *tomato* or *tomaattii* we expect images of the same concept. We therefore evaluate our method by calculating the distance between the representative colors found with the English and Finnish terms.

The distance ΔE between two colors $\boldsymbol{c_1}$ and $\boldsymbol{c_2}$ is defined as

$$\Delta E(\boldsymbol{c_1}, \boldsymbol{c_2}) = \sqrt{(\Delta \boldsymbol{L})^2 + (\Delta \boldsymbol{a})^2 + (\Delta \boldsymbol{b})^2} \quad , \tag{3.9}$$

where $\Delta \boldsymbol{L}$, $\Delta \boldsymbol{a}$ and $\Delta \boldsymbol{b}$ are the differences between the $\boldsymbol{L}$, $\boldsymbol{a}$ and $\boldsymbol{b}$ components of the representation of $\boldsymbol{c_1}$ and $\boldsymbol{c_2}$ in the ***CIE*** $\boldsymbol{L^*a^*b^*}$ color space. The Euclidean distance ΔE_{ab} in ***CIE*** $\boldsymbol{L^*a^*b^*}$ is a standard measure for color differences because of the near perceptual uniformity of the space. All colors produced by the method are assumed to be in the sRGB standard device dependent color space and D65 is assumed as a standard illuminant.

The terms selected are split up into three categories. The first one contains names of colors, where we expect small values for ΔE. The second category has terms that have a clear association with colors, where we also expect small ΔE-values. We compare these two categories with the last one, where we selected terms without an expected association with a color (category 3). The results of the experiments can be found in Tables 3.3 and 3.4. The above discussed difference is given in column ΔE. As expected the difference between the colors produced for the set of color names is smaller than the difference in colors obtained for the real objects and for the colorless terms. We observed that in the case of the real objects, most terms had a connection to more than one color, and in most of the cases, the order in which the colors were ranked in the set was different. Thus taking into account only the first representative color in the resulting set produced large errors.

To take into account the differences between the sets, and not only the most representative color, we used a measure of set distance between the computed sets. Given two sets $\boldsymbol{C}$ and $\boldsymbol{D}$ with n and m colors and $r(\boldsymbol{x})$ denoting the computed contribution of color $\boldsymbol{x}$, we define the set distance ΔE^* as

$$\Delta E^* = \frac{\sum_{i=1}^{n} r(\boldsymbol{c_i}) \min_j \Delta E(\boldsymbol{c_i}, \boldsymbol{d_j}) + \sum_{j=1}^{m} r(\boldsymbol{d_j}) \min_i \Delta E(\boldsymbol{c_i}, \boldsymbol{d_j})}{\sum_{i=1}^{n} r(\boldsymbol{c_i}) + \sum_{j=1}^{m} r(\boldsymbol{d_j})}, \ \boldsymbol{c_i} \in \boldsymbol{C},\ \boldsymbol{d_j} \in \boldsymbol{D} \tag{3.10}$$

or as a weighted sum of the minimum distances between the colors in the two representative color sets, where the weights are taken to be the computed contributions of individual colors.

As can be seen in Table 3.3, using the set distance instead of the most representative color distance lowers the differences between colors for some of the terms considerably. Terms that had only one color in the set of representative colors, on the other hand, have the same values for both ΔE and ΔE^*. The most visible improvement is in the low chroma colors and in the real objects group (category 2).

Using the mean shift algorithm on the same test terms produced results which followed the same trends and had the same subjective quality of the produced representative color sets.

TERM	FINNISH	ΔE	ΔE^*	Divergence
blue	sininen	7.483	7.483	0.202
yellow	keltainen	18.104	18.104	0.185
red	punainen	15.006	15.006	0.149
black	musta	31.457	13.703	0.059
white	valkoinen	21.898	7.455	0.070
green	vihreä	13.238	13.238	0.181
brown	ruskea	19.346	19.346	0.067
gray	harmaa	31.152	11.086	0.098
orange	oranssi	16.100	16.100	0.171
tomato	tomaatti	11.740	11.740	0.216
ocean	valtameri	12.619	12.619	0.113
beer	olut	68.847	23.056	0.053
glass of coke	coca-cola	52.365	33.086	0.103
parking lot	parkkipaikka	26.560	19.341	0.136
concrete	betoni	41.490	28.541	0.111
bricks	tiilet	37.345	37.345	0.139
banana	banaani	34.530	23.745	0.095
cheese	juusto	23.749	23.749	0.125
lemon	sitruuna	7.245	7.245	0.213
strawberry	mansikka	14.592	14.592	0.185
ice	jää	30.729	18.273	0.075
polar bear	jääkarhu	25.468	22.295	0.175
wood	metsä	37.173	32.917	0.075
peace	rauha	46.528	25.441	0.046
chair	tuoli	16.212	16.212	0.076
w. i. c.	l. i. y.	25.701	25.701	0.026
book	kirja	29.056	22.624	0.057
city	kaupunki	18.502	18.502	0.056
family	perhe	23.749	23.749	0.053
paradise	paratiisi	15.777	15.777	0.077

Table 3.3: ΔE, ΔE^* and Divergence values for the terms used in the evaluation. The two abbreviations in the table are w.i.c. for "wireless internet connection" and l.i.y. for "langaton internet yhteys".

Color distribution divergence

The presented distance measures operate on the set of representative colors and depend on the algorithm and parameters selected. The presented algorithms also need a measure of confidence on the computed results, which can also serve as a measure of suitability of a term for color extraction. The target measure would only depend on the set of images retrieved and not on the algorithm applied to produce the representative

	ΔE	ΔE^*	Divergence
category 1	19.3093	13.5023	0.1312
category 2	30.3179	22.0387	0.1296
category 3	24.3144	23.2554	0.0582

Table 3.4: Average ΔE, ΔE^* and Divergence values for the three categories.

color set. We observed that a good candidate measure that has the required property is the KL Divergence between the distribution of colors for a term and the general distribution used for the bias correction. The KL Divergence is an information theoretical measure for dissimilarity between distributions.

Figure 3.6 depicts the values for the KL Divergence for the terms in the three categories selected in the experiments. Both Table 3.3 and Table 3.4 also show the divergence for separate terms and the average divergence per category. As expected, the average divergence of the color names and real objects is close and an independent sample t-test found the difference not significant at $\alpha = 0.05$ level, while both categories have a significantly higher divergence value than the one in the colorless terms category.

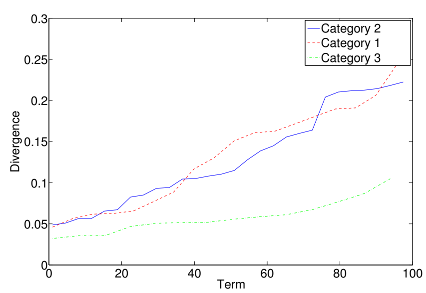

Figure 3.6: Divergence between the learned set distribution and the term distributions.

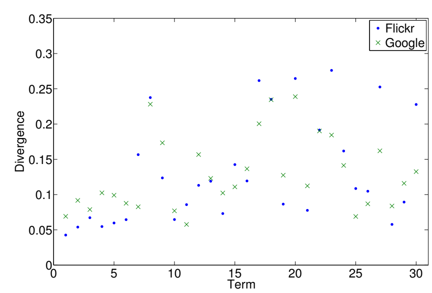

Figure 3.7: Divergence for images sets from different search engines.

Search engine influence

The results presented so far take images from both image search engines. It is interesting to look at possible differences between the engines, specially because they use two different kinds of indexing. We expect that Flickr, which has manual tagging would overall perform better. To present the differences we use the KL divergence. The expectation that Flickr would provide better results can be translated to an expectation of a higher divergence for the distribution of colors computed from the set of images acquired from Flickr.

The color distributions computed from the images acquired from Flickr had a mean divergence of 0.1769 with a variance of 0.1172. The Google based ones had a mean of 0.1315 and a variance of 0.0604. A paired t-test on the divergence values showed a significant effect at $\alpha = 0.05$.

On closer inspection of the data, we observed that the large variance in the Flickr data was due to a small number of terms (10 from the 60) which had a much higher divergence value than the others. The terms with this property were all in Finnish and the number of images acquired less than 50. On the other hand, all the Google image sets had 200 images, which was the limit used in the experiments. After the deviating terms were filtered, the divergence distribution for the Google data had a mean of 0.1286 with a variance of 0.0521. The new mean divergence for the Flickr data set was 0.1325 with a variance of 0.0753. While the Google mean divergence changed 2.2%, the Flickr one changed 25%. A repeated paired t-test for equality of means showed no significant effect of the search engine at $\alpha = 0.05$. The above results show that contrary to our prior belief, for the set of terms we used we did not find a significant effect of the search engine on the suitability of the produced image sets for representative color extraction. The number of images in the set, however, showed a significant effect. Figure 3.7 shows the divergence values for the color distributions computed from the images acquired from Google and Flickr search engines .

3.7 Conclusions

In this chapter, an unsupervised method to enrich textual applications with relevant images and colors is presented. We show that meaningful colors for a term can be computed using large image repositories and image processing. The most important problem in the representative color extraction identified was the one of correcting for the color bias present in the images. We show how the bias can be corrected and representative colors extracted using a histogram based approach or mean shift.

The method was evaluated using hue angle boundaries for high chroma colors and differences between sets of representative colors computed for terms in English and Finnish. By comparing the value of the KL Divergence for the three sets of terms selected, we show that it can be used as a measure of suitability of a term for color extraction.

A prototype system based on this method is presented where the method is applied to song lyrics. In order to identify terms within the text that may be associated with images and colors, we select noun phrases using a Part of Speech tagger. Large image repositories are queried with these terms and results gathered in a set of images per term. The images that are best ranked by the search engine are displayed on a screen, while the extracted representative colors are rendered on controllable lighting devices in the living room.

4

Specification and rendering of spatial distributions of light

"We did not need a special word for interactivity in the same way that we do not (yet) need a special word for people with only one head."

Douglas Adams

After the selection of a palette of colors for the desired lighting atmosphere, their distribution in the environment has to be defined. With a large number on independent light sources, their individual control becomes unfeasible. To overcome this, this chapter introduces the "virtual spot" concept that enables the creation of the spatial distribution based on a small number of atomic effects that user can understand. Using this abstraction together with codes embedded in the produced light, an interactive spatial light design algorithm is presented. Two prototypes of the proposed algorithm were built and tested for feasibility and user satisfaction.

This chapter is based on [87]

4.1 Introduction

Advances in solid state lighting (SSL) have enabled its use as a viable alternative to traditional light sources [85, 77]. The multitude of color tunable light sources that comprise a modern SSL system allows for finer control and more freedom in setting and adaptation of the color, and of spatial and temporal properties of the lighting system. This results in new possibilities for atmosphere creation and lighting effects.

The large freedom in settings, typical of such lighting systems, makes traditional controls ill-suited, even for professionals. The process of manipulating the parameters of each light source individually would be overwhelming to the user. The reason for this lies in the nature of traditional controls which are based on the *cause-effect* paradigm. For example, the user flips a wall switch to control the status of a light source (the *cause*), in order to create an illumination distribution in the room (the *effect*). This control paradigm will fall short when applied to extensive SSL systems, because of the large amount of parameters and consequently the substantial number of possible effects.

Various solutions for the control of light effects have been developed. Usually these solutions rely on a two step approach. The first step consists of the description of the complete target light effect. For example, the description can be given by means of a drawing that shows the distribution of the light colors over a scene. Alternatively, the method presented in Chapter 3 can be extended to include spatial information next to the extracted color information, resulting in a spatio-color map of an image or a term. The second step consists of the mapping of the target scene onto the real space. The mapping problem can be solved once the footprints of the available light sources are known.

The mapping is traditionally solved using computer simulations [25]. Often these simulators are based on either ray tracing or radiosity methods. In general, the accuracy of these simulators is only acceptable when the space, including furniture, can be accurately modeled. This can be very cumbersome for the light designer. Moreover, experiencing the designed light scene on a computer screen is very different from being in the actual scene.

Another approach is based on recording the various light source footprints with a camera [83]. In this approach all light sources have to be switched on and off sequentially. This time-consuming approach has the following shortcomings. First, the resulting light effect depends on the camera view point, hence, an accurate analysis on the best camera location is required. Furthermore, a new calibration is required every time something changes in the space, e.g. when the interior of the space is changed or when the location or orientation of luminaires are varied. Last, and most noteworthy, this approach does not allow direct interaction, since the creative step and the mapping step are decoupled.

Interactivity in the context of lighting design enables fast prototyping and iterations, both considered an essential part of an effective creative process [3]. Furthermore, it would be advantageous for the user to design the light effect directly in the space and not on paper. To this end, we present a novel solution that uses a radically different paradigm for lighting control. Our solution allows the user to control the effect directly on the spot where it should be and without worrying about the contribution required by every single light source. The system is based on the use of coded light [66, 67, 116] , where each lamp embeds invisible light source identifiers in its light output. In that way, by using a photo sensor, it is possible to discern the illumination contributions from all the light sources. A lighting system based on this technology allows users to create a light effect, in real time, by just choosing the desired effect (e.g. color, intensity, distribution) for the target location. The translation from desired effect to light source parameters is performed by an algorithm that uses the chromaticity properties and the intensity settings of all sources, as measured by the photo sensor.

This chapter is organized as follows. Section 4.2 gives a general overview of the system used in the rest of the chapter. Then, in Section 4.3, we present the coded light technology and we explain how it is used to estimate the contributions of the individual light sources at a given location. Subsequently, Section 4.4 describes the algorithms that allow light effect rendering using these estimates. A system setup that we used for testing of the presented concept and application development is then presented in Section 4.5. Finally, conclusions are drawn in Section 4.6.

4.2 System overview

Figure 4.1 schematically depicts a lighting system enabling the proposed lighting design method. The system consists of a large number (N) of SSL sources installed in the ceiling, all embedding a unique identifier in their light output. The system additionally consists of a remote control (RC) and a system controller (SC). The RC is operated by the user and it is able, using an optical sensor, to receive the coded light identifiers in the light. The user input and the received identifiers are sent to the SC, which controls the SSL light sources accordingly. During the interaction, the user can store the information associated with a location, which we will call a target or a target location.

4.3 Coded Light

The introduction explained that for the interactive control of a complex SSL lighting system it is of importance to estimate the localized illumination contribution of each light source. In the system proposed here, this estimation is achieved using a tech-

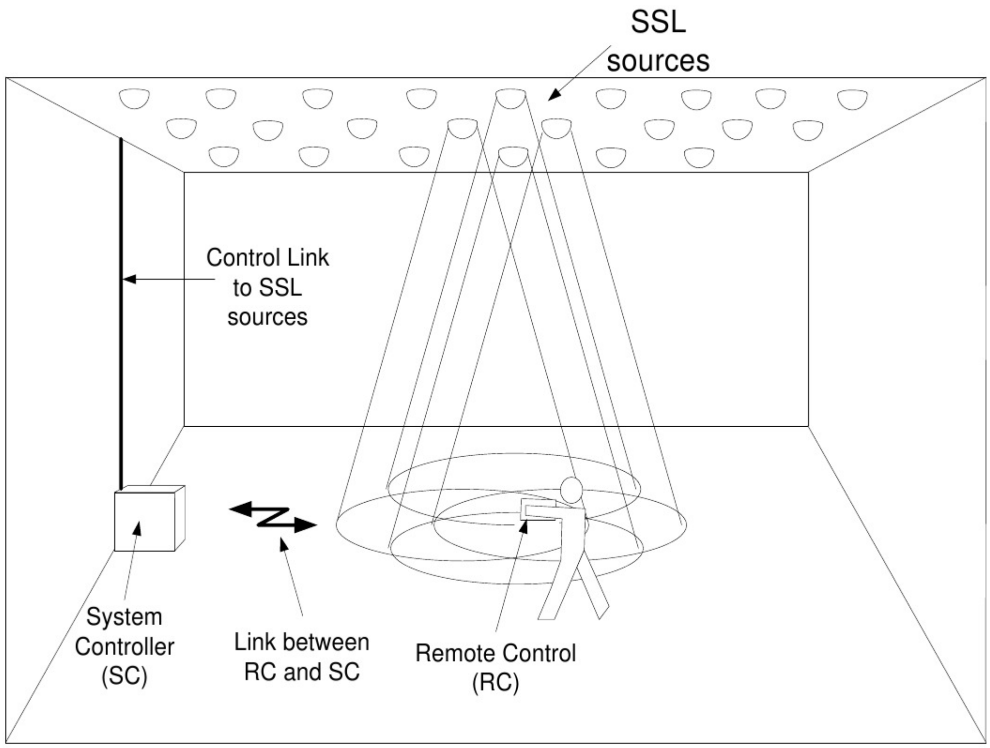

Figure 4.1: System overview. The system comprises SSL sources emitting coded light, a remote control for user control, and a system controller that sends control commands to the SSL sources.

nology called *coded light*, as earlier presented in [66, 67, 116]. Using this technology every light source can be identified in the system using a unique identifier embedded in the light. This identifier should be invisible to the human eye, but can be detected using the optical sensor device in the remote control.

4.3.1 Application requirements on light identifiers

To be able to use coded light in general lighting applications that enable immediate response on interaction, a few application requirements have to be fulfilled. The most important ones are:

1. Independent identifiers and illumination: The main function of the light emitting diodes (LEDs) in the lighting system is providing illumination, thus the embedding of identifiers should not affect the short-term average light output of the light sources. Also, the identifier embedding techniques should preferably be compatible with the typically applied pulse-width modulation (PWM) dimming

of LED light output to achieve efficient driving of the light sources.

2. Imperceivable: The modulation of the LED light, to embed the identifiers, should not create visible flickering, otherwise it would disturb the users of the lighting system. The invisibility can be achieved by minimizing the energy in low frequency components (approximately below 100 Hz) in the light source identifiers.

3. Number of LEDs: The system must be able to measure the contribution from each locally relevant light source individually and simultaneously. It should be able to operate in an environment with several hundreds of LEDs.

4. Short response time: The modulation method should allow fast light source identification and illumination estimation. This guarantees that the user experiences an immediate reaction after pressing a control button. Hence, a sensor should be able to identify and measure all relevant light sources within several tenths of a second.

4.3.2 Coded Light techniques

Different techniques meeting these requirements were previously presented in [66, 67, 116]. These techniques are compatible with PWM dimming of SSL light sources, and do not impact the illumination function. The proposed coded light techniques are based on coded division multiple access (CDMA) [66, 67] and frequency division multiple access (FDMA) [116].

The CDMA method is illustrated in Figure 4.2 with the solid line. Basically the bits of the light source identifier are embedded by slightly varying the length of the PWM pulse, a longer pulse identifying a 1 and a shorter pulse a 0. Figure 4.2 shows the resulting light output of the nth LED embedding the identifier code $c = [1\ 0\ 0\ 1]$. The dashed line identifies the light output for the normal PWM modulated signal, i.e. without embedding an identifier. The average duty cycle of the light is $d_n = 50\%$ for both signals in this example. From this we can conclude that the average illumination level is not changed due to the embedding of the identifier. The length of one PWM interval is T seconds, and the length of the code equals M bits. The whole code is acquired in MT seconds. To meet requirement 4, MT should be around 0.1 s to guarantee perceived instantaneous feedback [78]. The peak illumination level of the LED is denoted by a_n, while the actual illumination level equals $a_n d_n$ due to the PWM dimming. In the case of FDMA, every LED is assigned a unique repetition frequency $f_n = 1/T_n$ of the PWM pulses. For both the CDMA and FDMA approach the light modulation will be invisible when f is above the critical flicker frequency of the human visual system. Although the critical flicker frequency depends on many factors of among which the light level, the size of the flickering stimulus and the position of the stimulus in the visual field, 100 Hz can be taken as a safe lower limit [114] . It has to be noted that this limit only insures that the system does not produce directly visible flicker and does not take into account interactions with moving objects.

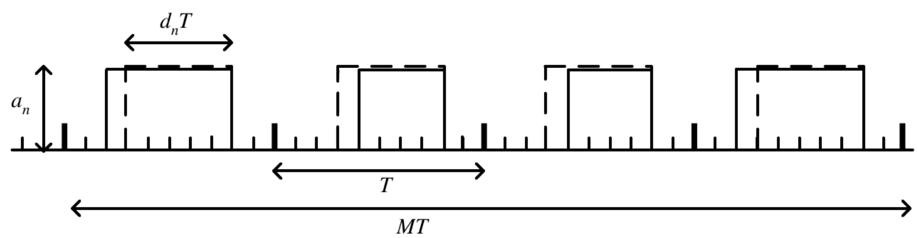

Figure 4.2: Embedding of the identifier code $c = [1\ 0\ 0\ 1]$ in the 50% PWM dimmed light output of the nth LED light source.

The aggregated illumination contribution of all N LED sources at one of the R target locations denoted as i can be written as

$$I_i = \sum_{n=1}^{N} a_n d_n \eta_{n,i} \tag{4.1}$$

where $\eta_{n,i}$ denotes the attenuation value for the light propagation from the nth light source to the ith target location.

In the remote control at a location i, the peak illumination contribution for each light source n $\hat{b}_{n,i} = a_n \eta_{n,i}$ is estimated individually. To this end the remote control is equipped with an optical sensor, e.g. a photo detector, which converts the optical signal into an electrical signal. Then by applying digital signal processing to the received coded light signal, the estimate $\hat{b}_{n,i}$ for all N LEDs is computed. These estimates are then consequently used as input to the interactive light design algorithm.

4.4 Interactive Light Design

The estimate of the individual illumination contributions $\{\hat{b}_{n,i} \mid 1 \leq i \leq R,\ 1 \leq n \leq N\}$ can be used to control the amount of light that the system produces at a point in space to satisfy a user requirement on the light level at that point. This is achieved by controlling the duty cycles $\{d_n \mid n \leq N\}$ of the LEDs.

We present the user-system interaction and the algorithms used in an order of increasingly complex user requirements. First, we show how a user can control the light intensity in a single point in space; second, we show how the chromaticity of the light at the point in space can be controlled; and finally we show an algorithm that controls the LEDs to satisfy multiple user requirements. In the rest of the section, we assume that all LEDs have the same peak illumination level, i.e. $a_n = a$. Under this assumption, the estimates of the light intensity contribution of a single LED n to a location i, $\hat{b}_{n,i}$, is inversely proportional to the distance between the LED and the location.

4.4.1 Single intensity requirement

The first problem that we solve using the estimates given above is the computation of the duty cycles required for the generation of specific light intensity in a point. In the case the required intensity at the target point is the maximum light intensity at that point, the solution to this problem is unique and trivial. However, if only a part of the total available illumination power is needed, the problem is under-constrained, as many combinations of LED duty cycle values can produce the desired light intensity.

To add a natural constraint to the distribution of the light we add a new parameter γ to the system that controls the distribution of the light intensity for the points around the target point. Using the estimates of the light intensity at the target point, $\{\hat{b}_n \mid 1 \leq n \leq N\}$, the required light intensity u and the spatial distribution parameter γ, the duty cycle for LED n is computed as

$$d_n = u \frac{(\hat{b}_n)^{\gamma-1}}{\sum\limits_{m=1}^{N} (\hat{b}_m)^{\gamma}} \tag{4.2}$$

It is easy to see, by substituting the computed duty cycles from Equation 4.2 into Equation 4.1, that when the required intensity is smaller than the maximum achievable intensity at that point, the aggregated illumination contribution I, will be equal to the required one u.

The parameter γ takes values from 1 up to ∞, producing a spot of light around the target point with a variable spatial extent. For $\gamma = 1$, the system produces a uniform distribution, corresponding to the maximum spot size, i.e. all LEDs with a non-zero illumination contribution at the target position are assigned the same duty cycle. When $\gamma \to \infty$, only the LED with the highest contribution (the optically closest LED) has a non-zero duty cycle, while all the others have a zero duty cycle. This results in the smallest spot size that can be produced by the system. Figure 4.3 shows the relative duty cycle of the LEDs as a function of the relative individual illumination contribution for different values of γ. In the figure, the relative individual illumination contribution is computed by dividing the individual illumination contribution by the maximal one, i.e. $\hat{b}_n/max\{\hat{b}_m \mid 1 \leq m \leq N\}$. Similarly, the duty cycle is normalized by the maximal one to compute the relative duty cycle, i.e. $d_n/max\{d_m \mid 1 \leq m \leq N\}$.

Due to the simplicity of the algorithm the latency of the system mainly depends on the communication latency. Additionally, because of the short message length needed to

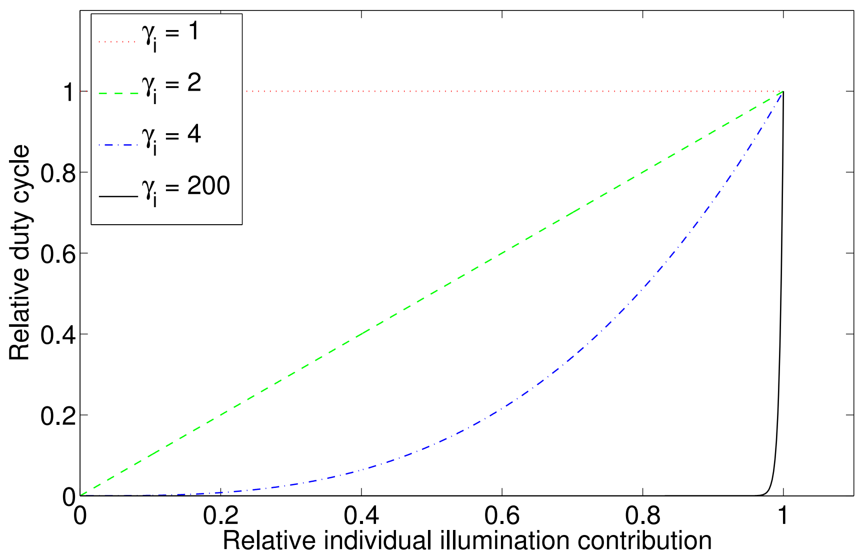

Figure 4.3: Relative duty cycles for LEDs as a function of the relative individual illumination contribution for different values of γ.

transfer the information needed in the algorithm, the parameters of the coded light can be tuned to enable a low latency interaction with the system that fulfills requirement 4.

4.4.2 Color requirements

In the case of a system consisting of LEDs having different chromaticities, their mixing can produce different colors at different positions. The different parts of the system, consisting of the sets of LEDs with the same chromaticities are referred to as the primary systems. As most additive color controllable systems consist of three primary systems with red, green and blue chromaticities, we will assume a system with three primaries in the rest of the section. We denote the chromaticities of the three primaries in the 1931 ***CIE xyY*** color space by $(x,y)_r$, $(x,y)_g$, and $(x,y)_b$, and their luminances at the target point as $\boldsymbol{Y_r}$, $\boldsymbol{Y_g}$, $\boldsymbol{Y_b}$. Given the ***CIE xyY*** coordinates, the 1931 ***CIE XYZ*** coordinates of the light incident at the target point from the red, green and blue system can be trivially computed and are denoted by $(\boldsymbol{X},\boldsymbol{Y},\boldsymbol{Z})_r$, $(\boldsymbol{X},\boldsymbol{Y},\boldsymbol{Z})_g$, and $(\boldsymbol{X},\boldsymbol{Y},\boldsymbol{Z})_b$.

Due to the additive nature of light and the trichromacy of human color vision, the problem can be subdivided in two parts. The first one is finding the combination of the intensity of the primaries that will produce the desired chromaticities and the second one is computing the duty cycles that produce the desired intensity per primary system.

Given the required chromaticity (x,y) and the required luminance $\boldsymbol{Y}$ for the target point, or $(\boldsymbol{X},\boldsymbol{Y},\boldsymbol{Z})$ in ***CIE XYZ*** color space, the required luminances for the primary systems u_r, u_g, and u_b are computed using a standard [114] transformation

$$\begin{bmatrix} u_r \\ u_g \\ u_b \end{bmatrix} = \begin{bmatrix} \boldsymbol{X_r} & \boldsymbol{X_g} & \boldsymbol{X_b} \\ \boldsymbol{Y_r} & \boldsymbol{Y_g} & \boldsymbol{Y_b} \\ \boldsymbol{Z_r} & \boldsymbol{Z_g} & \boldsymbol{Z_b} \end{bmatrix}^{-1} \begin{bmatrix} \boldsymbol{X} \\ \boldsymbol{Y} \\ \boldsymbol{Z} \end{bmatrix} \tag{4.3}$$

In the second step, using the required luminances for the primary systems u_r, u_g, and u_b and Equation 4.1, the duty cycles for the LEDs in each primary system are computed separately using Equation 4.2. The computation of the duty cycles for a single color requirement is given in Algorithm 4.1.

In the case of a number of primaries larger than three, only the first part of the algorithm has to be changed. In this case there is no unique combination of primary luminances that produce the required chromaticity at the required luminance. Hence, additional constraints or optimization criteria have to be used. One example criterion is the minimization of the power used to produce the required settings. Another is the maximization of the color rendering quality provided by the illumination system.

Algorithm 4.1 Method : $\{d_{\boldsymbol{c},n}\} = compute\_color\_requirement(\{\hat{b}_{\boldsymbol{c},n}\}, (x,y), \boldsymbol{Y}, \gamma)$

Input

estimates of the light intensity from LED n and primary $\boldsymbol{c}$, $\{\hat{b}_{\boldsymbol{c},n}\} = \{\hat{b}_{\boldsymbol{c},n} \mid 1 \leq n \leq N, \boldsymbol{c} \in \{\boldsymbol{r},\boldsymbol{g},\boldsymbol{b}\}\}$, target chromaticity (x,y), target luminance $\boldsymbol{Y}$, spatial parameter γ.

Output

duty cycles $\{d_{\boldsymbol{c},n}\} = \{d_{c,n} \mid 1 \leq n \leq N, \boldsymbol{c} \in \{\boldsymbol{r},\boldsymbol{g},\boldsymbol{b}\}\}$, producing the target luminance and chromaticity

---

$[\boldsymbol{X}\ \boldsymbol{Y}\ \boldsymbol{Z}]' = xyY\_to\_XYZ([\boldsymbol{x}\ \boldsymbol{y}\ \boldsymbol{Y}]')$

$$[u_{\boldsymbol{r}}\ u_{\boldsymbol{g}}\ u_{\boldsymbol{b}}\ ]' = \begin{bmatrix} \boldsymbol{X_r} & \boldsymbol{X_g} & \boldsymbol{X_b} \\ \boldsymbol{Y_r} & \boldsymbol{Y_b} & \boldsymbol{Y_b} \\ \boldsymbol{Z_r} & \boldsymbol{Z_g} & \boldsymbol{Z_b} \end{bmatrix}^{-1} [\boldsymbol{X}\ \boldsymbol{Y}\ \boldsymbol{Z}]'$$

for all $\boldsymbol{c} \in \{\boldsymbol{r},\boldsymbol{g},\boldsymbol{b}\}$ **do**

  for all $n \in \{1,\ldots,N\}$ **do**

  $d_{\boldsymbol{c},n} = u\frac{(\hat{b}_{\boldsymbol{c},n})^{\gamma-1}}{\sum_{m=1}^{N}(\hat{b}_{\boldsymbol{c},m})^{\gamma}}$

Figure 4.4 shows an example user interface that can be used to select the desired chromaticity, luminance and spatial distribution at the target point in space, determined by the position of the remote control. The desired chromaticity is selected on the chromaticity diagram. The luminance and spatial distribution are selected using up-down dials.

The system can be in active mode, where the target position follows the position of the remote, or in a passive mode, where the last active position of the remote is taken as the target position. Using the button ”Freeze” the user can change between the active and passive mode. The user interface can for instance be run on a PDA or smartphone that has the remote control sensor integrated.

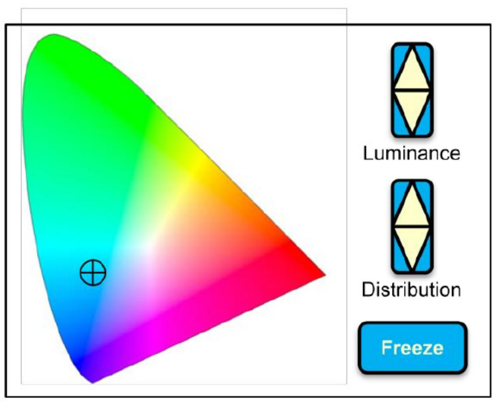

Figure 4.4: An example remote control user interface to select the chromaticity, luminance and spatial distribution at a target point.

4.4.3 Multiple requirements

In a practical system, the chromaticity, luminance and spatial distribution at several target locations are specified. Not all of these requirements can be satisfied simultaneously, so the system has to weigh the different requirements to come to a rendering of the scene. As a simple example, we can consider a three requirement case: two red spots at opposite sides of a room and a uniform white requirement over the whole room.

Given R target locations and a user requirement at the locations for a three primary system, the duty cycles $d_{n,i}$ that satisfy the individual requirements can be computed. A simple possible solution for the multiple requirement problem would be to add the individual requirement solutions for every LED, resulting in a high chance of requiring more intensity than is possible from an LED. Depending on the way this is handled, different artifacts can occur. In our example, if the solutions are simply added, the spots will disappear, as the white requirement dictates all the channels of all the LEDs to be on. If, however, the solution is normalized in a way that insures the maximum duty cycle is feasible, the result will be desaturated spots in a dim uniform field. Alternatively, the single solutions can be averaged, but this introduces even more problems as the zero duty cycles outside of the spot are also averaged.

A better solution would provide control between local and global accuracy. Similar to the spatial parameter in the single requirement case, a local priority parameter ρ_i is introduced. It takes values from 0 up to ∞ and controls the spatial weighting of the individual duty cycles $d_{n,i}$ as function of the distance between LED n and position i. The values of the weights for different distances and different importance parameter values ρ_i are similar to the ones shown in Figure 4.3. When $\rho_i \to \infty$, the value of the duty cycle of the nth LED, d_n, approaches the value of $d_{n,r}$, where r is the target point closest to the LED. So, for large values of ρ_i, requirements have mainly a local influence. When $\rho_i = 0$, $d_{n,i}$ will have the same influence independent of the distance between LED n and position i, i.e. the local and the global influence of the requirement is the same.

If we treat the solutions for the single requirements as points in N dimensional space and compute a weight for each point based on the local priority parameter ρ_i, we can define a sum of squares of weighted distances from any point in the space to the set of requirement solutions. A solution that minimizes this distance can be computed using least squares. For a suitable choice of a family of weighting functions (power functions), a closed form solution to the least square can be found. Equating the partial derivatives of the sum of squares to zero results in the solution given by

$$d_n = \sum_{i=1}^{R} d_{n,i} \frac{(\hat{b}_{n,i})^{\rho_i}}{\sum_{m=1}^{N} (\hat{b}_{m,i})^{\rho_i}} \tag{4.4}$$

The computation of the duty cycles for multiple requirements is given in Algorithm 4.2.

Figure 4.5 shows an example user interface for the selection of multiple requirements. Additional to the controls for the single requirement selection, given in Figure 4.4, the multiple requirements user interface includes a list of saved requirements and a control over the local importance parameter. Additionally, two buttons that control the addition and the removal of requirements from the list are added.

The list initially has one entry, denoted by the label “1”. The user initiates the interactive design by pressing the “Freeze” button, which initiates the active mode. After selecting the color and distribution parameters, the user can save the requirement using the “Save” button. This adds a new entry in the requirement list, and selects the new requirement as the active one.

Previously saved requirements can be changed by clicking on their label in the require-

Algorithm 4.2 Method : $\{d_{c,n}\}$ = $compute\_multiple\_requirements(\{\hat{b}_{c,n,i}\}, \{(x,y)_i\}, \{Y_i\}, \{\gamma_i\})$

Input

estimates of the light intensity from LED n and primary c to position i, $\{\hat{b}_{c,n,i}\} = \{\hat{b}_{c,n,i} \mid n \leq N,\ i \leq R,\ c \in \{r,g,b\}\}$,
target chromaticities $\{(x,y)_i\} = \{(x,y)_i \mid i \leq R\}$, target luminances $\{Y_i\} = \{Y_i \mid i \leq R\}$, spatial parameters $\{\gamma_i\} = \{\gamma_i \mid i \leq R\}$, local priority parameters $\{\rho_i\} = \{\rho_i \mid i \leq R\}$.

Output

duty cycles $\{d_{c,n}\} = \{d_{c,n} \mid n \leq N,\ c \in \{r,g,b\}\}$, producing luminance and chromaticity that is a weighted average of the requirements

---

for all $i^* \in [1 \ldots R]$ **do**

$\{d_{c,n,i^*}\} = compute\_color\_requirement(\{\hat{b}_{c,n,i^*}\}, (x,y)_{i^*}, Y_{i^*}, \gamma_{i^*})$

for all $c \in \{r,g,b\}$ **do**

$$d_{n,c^*} = \sum_{i=1}^{R} d_{c^*,n,i} \frac{(\hat{b}_{c^*,n,i})^{\rho_i}}{\sum_{m=1}^{N} (\hat{b}_{c^*,m,i})^{\rho_i}}$$

---

ments list, which activates the selected requirement. When activated, the previously saved requirements are in passive mode and have the position at which they were saved as a target position. Pressing the “Freeze” button enters active mode, where the target position for the selected requirement follows the position of the remote. The user can remove requirements by selecting one from the list and pressing the “Delete” button.

The example user interface shows one possible interaction with the system, which allows the control of chromaticity, luminance and spatial distribution for a number of points in space. This allows the user the same freedom of design as having a number of light spots with a variable color and beam angle. The advantage of the presented system is that the design is done without any mechanical movement or reconfiguration of the system. Hence, the requirements from a user point of view can be seen as a set of virtual lights, giving the definition of the desired effect. The system consequently translates these requirements into a set of controls for the light sources present in the environment. The user is never exposed to the complexity of the system, the number of light sources and their capabilities.

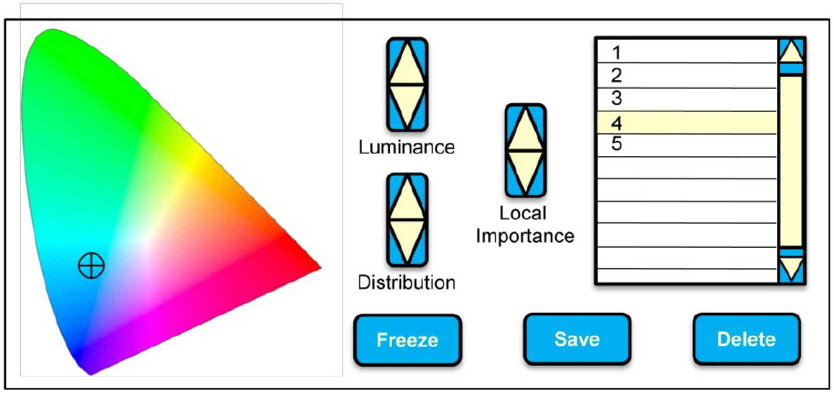

Figure 4.5: An example remote control user interface to select multiple requirements.

4.5 Prototypes and Usability Study

Physical setup

A prototype setup to test the feasibility of interactive light effect control based on coded light was realized. The system consisted of eight LED-based light sources installed in the ceiling, a RC and a SC. Each light source had three primaries, i.e., red, green and blue, which were independently controllable and assigned a unique identification code. Hence, 24 unique identifiers were embedded in the light, for which the CDMA technique was applied. The RC was implemented in a standalone unit equipped with a sensor that was able to estimate the illumination contributions corresponding to the different LEDs. The SC was implemented in a laptop. The interfaces between the RC and the SC, as well as between the SC and the SSL sources, were implemented as wired serial links. For simplicity of implementation, the user interface was not implemented in the remote control but in the laptop.

The algorithms for interactive light effect control described in Section 4.4 were implemented in the laptop in a combined LabView-Matlab software environment. The performed tests proved the robustness of these algorithms in effectively rendering up to eight light effects in eight distinct locations. Figure 4.6 captures an example of one of the tests. In the left diagram of Figure 4.6, the light effect editing is shown. During this phase, the remote control is placed on the target location and the desired color is selected via the user interface. In the right diagram of Figure 4.6, the result of the editing and the following rendering is shown.

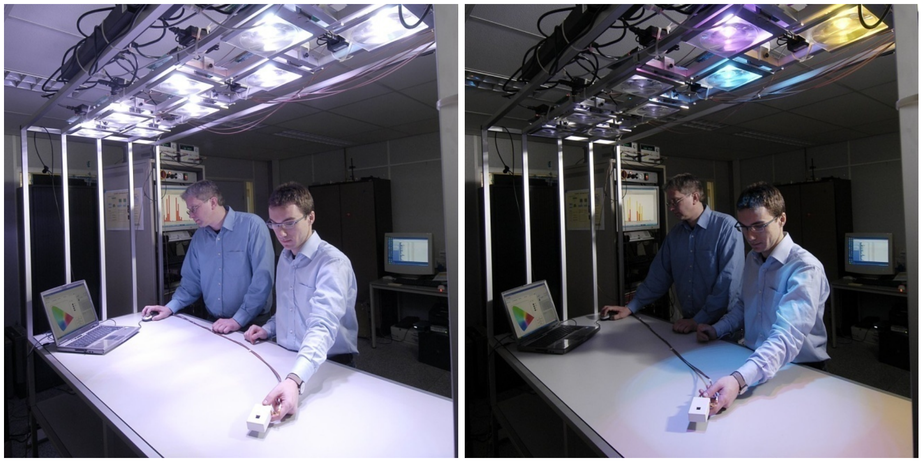

Figure 4.6: Interactive light effect control application as tested in the experimental setup. In the left diagram, the target location and color are selected. The right diagram shows the resulting light effect.

Simulated setup

Due to the limited hardware availability, a physical setup with a larger number of LEDs was not feasible. To test the usability of the interaction paradigm, a simulated environment was created. A single camera scene of an office space was created based on a 3d model with 63 equal light sources positioned on a grid covering the ceiling of the room. The UI from section 4.4.3 and the computed light effect were presented on a computer monitor. The position of the remote control in the scene could be selected using a computer mouse resulting in a real time change of the computed light effect.

Ten participants, experienced in lighting design and UI design, took part in a usability study using the simulated setup. After a short introduction to the study, they were given a set of three simple lighting design tasks :

- Create a spot light on the table.
- Create a warm spot light on the table and a cold uniform light in the rest of the room.
- Create the Italian flag on the wall of the room.

The interaction paradigm and the algorithm was not explained at the start of the experiment. After the participants fulfilled the first task, they were given a description of the interaction paradigm and the used interface. Nine out of the ten participants fulfilled the first task without any knowledge of the interaction paradigm or help from the experimenter.

After the participants fulfilled all three tasks, the USE usability questionnaire of Lund [4] was administered. It consists of 30 questions measured on a 7 point scale. The questions are divided in four groups: *usefulness*, *ease of use*, *ease of learning*, and *overall satisfaction*. The average score was 5.6. The highest scoring category was *ease of learning* with a score of 5.975. Similar scores had the categories *overall satisfaction* (5.728) and *ease of use* (5.63). The lowest scoring category was *usefulness* with a score of 5.1.

The high usability score was backed up by the comments from the subsequent interview. The interaction paradigm, its intuitive nature and ease of use were mentioned as a positive point by nine out of the ten participants. Most negative comments were concerning the UI. Issues with the labels used and the color picker were mentioned. Two proposals for improvement of the basic interaction paradigm were to add more types of light effect primitives and to add distinction between functional and decorative lighting.

4.6 Conclusions

In this chapter, an interactive lighting design and control approach based on coded light for large LED-based lighting systems was introduced. The coded light technology enables the online estimation of the illumination contribution of individual LED sources at a given point in space, using invisible light source identifiers. Algorithms were presented to allow light rendering for single and multiple intensity and chromaticity requirements at target positions with such a system. The algorithms were implemented in two test systems. Both the feasibility and the usability of the proposed system and algorithms were demonstrated using the test systems.

5

Creation and rendering of spatio-temporal distributions of color

"We are glorious accidents of an unpredictable process with no drive to complexity, not the expected results of evolutionary principles that yearn to produce a creature capable of understanding the mode of its own necessary construction."

Stephen Jay Gould

The previous two chapters demonstrated the selection of a color palette and an interactive method to define the spatial distribution of those colors. The produced light effect is however static. Contrary to this, most natural light effects are dynamic. Some, like the change of chromaticity of the sunlight and skylight, are dynamic at very long time intervals, and some, like the lightning strikes, at very short. Even more, most natural light effects are unpredictable from the point of view of a naive observer. In this chapter, a model that can be used for the creation of stochastic dynamic light effects is presented. A method to learn such a model from a video of an example natural light effect is presented. The learning method was tested in a large scale user experiment. Finally, possible extensions to spatio-temporal models and interactive systems are discussed. This chapter is based on [86].

5.1 Introduction

Recent advances in lighting introduced a revolution in the capabilities of lighting systems and elevate light to a status of a new medium. The main drivers of the advance, solid state light sources, enable lighting environments with a high spatial resolution, with fully controllable color and a wide range of dynamic capabilities.

One of the characteristics of modern lighting systems, which is a large differentiator, are the dynamic capabilities of the light sources used. The normal operating mode of most general use traditional light sources, disco lights being one of the few exceptions, is static and they only have an on/off functionality or a limited dimming range. This is contrary to lighting conditions in nature, which are inherently dynamic, from the slow change of the intensity and the color temperature of daylight during one day, to the fast flashes of lightning in a thunderstorm. Furthermore, most of the light effects we experience in nature are unpredictable on a certain timescale.

The new capabilities also introduce new challenges as the number of controllable parameters is much higher than for traditional lighting systems. As a result, the standard control paradigms used in traditional lighting systems become ineffective and a novel representation of the problem is needed to tackle the complexities.

One of the ways to simplify the control of modern lighting systems is to use a different medium, for example text, images, or video, as a representation of the desired ambiance and to use translation algorithms to translate the ambiance to the available light sources. This enables the users to control the lighting system by simple examples. Due to the fact that the user directly selects the final effect, while a set of algorithms compute the control values of the light sources to realize the effect, this control paradigm is an example of the so called effect driven control.

In this chapter, we propose a new method for generation of light effects which are locally unpredictable and non repeating, but resemble a natural light effect. The generation is done by simulating the execution of a stochastic process. The use of a simple stochastic process, a first order Markov chain, simplifies the creation of the stochastic models without a significant sacrifice on the expressive power. Furthermore, we present an unsupervised learning algorithm that produces a model based on a video of a natural light effect. To measure the recognizability and desirability of the produced light effects, a large scale user test was carried out, the results of which are presented. In the end, a generalization of the method to a spatio-temporal one is presented.

5.1.1 Stochastic models for media generation

A model of a *discrete time stochastic (random) process* X, $X = \{X_1, X_2, ..., X_N\}$, is a set of rules that characterize the joint probability distribution between all its random variables. The simplest model is the one where the random variables $X_1, X_2, ..., X_N$

are independent and identically-distributed (i.i.d.). For many natural processes, however, the assumptions, specially the assumption of independence of the random variables, are unrealistic. An often used generalization from the i.i.d. processes assumes that the distribution of a random variable X_N depends only on a subset of variables $\{X_{N-k},...,X_{N-1}\}$. Stochastic processes with this property fall in the class of *Markov processes*[38].

The probabilistic behavior of a Markov process is determined only by the dependencies between a subset of successive random variables. In the case of a first order Markov process the behavior is determined by the dependencies of immediate successors - between X_1 and X_2 , between X_2 and X_3 , etc. Despite their apparent simplicity and restrictions, Markov chains are rich in behavior, amenable to analysis, and adaptable to many applications, from weather to baseball prediction. They are centrally important to applied and theoretical probability. In the context of this work, we concentrate only on Markov models with discrete times and a finite number of states, or *finite state space Markov chains*.

Notably, one of the uses of Markov chains close to the one presented in this paper is for generation of media. The application that motivated the development of Markov models, text generation based on rules, has been used in a wide array of artificial intelligence applications. The rules in the model are in the form of conditional probabilities on the succession of words, learned from a large corpus of example content. Testing the performance of text processing algorithms often uses text content generated using Markov chain Monte Carlo [38]. An interesting and humorous example of the limited recognizability of text produced by a random process related to Markov chains is the random scientific paper generator [98], which produced a "scientific paper" that was accepted on a conference with a questionable review process. Markov chains have been used for modeling music [35] and for automatic generation of music [46].

In Markov chain applications, it is useful to think of the index t in X_t as a time index. X_t represents the state of the Markov chain model at "time" t. Formally, a Markov chain is a sequence $X = \{X_t\}_{t\geq 1} = \{X_1, X_2, ...\}$ of random variables taking values in a discrete set $\mathcal{E}$, and a matrix of conditional probabilities $\boldsymbol{P} = \{p_{ij}\}$, called the transition probability matrix. The elements of $\mathcal{E}$ are called states. The elements of the transition probability matrix p_{ij} give the probability that the process will be in a state j at "time" $t+1$, given that it was in state i at time t, $p_{ij} = P\{X_{t+1} = j \mid X_t = i\}$.

5.2 Light effect generation using stochastic models

A dynamic light effect is a sequence of colors at a set of discrete times. To use a finite state Markov chain as a model of a process, the process has to be represented as a time varying sequence of a finite number of unique states. In the case of light effect creation, there is a simple natural representation, mapping the states to single colors.

As an example, the set of states representing a night sky with lightning would have two colors, a dark blue one and a warm white one. The transition probability matrix this example is also simple. The dark blue color state would have a high probability of not changing state and a very small probability of changing state to the warm white state, while the warm white state would have a very small probability of not changing the state and a large probability of transition to the dark blue state. Figure 5.1 depicts the Markov model for this example. In more complex examples the states can be probability distributions instead of single colors. In the above example of a stormy night, the warm white state can be substituted with a three dimensional Gaussian distribution in a color space centered around a warm white color.

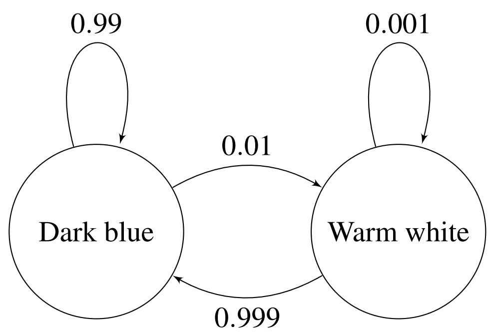

Figure 5.1: A simple light effect Markov Chain

To build more complex stochastic models manually would be impractical and another way of building the model is needed. In the next section an algorithm that creates a model based on a video of a natural light effect is given.

5.2.1 Learning

Given a video v of a target light effect, the following method for learning of a stochastic model of the light effect is proposed. The learning of the model constitutes of three main steps.

The first step is to extract representative colors from each video frame. This is done by computing a central tendency estimator, for example the sample mean or the sample color median, of the collection of colors present in the video frame or a region in the frame. The first step transforms the video into a discrete time sequence of colors $\{\boldsymbol{m}_t\}$ representing the video frames v_t.

The second step is the clustering of the representative colors from all the frames in the video $\{\boldsymbol{m}_t\}$ into a small number of classes, whose centroids $\{\boldsymbol{c}_i\}$ represent the states in the stochastic model. The clustering is done by using the blurring mean shift

Algorithm 5.1 Method : $(\{\boldsymbol{c}_i\}, \boldsymbol{P}) = \text{learn\_model}(v, hs)$

Input

example video v, kernel size hs

Output

color collection $\{\boldsymbol{c}_i\}$ of states, state transition probability matrix $\boldsymbol{P} = p_{ij}$

for all $t \in \{1, \ldots, |v|\}$ **do**

$\boldsymbol{m}_t = \text{extract\_color}(v_t)$

$(\{\boldsymbol{c}_i\}, \{n_t\}) = \text{mean\_shift\_cluster}(\{\boldsymbol{m}_t\}, hs)$

for all $i \in \{1, \ldots, |\{\boldsymbol{c}_i\}|\}$, $j \in \{1, \ldots, |\{\boldsymbol{c}_i\}|\}$ **do**

$\boldsymbol{P}^*(i, j) = 0$

for all $t \in \{1, \ldots, |v| - 1\}$ **do**

$\boldsymbol{P}^*(n(t), n(t+1)) = \boldsymbol{P}^*(n(t), n(t+1)) + 1$

for all $i \in \{1, \ldots, |\{\boldsymbol{c}_i\}|\}$, $j \in \{1, \ldots, |\{\boldsymbol{c}_i\}|\}$ **do**

$p_{ij} = \frac{p^*_{ij}}{\sum_{k=1}^{|\{c_i\}|} p^*_{ik}}$

algorithm [37, 21] on the representative colors in a nearly perceptually uniform color space (***CIE L*******a*******b****). The blurring mean shift is a density based algorithm that takes into account the local structure of the distribution in the clustering. The advantage of the mean shift algorithm over standard clustering algorithms like k-means is that the algorithm accepts the size of the kernel *hs* as an input parameter, instead of the number of final clusters. The size of the kernel *hs* can be set using a meaningful criterion such as the minimum distance between two cluster centroids that are not merged during the clustering. Furthermore, the mean shift algorithm is deterministic. After the clustering, the colors in the time sequence $\{\boldsymbol{m}_t\}$ are substituted with their respective class representative index $\{n_t\}$ producing a quantized time sequence. In case of complex states, the probability distribution of every state can be estimated from the colors that were clustered together to form the state.

As a third step, using the time sequence of class representatives, the state transition probabilities p_{ij} are estimated using the frequency of transitions between consecutive states in the source material. Assuming the time sequence is generated from a Markov chain, the frequency of state transitions between consecutive states is a maximum likelihood estimator of the transitions probabilities.

Algorithm 5.1 depicts the learning method. The means_shift_cluster() method takes as input a collection of colors $\{\boldsymbol{m}_t\}$ and the size of the kernel hs and produces as output a collection of the cluster representative colors $\{\boldsymbol{c}_i\}$ and the cluster index array $\{n_t\}$ for each of the input colors.

5.2.2 Simulation

Algorithm 5.2 Online simulation method

Input
color collection $\{\boldsymbol{c}_i\}$ of states, state transition probability matrix $\boldsymbol{P}$

Output
stochastic light effect

On initialize_model **do**

$state = \text{random}(1, \ldots, |\{\boldsymbol{c}_i\}|)$
output_color($c(state)$)

- - - - - - - - - - - - - - - - - - - -

On next_step **do**

$r = $random$([0,1))$
$k = 0$
$sum = \boldsymbol{P}(state, 0)$
while $sum < r$ **do**
  $k = k + 1$
  $sum = sum + \boldsymbol{P}(state, k)$
$state = k$
output_color($c(state)$)

Given a Markov chain with a set of states $\{\boldsymbol{c}_i\}$ and the transition probability matrix $\boldsymbol{P}$, a light effect can be generated by simulating the Markov process. The starting state of the system can be a user set state, a random state or the most probable state of the Markov chain. At regular time intervals, a new state of the system, and thus a new color of the luminaire is computed by sampling from the distribution given by the transition probability matrix. As the new state only depends on the current state of the system, the sampling of the distribution is carried out by generating a uniform $U([0,1))$ random number and based on that number a search through the cumulative distribution of the transition probabilities for the given current state. Algorithm 5.2 depicts the simulation method.

The implementation of this operation is straightforward and computationally cheap, making it ideal for implementation on embedded platforms. The suitability for embedded use is boosted by the small memory requirement of the generated model, that is quadratic to the number of states. The algorithm above uses a linear search, but for larger number of states, a binary search can be easily implemented by storing the cumulative sum of the transition probabilities instead of the probabilities themselves. Additionally, if a central control is used to control multiple light sources, only one copy of the model is needed for all of the light sources.

5.3 User study

To validate recognition and the desirability of the produced light effects, a user study was conducted during an exhibition. The study was designed with an application of home scene setting in mind, which influenced the setup and the method used.

5.3.1 Setup

Two luminaires were used, each having three independent RGB LED light sources and showing the same light effect. Both luminaires used the same model, but the states were not synchronized. The participant could only observe the light reflected from a white surface and couldn't see the light sources or the encasing luminaires.

Stimuli

Four stimuli were used, three automatically created using the proposed method and one hand crafted.

The first stimulus, *fire*, was created from a video of a beach fire. The second stimulus, *underwater*, was created from a low resolution representation of an underwater image. The transition probabilities of the model were computed using the spatial neighbor probabilities, resulting in an effect equivalent to the one produced from a video of a camera randomly moving over the scene with a constant speed. Even though the spatial information from the input image was used in the building of the model, the resulting model did not include spatial information. The third stimulus, *fireworks*, was a manually built impression of multicolored fireworks with periods of faster and slower dynamics. The fourth stimulus, *clouds*, was created from a time lapse video of a cloudy sky.

Two of the stimuli, *fire* and *fireworks* had fast dynamics, contrary to the other two, which had long term, smooth dynamics. The stimulus *fire* had predominantly warm colors, the stimuli *underwater* and *clouds* had cold colors and the stimulus *fireworks* didn't use a specific selection of colors.

Method

The desirability of the produced light effects was measured using a seven point Likert scale [62]. For each of the stimuli, the participants answered the question ”I would like to have this light effect in my living room“ on a scale from ”Not at all“ (-3) to ”Very much“ (3). Additionally, for each stimulus four recognition questions were asked: ”This effect looks like a fire“, ”This effect looks like an underwater scene“, ”This effect looks like fireworks“, and ”This effect looks like a cloudy sky“. The possible answers ranged from ”Not at all“ (-3) to ”Very much“ (3).

The authors are aware of the possible bias on recognition produced by naming the effects that were presented, but considering the application of home scene setting, where the user already knows the target effect she picks, this was not considered a problem in the context of this study. Furthermore, including all four effects in the questionnaire for each stimulus enabled a computation of a confusion matrix, as given in section 5.3.2. Using another method as for example free association, would have resulted in a considerably longer testing time and would require subjective evaluation of the results.

The order of presentation of the stimuli was balanced over the participants. Additional to the questions, participants could provide comments.

Participants

The study was conducted with 202 participants, 155 of which male, with a minimum age of 24, a maximum age of 59 and a median age of 35.5. 64 of the participants had experience working with atmosphere providing light sources.

5.3.2 Results

Results on suitability of the produced light effects for use in a living room environment showed that the most desired stimulus was *clouds*, with a median result of 1, followed by *underwater* with a median result of 1, and *fire* with a median result of 0. The stimulus *fireworks*, which was scored as highly unsuitable for the context and had a median score of -3. As the stimulus *fireworks* was very dynamic, this result was not surprising. A number of participants however noted that they could imagine a limited use of such an effect albeit in a different context. It was unclear, both from the answers to the questions and the additional comments given, what the reason for the difference in desirability of the other three stimuli was. Some people mentioned the color temperature as one of the reasons they scored a certain light effect high or low, while others indicated the dynamics as the primary reason. As there was no slow warm stimulus, a conclusion on the relative importance of these factors could not be given.

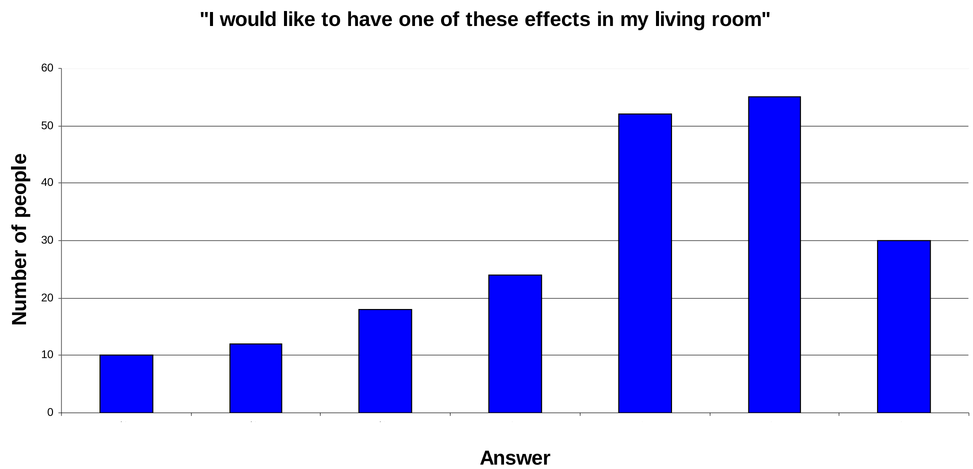

Figure 5.2: Histogram of the maximum score over the four stimuli on the question "I would like to have this light effect in my living room" on a scale from "Not at all" (-3) to "Very much" (3)

To judge the overall desirability of the produced light effects, the maximum score of each participant over the four stimuli was computed. Figure 5.2 shows a histogram of these resulting scores. As can be seen from the histogram, the overall desirability was scored high, with a median score of 1. This shows, together with the comments, that people think that dynamic light effects can be suitable for use in their living room, but only if the effects are very localized like the *fire*, or very slow and subtle as the *clouds* and *underwater* stimuli.

Figure 5.3 shows the median result on the answers from the recognition questions. The images in the row header depict the stimulus that was presented to the participant, whereas the images in the column header represent the effect mentioned in the question. The general recognition of all the effects was high, with most of them having a median score of 2 for the matching stimulus. The most easily recognizable effect was the *fireworks*, while the most confused one was the *underwater*. An interesting effect can be seen in the confusion of the *underwater* and the *clouds* stimuli. When presented with the *underwater* stimulus, participants scored equally high for both *underwater* and *clouds*, but when presented with the *clouds* stimulus, the score for *clouds* was significantly higher than the score for *underwater*. It was also observed that the

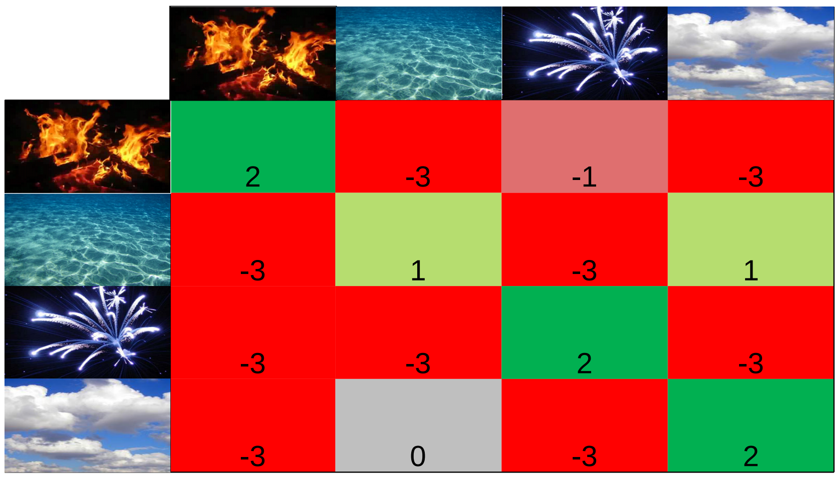

Figure 5.3: The median response of the participants to the question ”This effect looks like ...“ for all stimuli.

effect persisted over different orderings of the stimuli. This additionally shows that different light effects have a different range in which they can be recognized and for some of them that range can be very large.

5.4 Extensions

Two proposals for extension of the base effect were mentioned most often by the participants in the study. The first one was the addition of audio effects synchronized with the light effect. The second one was to extend the effect to a larger number of dependent, synchronized light sources. The first proposal needs a definition of states in a parametric space of a new medium, audio, and as such can be a topic of a wider research effort. The second proposal, however, can be implemented by a simple extension of the proposed method. Moreover, we present a extension that enables the creation of interactive effects.

5.4.1 Multiple dependent light sources

In the single light source case a single color was extracted from each frame resulting in a very low resolution, zero dimensional, representation of the video. The main idea of the extension is to increase the spatial resolution of the color representation and to make the transition probability distribution depend not only on a "previous" state of a single light source, but also on the state of its neighbors. Depending on the shape of the neighborhood, the model can be used to simulate a dependent effect on a one dimensional arrangement (a line) or a two dimensional arrangement (a plane) of light sources. As the training material is two dimensional, extension beyond two dimensions require assuming additional symmetries. In the rest of the section, a two dimensional arrangement is discussed.

To add spatial information to the model, the source material is subdivided into d non-overlapping regions $\{R_s\}_{s\in 1,\dots,d}$ and representative colors $\{\boldsymbol{m}_t|_{R_s}\} = \{\boldsymbol{m}_{t,s}\}$ extracted from each of the regions. The clustering of the colors is done in the same way as in the independent light source case using the mean shift algorithm on the color collection $\{\boldsymbol{m}_{t,s}\}$, producing the collection of representative colors $\{\boldsymbol{c}_i\}$ for the states in the state space $\mathcal{E}$ and the index array $\{n_{t,s}\}$.

The dependence on the state of the neighbors in a neighborhood $\mathcal{N}$ can be built into compound states $\boldsymbol{c}' \in \mathcal{E}^{|\mathcal{N}|+1}$, where $|\mathcal{N}|$ denotes the number of neighbors in $\mathcal{N}$. In this case the transition matrix P is of size $|\mathcal{E}|^{|\mathcal{N}|+1} \times |\mathcal{E}|$ is equivalently to a tensor $\mathcal{P}$ of size $\underbrace{|\mathcal{E}| \times \ldots \times |\mathcal{E}|}_{|\mathcal{N}|+1} \times |\mathcal{E}|$. The size of the transitions tensor increases exponentially with the number of neighbors, which results in both considerably larger storage requirements as well as need for longer training material. An additional problem that the spatial extension introduces is the special treatment of the "edge" and "corner" light sources. As the number of neighbors for these light sources is in general different from the number of neighbors for the light sources in the "middle", a different model has to be learned and simulated.

As an example for the learning and simulation of a model for a set of dependent light sources, we take a planar arrangement of light sources with a neighborhood consisting of the four direct neighbors in the positive and negative direction of two arbitrary axes. The learning algorithm for a "middle" point in a two dimensional dependent arrangement is given in Algorithm 5.3. For presentation simplicity we will denote the neighborhood by $\mathcal{N} = \{1, x+, x-, y+, y-\}$. For every position $\boldsymbol{s}$, the function $1(\boldsymbol{s})$ maps the position to itself, the function $x+$ to its "right" neighbor, $x-$ to its left, and $y+$ and $y-$ to its lower and upper neighbor. By $\mathcal{P}(n_{t,\mathcal{N}(\boldsymbol{s})}, n_{t+1,\boldsymbol{s}})$ we denote the element $(n_{t,1(\boldsymbol{s})}, n_{t,x+(\boldsymbol{s})}, n_{t,x-(\boldsymbol{s})}, n_{t,y+(\boldsymbol{s})}, n_{t,y-(\boldsymbol{s})}, n_{t+1,\boldsymbol{s}})$ of the tensor. Due to the possible sparsity of the resulting transitions tensor specially for larger neighborhoods, there is a possibility of a reachable compound state not having any positive transition probability. Depending on the severity, a number of strategies might be employed, the simplest of which is to leave the state at the position with a terminal compound state unchanged, a strategy that is easy to implement in the learning phase.

For the simulation of the model, every light source has to be given a position $\boldsymbol{s}$ and its neighbors $\mathcal{N}(\boldsymbol{s})$ defined. In a distributed system, the state of a light source has to

Algorithm 5.3 Method : $(\{\boldsymbol{c}_i\}, \mathcal{P}) = \text{learn\_model}(v, hs, d, \mathcal{N})$

Input

example video v, kernel size hs, number of segments per dimension d, neighborhood $\mathcal{N}$

Output

color collection $\{\boldsymbol{c}_i\}$ of states, state transition probability tensor $\mathcal{P}$

---

for all $t \in \{1, \ldots, |v|\}, \boldsymbol{s} \in \{1, \ldots, d\}$ **do**

$\{\boldsymbol{m}_{t,\boldsymbol{s}}\} = \text{extract\_color}(\{v_{t,\boldsymbol{s}}\})$

$(\{\boldsymbol{c}_i\}, \{n_{t,\boldsymbol{s}}\})$ = mean_shift_cluster$(\{\boldsymbol{m}_{t,\boldsymbol{s}}\}, hs)$

$\mathcal{P}^* = \boldsymbol{0}$

for all $t \in \{1, \ldots, |v| - 1\}, \boldsymbol{s} \in \{1, \ldots, d\}$ **do**

$\mathcal{P}^*(n_{t,\mathcal{N}(\boldsymbol{s})}, n_{t+1,\boldsymbol{s}}) = \mathcal{P}^*(n_{t,\mathcal{N}(\boldsymbol{s})}, n_{t+1,\boldsymbol{s}}) + 1$

for all $i \in \{1, \ldots, |\{\boldsymbol{c}_i\}|\}, \boldsymbol{j} \in \{1, \ldots, |\{\boldsymbol{c}_i\}|\}^{|\mathcal{N}|+1}$ **do**

$\mathcal{P}(\boldsymbol{j}, i) = \frac{\mathcal{P}^*(\boldsymbol{j}, i)}{\sum_{k=1}^{|\{\boldsymbol{c}_i\}|} \mathcal{P}^*(\boldsymbol{j}, k)}$

---

be communicated to its neighbors on a state change. Given these, the simulation is equivalent to the independent light source case.

5.4.2 Interaction

An additional advantage of a stochastic model over a fixed low resolution representation of a video is the possibility of interaction. The user can at any time instance force a light source at a position $\boldsymbol{s}$ into a specific state. The model will continue to evolve using the same probability distributions. As an example, using a model built from a video of clouds moving in the sky, the user can have a few interactive possibilities. One of them, arguably desired by most Dutch people, is to force all the light sources to a ”clear sky“ state. Using a spatial selection tool, individual light sources can be forced into one of the ”cloudy“ states, effectively painting moving clouds.

5.4.3 Temporal artifacts

The light effects produced by the stochastic model change at discrete time instances from one digital color to another. If these changes are not fast enough or too large, the effect can be perceived as unsmooth or flickering. Even though this might be desirable for some effects, for most effects this is not the case. In the user study, a considerable number of people stated that they disliked one of the effects due to the abrupt, unsmooth transitions. The question of perceived smoothness and perceived flicker for a general transition through a color space is a complex one and studied more in depth in Chapter 6.

5.5 Conclusions

A method of dynamic light effect generation using stochastic models was presented. Similar to dynamic lighting scenes in nature, the resulting light effects are unpredictable, yet recognizable. Next, a method to learn the stochastic models from a video source of a natural scene was shown. The method extracts the representative colors from the video and subsequently learns the typical transitions between the colors. After the model has been learned, the rendering of the effects has low memory and processing requirements, making it suitable for implementation even on embedded platforms. The recognition of the produced light effects was tested using a large user base and four models. The results show the suitability of the method for dynamic atmosphere creation, but also an appreciation of the produced light effects.

Possible extensions of the light effect to multiple dependent light sources and interactivity are proposed. Due to the richness of possibilities stochastic models use, this chapter is only able to explore the simplest of models and will hopefully lead to a more widespread use of unpredictability in the creation of light effects.

Part II

Temporal quality of light

6

Smoothness and flicker perception of temporal color transitions

"All models are wrong, but some are useful."

George E. P. Box

The first part of this thesis consisted of studies concerning one of the basic challenges in modern lighting systems, the creation of light effects. This chapter marks the beginning of the second part, which studies the means of measuring and controlling the quality of the produced light effects. As the overall quality depends on many factors, making the topic too broad for the scope of a single thesis, a subset of the issues was selected. Two of the largest differentiating features of solid state lighting systems are their dynamic capability and digital control. To further leverage these capabilities, the subset of light quality selected was the study of the perceived quality of temporal light effects and the visibility of digital artifacts.

As the final evaluation of the quality of the produced light and the visibility of artifacts is done by a human observer, a detailed knowledge of the human visual system and its sensitivity is needed. Unfortunately, no models exist that are specifically built to predict the response of the human visual system to dynamic light patterns beyond the ones for achromatic flicker. Furthermore, this chapter demonstrates the unsuitability of the existing models built to characterize spatial effects for the prediction of temporal ones.

Taken all this into account, this part of the thesis does not concentrate on algorithms, but mostly uses psychophysical methods to explore and model the effects of interest.

This chapter is based on [89, 90]

6.1 Introduction

Advances in lighting, especially in Solid State Lighting, enable new uses of light. Having improved spatial and temporal resolution, more saturated primaries and lower power consumption, LED based lighting systems can be used to design more complex and attractive lighting atmospheres. One of the largest differentiators of such lighting systems are their dynamic capabilities. However, the produced dynamic lighting atmospheres need certain properties to be attractive to the users.

Perceived smoothness is one of those desirable properties. Aside from a limited set of applications, such as disco lights, concerts, or attention attractors, abrupt changes in environment lighting are hardly perceived as pleasant. Other properties of produced lighting atmospheres that might have an influence on the attractiveness such as the hue composition or the spatial distribution, moreover, are more subjective.

The work presented in this chapter studies temporal properties of human color vision relevant to dynamic lighting applications, namely thresholds for smoothness perception of linear transitions and flicker visibility. To understand the possible source of problems connected to smoothness perception, we discuss the design of modern lighting systems and applications first.

Modern lighting applications use a digital control of the light source, with a limited number of intensity levels. Contrary to the analogue systems which have a continuous change in color, in digital systems the smallest distance between two colors, both in color and time, is limited by the resolution of the system. Similar to spatial color perception, an inappropriate minimum distance between colors can introduce perceived discontinuities.

Existing dynamic lighting systems use the device color space (usually $\boldsymbol{RGB}$) of the lights to control temporal changes. To produce smooth light transitions, low pass filters are applied on the individual color channels. Under some conditions, especially for light effects computed from another medium (such as a highly dynamic video signal for the Philips AmbiLight$^{\text{TM}}$), this leads to seemingly transitions between chromatic colors that are perceived as too slow. In the case of content dependent dynamic lighting, notably for video, this introduces a mismatch between the color of the lighting and the representative color of the video frames during the transition. A video transition from a red sunset to a blue underwater scene is followed by a light transition being purple for a noticeable amount of time. This behavior is deemed undesirable by most users.

The above mentioned problem is present in all dynamic lighting systems that control temporal changes in a device color space. The core of the problem is that the properties of the human visual system are not taken into account in device color spaces. Previous work on the temporal properties of the human visual system shows differences in the way intensity and chromaticity changes are perceived. Namely, the human visual system processes intensity changes faster than chromaticity changes [22, 8]. Moreover, the changes in chromaticity are smoothed by the human visual system more than the changes in intensity [97, 99]. Using a device color space to control the temporal changes does not allow the use of such results.

To compute the maximal distances between colors that still produce a smooth spatial pattern, the notions of visibility threshold and just noticeable difference [69] were introduced. The continuation of the work on spatial just noticeable differences led to the development of, among others, the ***CIE Luv***, ***CIE*** $\boldsymbol{L^*a^*b^*}$, and ***CIECAM02*** color spaces [32], which show a relatively good correspondence between the perceived and the predicted differences between colors and smoothness of spatial transitions.

Unfortunately, no spaces that predict the perceived smoothness of temporal patterns exist. The fact that the perception of temporal transitions depends on the frequency at which the changes are made, further complicates the representation and smoothness prediction in the temporal case. To gain better understanding of the way the human visual system processes temporal patterns in the context of dynamic lighting applications, we designed two exploratory experiments and an additional confirmatory experiment, of which we present the results in this chapter.

Previous work on temporal properties of the human visual system closest to the topic of interest of this chapter comes from the area of flicker sensitivity. In [23, 24], De Lange describes flicker sensitivity at different frequencies and for different average luminance levels and types of stimuli. The results of De Lange are supported by results of Kelly [54, 55], in which he additionally studies the effect of the surround average luminance and spatial context on flicker perception.

Using different methods, several authors report differences between sensitivity to luminance and chrominance flicker. In [99], the response of the visual system is modeled as a finite impulse response filter and differences in the properties of luminance and chrominance flicker are demonstrated. Kelly [57] uses spatio-temporal properties to show a difference between luminance and chrominance flicker effects. In all the above studies, the number of color pairs used as the two settings in the chromaticity flickering stimulus was limited.

This work differs from previous work in a number of points. First, in one of the experiments presented we use discrete linear transitions, in line with the application requirements discussed above, while in prior work either flicker or a fixed number of rectangular pulses are used. Second, the chromaticity changes are further subdivided in chroma and hue changes. And third, the thresholds are computed not only for a number of luminance levels, but also for different levels of chroma and hue, to be able to model the distribution of smoothness thresholds over the color space, similar to the spatial McAdam ellipses.

To be able to compare the temporal sensitivity results to results from spatial visibility thresholds, the presented experiments use the ***CIE $L^*a^*b^*$*** color space and the ΔE_{ab} metric. The geometric middle point, in ***CIE $L^*a^*b^*$***, between the starting and ending color is called a *base color point* in the rest of the chapter.

Based on previous studies on temporal properties of human vision described before, we form and test the validity of two main hypotheses:

- There is a difference in sensitivity for lightness and chromaticity changes. The thresholds for lightness are lower.
- There is an effect of the base color point on the visibility thresholds for lightness, chroma and hue changes.

6.2 Smoothness thresholds for linear color transitions

The first experiment was designed to measure the sensitivity for discontinuities in linear temporal color transitions. The transitions are built in the ***CIE $LCh_{a^*b^*}$*** color space and the threshold are expressed in ΔE_{ab}, the Euclidean distance in the ***CIE $L^*a^*b^*$*** color space.

Given the results of previous research on related topics discussed above, we investigate the effect of *frequency* of the changes, *base color point* and *direction* of change on the visibility threshold tested. For possible directions of change, directions parallel to the axes of the ***CIE $LCh_{a^*b^*}$*** color space were taken. The smoothness visibility threshold was measured using a tuning procedure.

6.2.1 Method

Equipment

As a light source, a LED lamp was used, with three RGB LEDs at the bottom and three RGB LEDs on the top of a light guide. The light guide was used as a mixing chamber and was painted with a diffuse white paint on one side to insure outcoupling of the light in the desired direction. The LEDs were driven using pulse width modulation (PWM) with a driving frequency of 500 Hz and 11 bit levels. The driver accepted ***RGB*** values in the device space of the LEDs. The target stimuli were defined in ***CIE $LCh_{a^*b^*}$*** and transformed via ***CIE XYZ*** to the ***RGB*** device space of the light units using a computer program running on a PC connected to the light units.

For easier generalization, the D65 white point, not the white point of the system, was used in the transformations. Care was taken that the produced colors were inside the device gamut of the light source.

In order to validate that the light sources produced the expected colors and temporal patterns, three experiments were carried out. First, we tested the independence of the L, C and h axis using a colorimeter. Second, the linearity of each channel separately and all the channels together was tested using a colorimeter. Third, the temporal responses of the lights were tested using a high speed camera. The results of these experiments didn't show significant differences between the model used and the physical properties of the light produced by the hardware.

Stimuli

The stimuli were discrete linear temporal transitions around a *base (middle) color point* LCh_{base} in a *direction* $d \in \{L, C, h\}$ and *frequency of change* f. The transitions were characterized with a total length A and a step size S applied every $\frac{1}{f}$ seconds. The stimuli consisted of $\frac{A}{S} - 2$ number of steps with size S and additional $2p$ number of steps with step sizes $\frac{S}{2^i}$, $1 \leq i \leq p$. For the experiment, a value of $p = 6$ was used. Example stimuli for a change in direction d are depicted in Figure 6.1.

Single transitions were repeated in alternating directions along the axes of the color space to make the user task easier. To diminish the perceived effects of the edges of the repeated transitions, the length of the transition was at least one order of magnitude larger than the step size and smoothing of the edges of the repeating pattern was applied, as described above.

The linear transitions were varied in *frequency*, *base color* and *direction of change*. During each tuning, the step size S was changed based on the user input. This means that the frequency and total length of the transition were kept constant, but the speed and the time of a transition changed. A pilot experiment showed that for frequencies larger than 30 Hz, the speed was too high to detect discontinuities for changes in chroma and hue. Therefore the frequencies used in the experiment were 5, 10 and 20 Hz for all base colors and 30 and 40 Hz for a subset of the base color points, as shown in Table 6.1.

Two criteria were taken into consideration in the choice of the base chromatic colors. First all the three primary hues were included. Second, the available amplitude of change at those base points had to be large enough. Using primaries and a secondary color at a moderate chroma level fulfilled both criteria. Beside the chromatic colors, three achromatic color points at different luminance levels were included.

For possible directions of change, directions parallel to the axis of the ***CIE*** $\boldsymbol{LCh_{a^*b^*}}$ color space, lightness, chroma and hue, were selected. The stimuli with achromatic

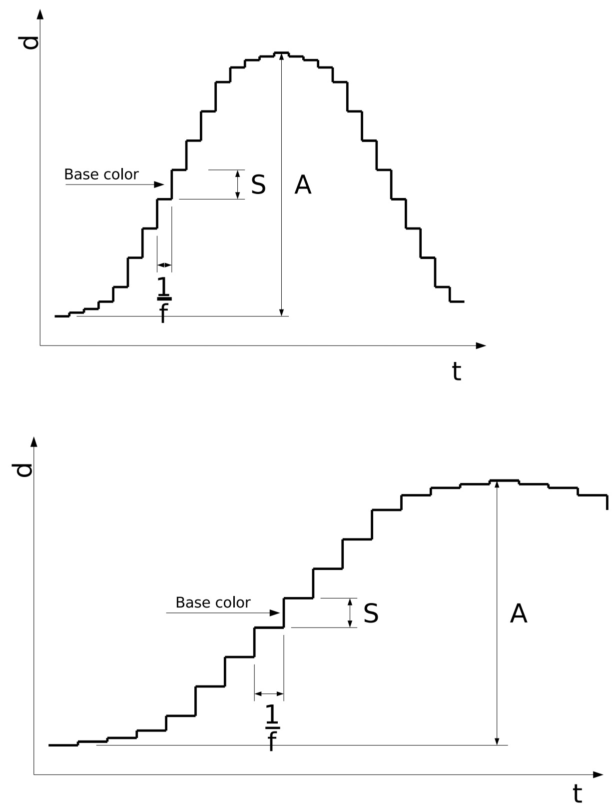

Figure 6.1: Base stimuli for different frequencies.

Name	L	c	h	x	y	Y	Frequency (Hz)
Magenta	60	60	350	0.397	0.254	0.281	5 10 20
Red	60	60	45	0.506	0.370	0.281	5 10 20 30
Green	60	60	150	0.274	0.484	0.281	5 10 20
Blue	70	60	290	0.224	0.201	0.408	5 10 20
High L	75	0	0	0.313	0.329	0.483	5 10 20 30 40
Middle L	50	0	0	0.313	0.329	0.184	5 10 20 40
Low L	30	0	0	0.313	0.329	0.062	5 10 20 30 40

Table 6.1: Base color points in (***CIE*** $\boldsymbol{LCh_{a^*b^*}}$) and (***CIE xyY***) used in the experiment, with corresponding frequencies.

base color points were only varied along the lightness axis. The length A of all transitions in direction L was $50\Delta E$, for direction C - $120\Delta E$, and for direction h - $100\Delta h$, where Δh is the angular difference in degrees between the extremes of the transitions.

The total amount of stimuli was 53. The stimuli were presented in two parts, a training set of stimuli given in the same order for every participant and a testing set of stimuli which was randomized. The randomization was done to average over possible learning effects, tiredness and flicker adaptation [92].

Participants

Nine females and thirteen male subjects participated in the experiment. Their ages ranged from 22 to 35 years. All participants had normal color vision, tested with the Ishihara's test of color deficiency. Three of the subjects had experience with viewing dynamic color patterns.

Procedure

The above described LED lighting system was used as a light source in the experiment. The light coming from the LED shone on a wall and the participant was only able to see the reflected light. The light source itself was covered by black paper so the participants were not able to look directly into the light source. The participants were seated on a distance of 3 meters from the wall at which the maximum luminance of the reflected light from a $D65$ stimulus was $35cd/m^2$.

Before the start of the session the participants received instructions. Participants were instructed to look at a fixation point in the middle of the light spot reflected on the wall and to judge whether the discontinuity of the color transition was visible or not. By selecting a button on a control keyboard they could

1. Increase the step size,
2. Decrease the step size, and
3. End the trial and go to the next.

Participants were asked to select the step size at which the discontinuity was just *not* visible, i.e. the transition was perceived as smooth. Preceding the experiment a training session was implemented to get used to the stimuli and the test.

Each trial started with the largest step size which was $\Delta E_{ab} = 15$ for hue and chroma changes and varied between $\Delta E_{ab} = 2$ and $\Delta E_{ab} = 7$ for lightness changes depending on frequency. The largest step size was selected as a largest step possible to keep a desired ratio between the total length and the step size or a step size that produces a clearly visible step. The step size could be increased or decreased by $\Delta E_{ab} = 0.1$ for lightness and $\Delta E_{ab} = 0.5$ for chroma and $\Delta h = 0.5$ for hue. After each hue tuning the selected step sizes for hue changes in Δh were transferred to ΔE_{ab}.

After five participants, the data showed that the starting step for all the conditions with lightness changes at frequencies of 20 Hz and 30 Hz was too low. After these five participants the starting thresholds were increased for these 2 conditions resulting in the values described above.

During the experiment the participants could chose their own break when they got tired. Because the tunings where randomized the effect of this break is averaged over all the tunings.

6.2.2 Results

The step size, given in ΔE_{ab}, for which the participant perceived the transition as just smooth is defined as the threshold. For lightness changes at 20 and 30 Hz, the first 5 participants had a lower threshold than the other participants due to the lower starting step size. Although this effect was not significant, the data for these conditions and participants was excluded from the analysis.

For each condition the distribution of the results over the participants was tested for normality. For some conditions the data was skewed. As for all of these the results were close or at the maximum step size in the experiment, some participants perceived the transition with the starting step size as smooth and the actual threshold is larger than the starting step size. The results were skewed for lightness changes for 40Hz, and chroma and hue changes at 20 and 30Hz for all participants. The skewed results were excluded from the analysis.

The mean thresholds and 95% confidence intervals are given in Figure 6.2. The dashed line in the figure represents the maximum step size given to the participants in the experiment. Figure 6.2 includes the conditions for which the date was skewed for illustration. Note that for those conditions for which the distribution is skewed, the confidence intervals are not reliable.

For a overview of all effects, the results were fitted to a linear mixed model and analyzed with an ANOVA (Analysis of Variance [73]) with *base color point*, *direction of change*, and *frequency* as fixed factors and *participant* as random factor. To ensure a complete design, only the *base colors* and *frequencies* that that appear for all *directions* were taken in the model. All main effects for both the fixed and the random factor and all interactions between the fixed factors were used in the model. Additionally, the two-way interaction effects between the random factor and the fixed factors were included. A significant main effect of *direction* was found ($F(2,680) = 1078,\ p < 0.001,\ r = 0.87$), as well as for *frequency* ($F(2,680) = 188,\ p < 0.001,\ r = 0.60$), *base color point* ($F(3,680) = 31,\ p < 0.001,\ r = 0.34$) and *participant* ($F(20,680) = 12,\ p < 0.001,\ r = 0.51$). A significant interaction effect between *frequency* and *base color point* ($F(6,680) = 7,\ p < 0.001,\ r = 0.24$) was found, as well as between *frequency* and *direction* ($F(4,680) = 26,\ p < 0.001,\ r = 0.37$), *base color point* and *direction* ($F(6,680) = 21,\ p < 0.001,\ r = 0.40$) and a three-way interaction between *frequency*, *base color point* and *direction* ($F(12,680) = 4,\ p < 0.001,\ r = 0.27$). None

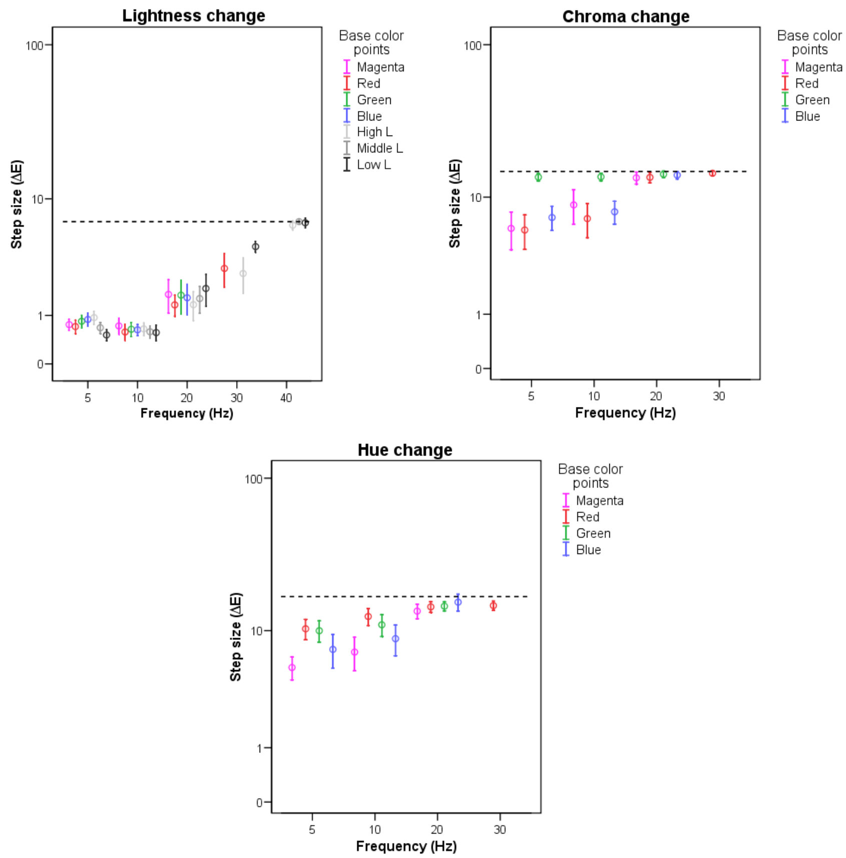

Figure 6.2: Thresholds for smoothness perception of linear temporal transitions at different frequencies, different base color points and directions of change.

of the interaction effects with *participant* were significant.

The large effect size of *direction* and the interaction effects of *direction* with both *frequency* and *base color point* suggests that splitting up the data for different *directions* of change is interesting. A number of smaller models using a subset of the factors that showed significant interactions were fitted. The most important effects from the overall model and the subdivided models are further discussed below.

Lower thresholds for lightness changes compared to chroma and hue changes
The results from the analysis of the overall model demonstrated that the effect of *direction of change* was significant. A post hoc test revealed that the thresholds for lightness changes were significantly lower than the thresholds for changes of Hue and Chroma. This result validated, in the context of this experiment, the first hypothesis.

The result is also in accordance with previous results on flicker sensitivity that show a difference between luminance and chrominance flicker sensitivity [99]. It has to be noted, however, that previous studies use modulation thresholds to describe the sensitivity, while we use absolute step size thresholds. Furthermore, taking into account that the color space and distance metric used in the experiment are near to uniform for spatial differences in color, this result shows that the sensitivity to temporal changes in lightness is higher than the sensitivity sensitivity to changes in chroma and hue.

Independence of smoothness thresholds for lightness change on the chromaticity of the base color A model with *base color point* (for both the chromatic and the achromatic points) and *frequency* as fixed factors and *participant* as random factor for the subset of the data in the lightness *direction* showed that there is a significant effect of *frequency* ($F(2,365) = 135,\ p < 0.001,\ r = 0.65$) and *participant* ($F(20,365) = 18,\ p < 0.001,\ r = 0.70$), and an interaction effect of *frequency* and *base color point* ($F(12,365) = 2,\ p < 0.007,\ r = 0.27$). Due to the significant interaction effect, smaller models per *frequency* were built. A post hoc analysis showed there was no significant difference of the thresholds for lightness changes for the chromatic color points (“Magenta”, “Red”, “Green”, “Blue”) at any frequency. Significant main effect of *base color* were found for 5 Hz ($F(6,120) = 9,\ p < 0.001,\ r = 0.56$) and 20 Hz ($F(6,90) = 2,\ p < 0.014,\ r = .39$). A post hoc test revealed that the effect was caused by the condition “LowL” for both frequencies. For 10 Hz there was no main effect of *base color point*. This shows independence of the thresholds to changes in chromaticity and dependence on lightness. The independence of the thresholds for lightness changes on chromaticity of the *base color* is in contradiction to the second hypothesis.

Dependence of thresholds for chroma and hue changes on the chromaticity of the base color Contrary to lightness changes, both thresholds for chroma and hue changes had a significant dependence ($F(3,220) = 40,\ p < 0.001,\ r = 0.59,\ F(3,220) = 18,\ p < 0.001,\ r = 0.44$) on the *base color point*. For chroma changes, a post hoc test revealed that thresholds for “Green” were higher compared to all other base colors. For hue changes, a post hoc test revealed that both “Green” and “Red” had higher thresholds compared to “Magenta” and “Blue” that had the lowest threshold. This proves the second hypothesis for changes in chroma and hue.

Difference in behavior of chroma and hue changes A surprising effect can be observed in the behavior of the thresholds for the *base color point* “Red”, hinted by the significant interaction effect between *base color point* and *direction*. While for Chroma changes, the thresholds for “Red” are not significantly different from the thresholds for “Magenta” and “Blue”, the Hue changes for the same condition are significantly different from the thresholds for “Magenta” and “Blue” and are in the same group as thresholds for “Green”. This shows that possibly different mechanisms are used in the processing of hue and chroma. One of the possible causes of this effect is a difference in speed of chromatic adaptation under changing hue or chroma.

Effect of frequency The behavior of the smoothness thresholds of lightness changes for different frequencies is in accordance with the behavior of the visibility thresholds for luminance flicker that can be found in literature, for example [23]. For easier comparison, the absolute thresholds are given as contrast sensitivity, or inverse contrast, computed using Michelson’s formula, and are depicted in Figure 6.3. The thresholds for chroma and hue changes were only tested at three frequencies, which makes a direct comparison with results from chromatic flicker sensitivity hard.

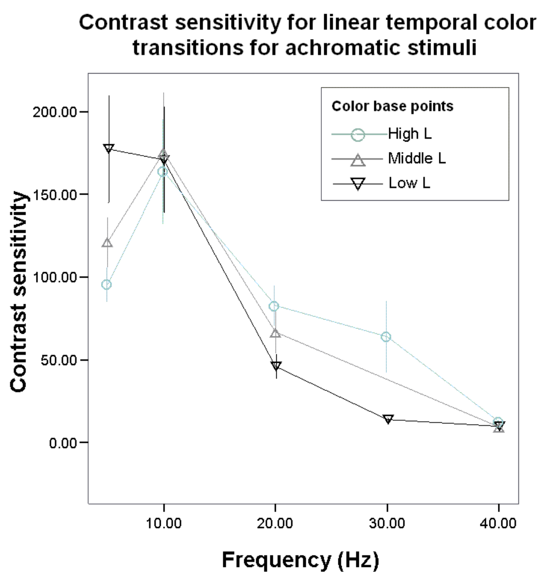

Figure 6.3: Thresholds for smoothness perception at different frequencies and different base color points for lightness changes given in contrast sensitivity.

6.3 Speed of smooth transitions

Another way to represent the behavior of the smoothness thresholds for different frequencies is by using the speed of the transitions that are just smooth, i.e. the speed of linear transitions with a step size equal to the threshold step size. We define the speed of the transition (v) in terms of the threshold step size (s) and the frequency of change

(f) as $v = f \times s$. Figure 6.4 depicts the speed of just smooth transitions for a set of frequencies for different *directions of change* (lightness, chroma and hue), averaged over the *base points* and *participants*. The speed is calculated based on the results for the step thresholds of the experiment described above.

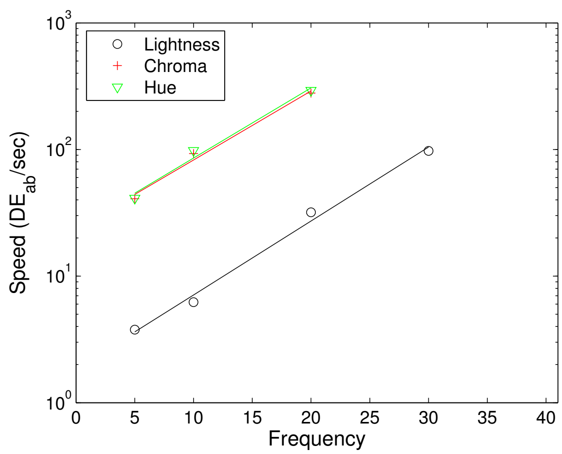

Figure 6.4: Dependence of the speed of just smooth transitions on frequency and a linear fit of the data. Speed is given on a logarithmic scale.

Surprisingly, the results for the natural logarithm of the averaged speed thresholds show a clear linear relationship to the frequency. Furthermore, the fitted lines for the different directions of change all show a similar slope (0.135, 0.125 and 0.128 for lightness, chroma and hue respectively) and differ only in their intercept values. The different intercept values account for the absolute difference in the speed threshold between luminance chromaticity changes.

As interesting as this result is, it is based on a fit of 4 (for lightness) and 3 (for chroma and hue) averaged points, which makes the evaluation of the goodness of the fit hard. To validate the above relationship, an additional experiment was carried out. The additional experiment was a repeated measurement with three participants, denoted by IV, RC and DS. As there were no interaction effects of participant to any other factor in the previous experiment, we expect the same trends in the data for a large set of participants in the first experiment and the limited set of participants in the validation experiment. The absolute value of the speed is, however, expected to be different in the validation set because the participants were all experienced with observing temporal color changes.

The same experimental setup as in the previous experiment was used. The stimuli for the validation experiment were selected based on the stimuli given in Table 6.1 and by adding additional frequencies. An overview of the base points and frequencies of the stimuli used in the validation experiment is given in Table 6.2. All base points and frequencies were tested in three directions of change, except for frequencies and base points marked with a (∗), which were used only for lightness changes.

Name	L	c	h	Frequency (Hz)
Red	60	60	45	2.5 5 10 15 20 25 30 40∗
Green	60	60	150	2.5 5 10 15 20 25 30 40∗
Blue	70	60	290	2.5 5 10 15 20 25 30 40∗
High L^*	75	0	0	2.5 5 10 15 20 25 30 40
Middle L^*	50	0	0	2.5 5 10 15 20 25 30 40
Low L^*	30	0	0	2.5 5 10 15 20 25 30 40

Table 6.2: Base color points in (***CIE LCh***$_{a^*b^*}$) used in the validation experiment, with corresponding frequencies.

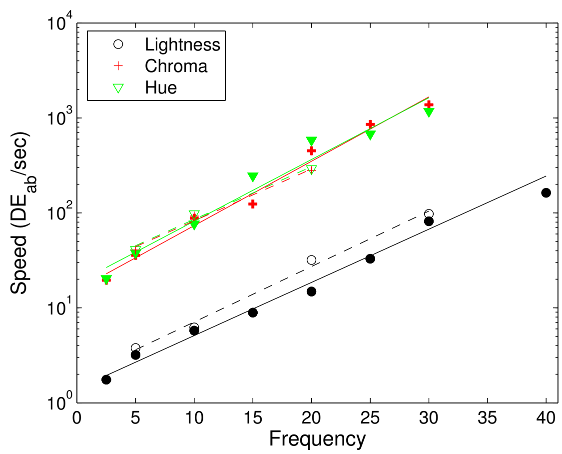

Figure 6.5: Dependence of the speed of just smooth transitions on frequency and a linear fit of the data in the original (dashed lines) and in the validation (solid lines) experiment. Speed is given on a logarithmic scale.

Figure 6.5 depicts the results of the validation experiment compared to the original experiment. The solid lines represent the data from the validation experiment.

The results are in accordance with the results from the original experiment with resulting fit slopes of 0.135 for lightness, 0.141 for chroma and 0.145 for hue changes. Additionally, the higher number of frequency points allows to better estimate the goodness of the fit. The R^2 of the fit was 0.973 for lightness, 0.978 for chroma and 0.944 for hue changes. As expected, the absolute value speed threshold for lightness was lower in the validation experiment due to trained observers.

The above results show that for the range of $2.5Hz$ to $30Hz$ (or $40Hz$ for lightness), the speed threshold of a transition in a direction parallel to the axes of ***CIE LCh***$_{a^*b^*}$

can be determined using a simple exponential law of the form

$$v = a_0 \mathrm{e}^{slope}, \tag{6.1}$$

simplifying significantly the modeling of the sensitivity to discontinuities of linear temporal transitions.

6.4 Flicker visibility thresholds

One of the limitations of the first experiment was that the amplitude of the change was limited and hence for large step sizes the transition speed became too large to find the smoothness threshold. This resulted in unreliable data for higher frequencies. The amplitude limitation has a smaller impact in the case of flicker where the light alternates between two values. Testing the flicker sensitivity in the same setup can show a possible connection between the smoothness thresholds for linear transitions and the visibility thresholds of flicker. A possible connection between the smoothness and flicker thresholds can be a step towards building a model for smoothness perception of temporal color changes based on a more easily obtainable model of flicker visibility. Even thought chromatic flicker has been studied earlier, the color points around which it has been modeled is very limited. To be able to directly compare the results of smoothness and flicker visibility, a second experiment was designed. The same hypotheses as in experiment one were used in this experiment. Additionally, the same effects as in the first experiment were expected.

6.4.1 Method

The same setup as in the first experiment was used.

To be able to compare the results with the results from the first experiment, similar stimuli were used. An overview of these stimuli is given in Table 6.3. Standard square shaped flicker stimuli were used in the experiment, alternating between $LCh_{base} - \frac{S}{2}$ and $LCh_{base} + \frac{S}{2}$ every $\frac{1}{f}$ seconds. As before LCh_{base} denotes the base (middle) point, S the step size, in this case equal to the amplitude of the flicker, and f the frequency.

Participants

Nine females and thirteen male subjects participated in the experiment. Their ages ranged from 22 to 40 years. All participants had normal color vision, tested with the Ishihara's test of color deficiency. Three of the subjects had experience with viewing dynamic color patterns.

Name	L	c	h	x	y	Y	Frequency (Hz)
Magenta	60	60	350	0.397	0.254	0.281	10 20 40 60
Red	60	60	45	0.506	0.370	0.281	10 20 40 60
Green	60	60	150	0.274	0.484	0.281	10 20 40 60
High L	75	0	0	0.313	0.329	0.483	5 10 20 40 100
Middle L	50	0	0	0.313	0.329	0.184	5 10 20 40 100

Table 6.3: Base color points in (***CIE*** $\boldsymbol{LCh_{a^*b^*}}$) and (***CIE*** $\boldsymbol{xyY}$) used in the experiment, with corresponding frequencies.

Procedure

The same experimental procedure as in the first experiment was used.

6.4.2 Results

The results from the second experiment are the visibility thresholds on the flicker amplitude for which no flicker was visible (just not visible flicker boundaries).

The resulting mean thresholds and 95% confidence intervals from the experiment are given in Figure 6.6.

The results were analyzed using an ANOVA with *base color point*, *direction* of change and *frequency* as fixed factors and *participant* as random factor. A similar set of models as in the first experiment was used. A number of significant effects, also visible in the Figure 6.6, were found. The independence of the thresholds for lightness changes on the chromaticity of the base color could not be checked due to the limited set of base colors used in the experiment.

Lower thresholds for lightness changes compared to chroma and hue changes As in the results from the first experiment, the visibility threshold for lightness flicker was significantly lower ($F(1,506) = 9,\ p = 0.003,\ r = 0.13$) than for chroma and hue flicker at the same frequency.

Dependence of thresholds for chroma and hue changes on the chromaticity of the base color A significant dependence on chromaticity of the base color point on the visibility thresholds for hue ($F(2,242) = 36.859,\ p < 0.001,\ r = 0.48$) and chroma ($F(2,242) = 18,\ p < 0.001,\ r = 0.36$) flicker was found, in accordance with the results from the first experiment.

Difference in behavior of chroma and hue changes The same behavior for the base color “Red” as in the first experiment can be observed in the results for flicker

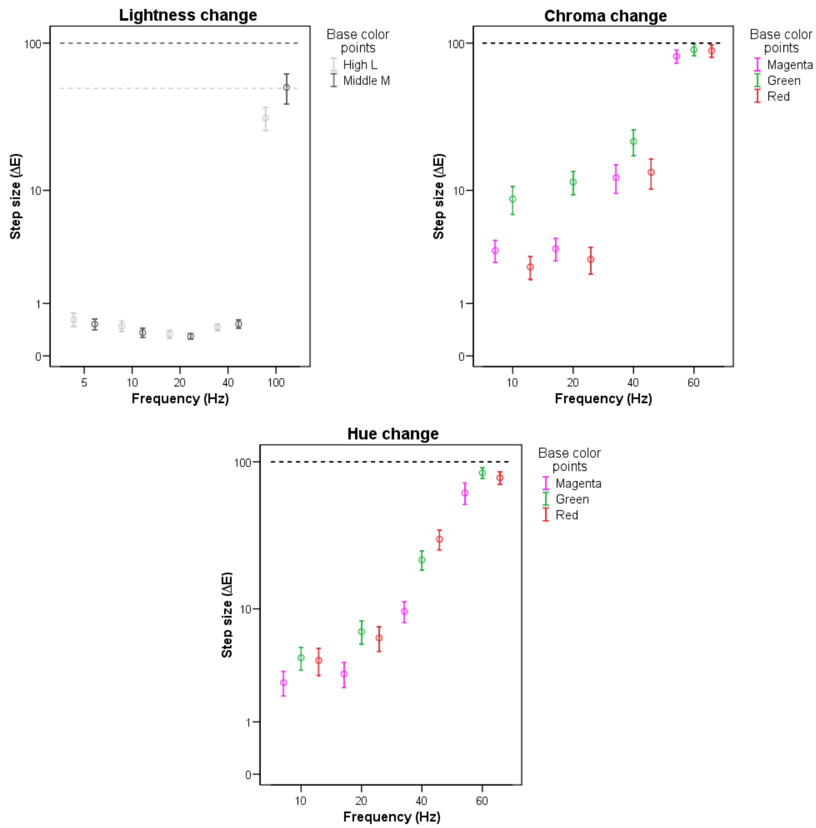

Figure 6.6: Visibility thresholds for flicker at different frequencies, different base color points and directions of change.

visibility.

Effect of frequency The behavior of the flicker thresholds of lightness, chroma and hue changes for different frequencies is in accordance with the behavior of the visibility thresholds for flicker that can be found in literature [99]. The contrast sensitivities computed from the flicker visibility thresholds, are depicted in Figure 6.7. It has to be noted, however, that the flicker frequency as given in literature is usually computed as the number of full cycles per second. In the results presented here, the frequency used is the frequency of change, thus twice the number of cycles per second. An interesting observation connected to the definition of frequency is that the peak in the sensitivity for linear transitions is around 10 Hz, while for flicker sensitivity it peaks at 20 Hz, corresponding to a frequency of 10 cycles per second.

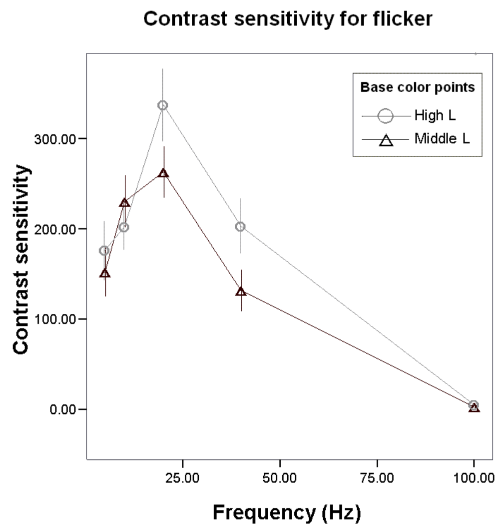

Figure 6.7: Thresholds for flicker visibility at different frequencies and different base color points for lightness changes given in contrast sensitivity.

Dependence of the type of change A significant difference was found between the smoothness and flicker visibility thresholds at all conditions in the intersecting set of the two experiments presented. In all cases the flicker thresholds were lower, showing a higher sensitivity to discontinuities in flicker. However, since all the effects found for smoothness perception found in the first experiment show the same tendencies as the thresholds for flicker in the second experiment, this suggests that it is possible use the results from flicker sensitivity to predict the differences for smoothness threshold under different conditions.

6.5 Speed of transition for flicker

Analogous to the definition of speed for the linear transitions, we can define the amount of change in one second as the speed of flicker and determine its dependence on the frequency. We define the speed (v) in terms of the step size (s) and the frequency (f) as $v = f \times s$. Figure 6.8 depicts the speed thresholds of just perceivable flicker for a set of frequencies for different directions of change (lightness, chroma and hue), averaged over the base points and participants, where the speed was determined from the flicker thresholds of the second experiment.

The results show an exponential relationship as in the results for the speed of linear

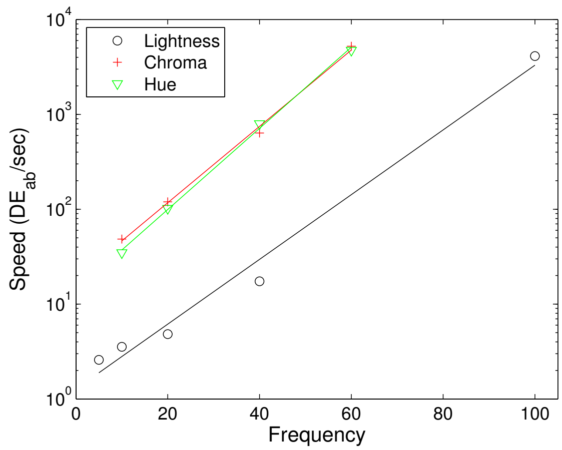

Figure 6.8: Dependence of the speed thresholds of just not perceivable flicker on frequency and a linear fit of the data. The speed is given on a logarithmic scale.

transitions. In the case of flicker, the computed slopes are 0.078 for lightness, 0.093 for chroma and 0.098 for hue.

An additional validation experiment for flicker was carried out using the same setup as the validation experiment for the speed of linear transitions. The same base points and frequencies were used also to allow for better comparison.

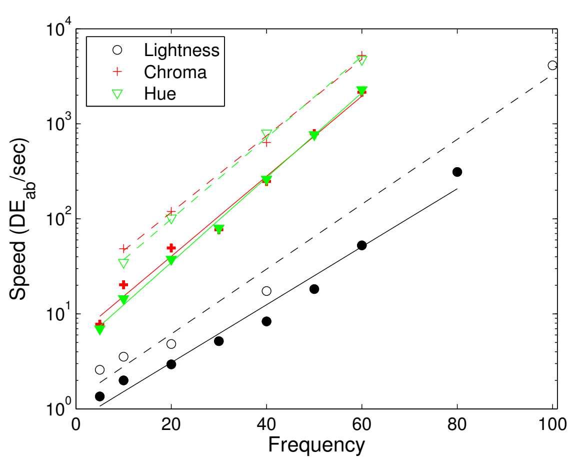

Figure 6.9: Dependence of the speed thresholds of just not perceivable flicker on frequency and a linear fit of the data for the original (dashed lines) and the validation (solid lines) experiment. The speed is given on a logarithmic scale.

Figure 6.9 depicts the results of the validation experiment. The same trends as in the results for linear transitions can be observed. The resulting slopes were 0.07 for lightness changes, 0.097 for chroma changes and 0.103 for hue changes, with R^2 of 0.971, 0.993 and 0.996.

For comparison, Figure 6.10 depicts the results of both linear transitions and flicker with their respective linear fits. The results depicted are taken from the validation experiments. The linear transition results are depicted using solid lines.

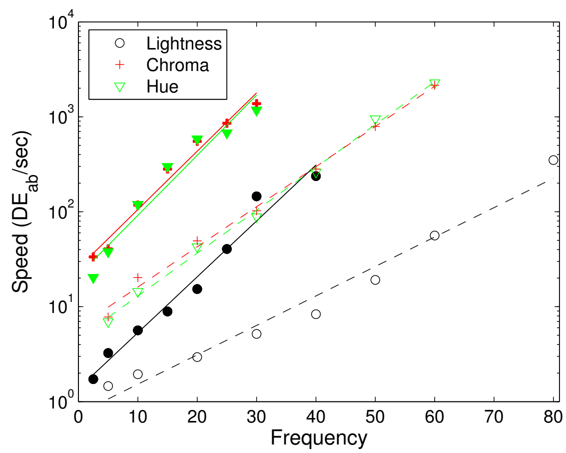

Figure 6.10: Dependence of the speed of just smooth transitions on frequency and a linear fit of the data in the validation experiment. Speed is given on a logarithmic scale.

6.6 Temporal control

Based on the above results, a simple algorithm for the perceptually inspired control of light effects can be devised. Knowing the maximum speed in different directions of change that produces smooth and flicker free transitions, the system can limit the maximum allowed change in color, depending on the system parameters.

The results can also be used as input for the design of the hardware and the system. For example, to design a light system that can produce smooth linear color transitions at a speed of $50\,\Delta E/s$ for lightness levels above $2\,cd/m^2$, the frequency of change has to be at least $20\,Hz$ and the smallest step $0.06\,cd/m^2$ or smaller (a precision of 1024 levels for a $35\,cd/m^2$ system)

An additional problem arises when the system is coupled to another medium, like in the example AmbiLight application. In this case, the extraction algorithm might dictate a transition to a color that will result in an unsmooth transition. Knowing the differences in sensitivity to luminance and chrominance changes, its easy to imagine a special case in which only the luminance part of the change is above the smoothness limit. The way the algorithm than limits the step is ambiguous and more studies to find the best means are needed. A way in which this can be answered is by finding the optimal temporal path for the transition between two colors. This problem is the object of study of the next chapter.

6.7 Conclusions

Results from two experiments that support previous results from flicker sensitivity research and demonstrate new effects relevant to production and control of dynamic lighting were presented. Using a spatially nearly uniform color space, ***CIE LCh***$_{a^*b^*}$, we show that the sensitivity for changes in the lightness direction is higher than in the chroma and hue directions. Independence on the chromaticity of the base color point for lightness changes is demonstrated. Temporal chroma and hue changes show dependence on the base color point. The size of the above effects justify the use of a model of perceived smoothness for control of dynamic lighting over the traditional control in a device dependent color space.

Even though a difference in the absolute thresholds and different dependence on frequency is shown for smoothness thresholds for linear transitions and flicker visibility thresholds, the same base color and direction of change effects were demonstrated. The correlation of the results for different types of dynamic transitions enables building a model for perceived smoothness based on flicker visibility thresholds.

The results of the experiments can be used to determine system requirements and devise temporal control algorithms which take into account the properties of the human visual system. The results for the dependence of the speed on frequency give a straightforward algorithm to determine the speed of a transition between two points given the frequency of the lighting system.

7

Discrimination and preference of temporal color transitions

"A straight line may be the shortest distance between two points, but it is by no means the most interesting."

Doctor Who

From the results in the previous chapter, the speed of a transition that appears smooth can be computed. To get to this desired speed, however, often new points between the points already in the transition have to be introduced. Thus, the question of the perceptually most preferred path between two points in a color space arises. In this chapter, the answer to this question is given in two steps. First, the difference between temporal transitions connecting two points that are perceived as distinct is explored. With this knowledge, in the second part of the chapter, the most preferred transition among a set of distinct possibilities is found. This chapter is based on [109, 110].

7.1 Introduction

Many lighting applications generate colored light that varies over time. For instance, the color of the light surrounding an AmbiLight TV changes in accordance with the color of the content shown on the display to enhance the experience of watching TV. However, when the light follows the content from frame to frame, the resulting step-like pattern will be perceived as jerky or flickering. In order to produce slowly changing, smooth light effects, frame skipping and temporal sub-sampling can be used. As the color is specified only at certain moments in time, intermediate colors have to be calculated such that the light effect is perceived as consistent with the video.

Dynamic light can also be used for atmosphere creation. By changing the color of the illumination of environments, such as shops or theaters, various ambiances can be created. When the illumination has to change gradually from one color to another, a suitable color transition has to be defined. Currently, there is not much scientific knowledge on the discrimination and preference of temporal light effects. Since color spans a three dimensional space, there are innumerable ways to generate a transition between two colors. In applications where the light source is driven by an RGB signal, temporal light effects are usually generated by a linear interpolation in RGB space. However, since RGB is a device dependent color space, the actual color transition will depend on the physical characteristics of the light source. A second drawback is that RGB is not an appropriate space to describe color perception. Therefore, it is better to define color transitions in a device independent and perceptually uniform color space, such as $\boldsymbol{CIE\ L^*a^*b^*}$.

The choice of an appropriate perceptual color space does not necessarily lead to color transitions that are appreciated by human observers, as there are still many algorithms that seem to be reasonable. Therefore, this chapter aims at determining guidelines for the design of temporal transitions between two colors. In order to study what color transitions are preferred, one should know what color transitions can be distinguished. It makes no sense to ask people which of two temporal light effects they prefer, if they cannot see the difference. Therefore, two experiments are performed. The first experiment investigates when two different temporal transitions between two fixed colors are perceived as distinct. The results are used in the second experiment to develop six algorithms to create temporal color transitions. Participants evaluate these algorithms for two different applications: AmbiLight TV and atmosphere creation.

Although temporal properties of the human visual system have been subject of many psychophysical studies, only few studies bear reference to our research. First, studies on temporal contrast sensitivity have shown that the detection threshold for temporal sinusoidal fluctuations in luminance decreases with frequency up to about 10 Hz and then increases with frequency [23, 24]. Thresholds for temporal chromatic variations are constant for frequencies up to about 4 Hz and then increase with frequency [99]. If the mechanisms underlying the perception of gradually changing light patterns are similar to those of the perception of alternating light patterns, one would expect that the ability to discriminate between two temporal color transitions depends on duration for luminance variations but not for chromatic variations (at least for durations larger than $1/4$ s).

Second, studies on color discrimination have shown that the visual system is more sensitive to simultaneous spatial color variations compared to variations presented with a temporal delay. For example, the discrimination threshold for simultaneously viewed, closely juxtaposed colored patches is about two times smaller than the threshold for successively presented colors with a temporal delay of $200-550\,\text{ms}$ [103]. Therefore, discrimination thresholds for temporal color transitions are expected to be larger than those for simultaneously viewed color pairs, which are usually around $\Delta E_{ab} = 1$ [32].

Finally, Montag [72] showed that the color with the appearance halfway between two other colors in $\boldsymbol{CIE\ LCh_{a^*b^*}}$ color space that have equal chroma and a hue difference of about 50° did not have the same chroma as the color pair. The selected color was closer to the geometric midpoint in $\boldsymbol{CIE\ L^*a^*b^*}$ of the color pair. Hence, a linear transition between two colors in a perceptual color space might be a promising algorithm for color transitions.

7.2 Discrimination of temporal color transitions

This experiment aimed at determining how large the difference between two temporal color transitions should be before people can perceive the difference. Since the final aim was to obtain guidelines for the light effects of AmbiLight TV and atmosphere creation, it was decided to use a set-up that could be applied for both applications. A linear transition in $\boldsymbol{CIE\ L^*a^*b^*}$ was used as a reference stimulus and the just noticeable difference from the reference found using a tuning procedure.

7.2.1 Method

Set-up

A 42″ Wide Screen Plasma TV was positioned at a distance of 23 cm in front of a white wall in a room of 4 by 6 m. Eight ***RGB***LED-units were mounted at the back of the TV such that the LEDs were not directly visible but only the light reflected from the wall. The maximum luminance on the wall was $100\,\text{cd/m}^2$. A couch was placed at a distance of 4.5 m from the screen. The display and LEDs were controlled by a computer system. The chromaticity coordinates of the TV primaries were close to the EBU [80] primaries. The driving values for the LEDs and display were calculated such that the white point corresponded to CIE D65 and all colors were located within the gamut of the display. This was possible because the color gamuts of all LED-units were larger than that of the display. The refresh rate was 60 Hz for the TV and 50 Hz for the LEDs.

Stimuli

To determine the discrimination ability of temporal color transitions four *color pairs* defining the start and end color of the transition were used: “blue-green”, “green-magenta”, “magenta-green”, and “magenta-yellow”. The order of start and end color were reversed in the second and third color pair, to measure the effect of chromatic adaptation. The chromaticity and luminance of the colors were based on four images that were used in the second experiment (see Section 7.3.1 and Figure 7.3) and are presented in Table 7.1.

	Y	x	y	L	a	b
Blue	0.14	0.189	0.126	45	43	-74
Green	0.23	0.320	0.575	55	-50	55
Magenta	0.14	0.418	0.242	44	57	-9
Yellow	0.31	0.473	0.481	63	4	75

Table 7.1: Luminance and chromaticity of the four colors expressed in 1931 ***CIE xyY*** and ***CIE*** $\boldsymbol{L^*a^*b^*}$.

For each *color pair* several color transitions were generated using MatLab software. The reference transition was a linear interpolation between the start and end color in ***CIE*** $\boldsymbol{L^*a^*b^*}$. The test transitions were arcs defined in one of two planes:

1. the plane through start and end color parallel to the lightness axis, called the lightness-plane and
2. the plane through start and end color perpendicular to the first plane, called the chromaticity-plane (see Figure 7.1)

The arcs were defined by three points: the start color, the end color and a color in the corresponding plane at a distance d from the color halfway the start and end color. Colors of the test transitions located in the lightness-plane had equal hue and chroma compared to those of the reference transition. However, the lightness of the colors was larger (“L_+”) or smaller (“L_-”). Colors of test transitions located in the chromaticity-plane had the same lightness compared to those of the reference transition. However, the colors varied in hue and chroma. The arc could bend towards the center of the gamut (“C_{in}”) or towards the boundaries of the gamut (“C_{out}”). Figure 7.1 shows examples of the *directions* “C_{in}” and “C_{out}” for each color pair. The ***CIE*** $\boldsymbol{L^*a^*b^*}$ values of a color transition were transformed into ***RGB*** values for each LED-unit separately.

Each color transition was composed of three parts: first the start color was shown for 2 s, then the transition (0.5 s or 4 s) and finally the end color was shown for 2 s. The two *durations* of the color transitions are realistic speeds of fast and slow transition for both AmbiLight TV and atmosphere creation.

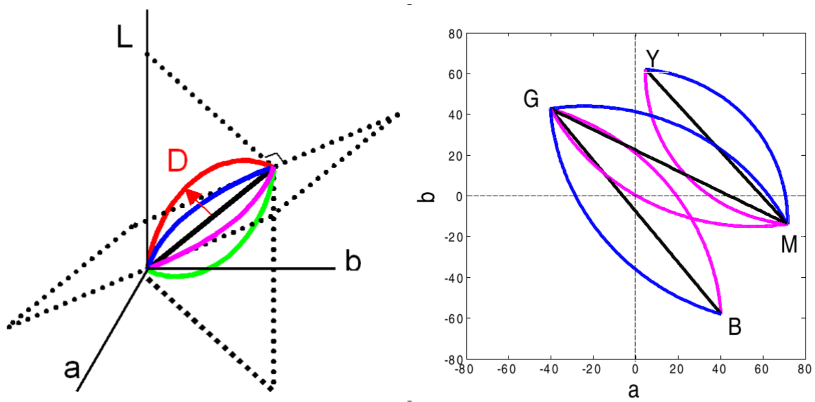

Figure 7.1: (a) Examples of the reference transition (black line) and the test transitions with *direction* "L_+" (red arc), "L_-" (green arc), "C_{in}" (magenta arc) and "C_{out}" (blue arc). (b) Projection of the reference transition and test transitions "C_{in}" and "C_{out}" on the a^*b^*-plane for each color pair.

Participants

Ten male and nine female participated in the experiment. Their age ranged between 20 and 46 years, with an average of 29 years. All participants had normal color vision, as measured with the Ishihara test of color deficiency [48].

Procedure

Participants were seated at the couch in front of the display. During the experiment the ambient illumination and the display were turned off. Two temporal color transitions were shown on the LEDs immediately after each other: first the reference and then a test transition. Both transitions had the same start and end color and the same duration. After the second transition, all LEDs emitted white light, in order to avoid continuous change of adaptation state.

Participants had to judge whether the two transitions were equal or not. By selecting a button on the screen of a notebook, participants could:

1. repeat the trial
2. increase the difference between the transitions
3. decrease the difference between the transitions
4. record the value and go to the next trial

Participants were instructed to find the point at which the difference between the two color transitions was just detectable. The test transition that was initially shown was clearly different from the reference one and had a difference of $d = 25\,\Delta E_{ab}$. The step size of the tuning procedure was initially $1\,\Delta E_{ab}$ and decreased to $0.5\,\Delta E_{ab}$ for test transition smaller than $d = 2\,\Delta E_{ab}$.

The experiment consisted of 32 conditions: 2 *durations*, 4 *color pairs* and 4 *directions* of the test transitions. Participants started with 6 practice trials to get used to the task before the 32 trials were presented. The experiment took about one and a half hour per participant.

7.2.2 Results

The maximum color difference d (in ΔE_{ab}) between the reference transition and the test transition that could just be distinguished is defined as the *discrimination threshold*. Figure 7.2 presents the thresholds as a function of the *direction* of the test transition (a) per *color pair* and (b) per *duration*.

An ANOVA was performed with *color pair* ("blue-green", "green-magenta", "magenta-green", and "magenta-yellow"), *direction* ("L_+", "L_-", "C_{in}", and "C_{out}") and *duration* ("0.5 s" and "4 s") as fixed factors and participant as random factor. A model consisting of all main effects and all second order interactions of the fixed factors was fitted. The analysis revealed a significant main effect of *color pair* ($F(3,527) = 5.18,\ p < 0.01,\ r^2 = 0.029$), *direction* ($F(3,527) = 45.55,\ p < 0.01,\ r^2 = 0.206$) and *duration* ($F(1,527) = 6.93,\ p < 0.01,\ r^2 = 0.013$) and a significant interaction effect between *color pair* and *direction* ($F(9,527) = 5.01,\ p < 0.01,\ r^2 = 0.079$) and between *duration* and *direction* ($F(3,527) = 23.80,\ p < 0.01,\ r^2 = 0.119$). The interaction between *color pair* and *duration* was not significant ($F(9,527) = 0.539,\ p = 0.66$).

The effect of *color pair* was caused by the "blue-green" transition, for which the threshold was on average slightly lower compared to the other color pairs. However, this was mainly due to the interaction between *color pair* and *direction*. The difference between *color pairs* was only significant for the *direction* "C_{in}" (see Figure 7.2).

The effect of *direction* was caused by a significant difference between transitions in the lightness-plane and transitions in the chromaticity-plane. Thresholds were on average lower for "L_+" and "L_-" compared to "C_{in}" and "C_{out}". However, the difference between *directions* depended on *duration*. At "0.5s" the difference between the two groups of transitions was large, but decreased for the slower transition.

Discrimination thresholds were on average smaller for color transitions of "0.5s" compared to "4s". However, as mentioned above, the effect of *duration* depended on the *direction* of the transition. Thresholds increased with *duration* for the transitions in the lightness-plane, and slightly *decreased* with duration for the transitions in the chromaticity-plane.

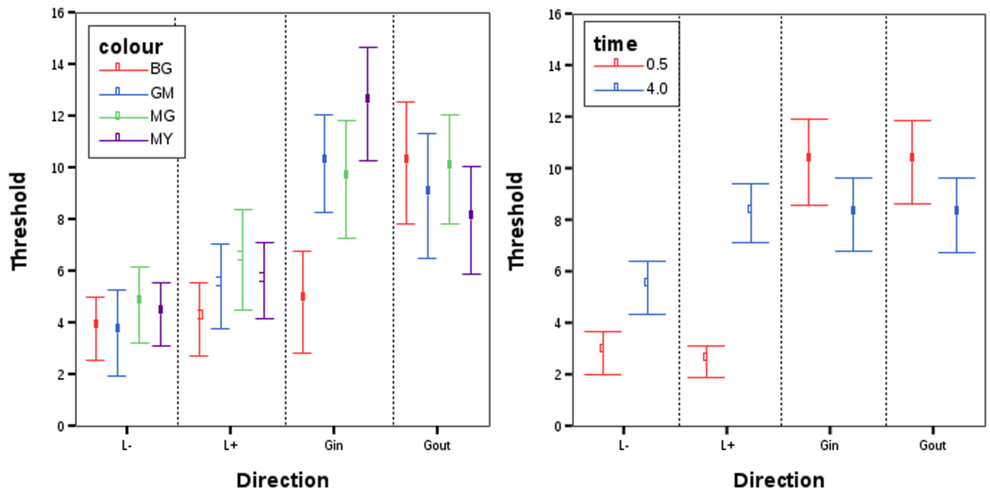

Figure 7.2: Discrimination threshold (in ΔE_{ab}) for the direction of the test transition (a) per *color pair* (averaged over the 2 *durations*) and (b) per *duration* of the transition (averaged over the 4 *color pairs*. The error bars correspond to the 95% confidence intervals.

7.3 Preference of temporal color transitions

This experiment aimed at finding the best way to make a temporal transition between two fixed colors. Since there are innumerable ways to create a transition between two colors in a 3 dimensional color space, it is hardly possible to find the most optimal solution. In addition, there might be a large range of transitions that are assessed as "very appealing". Therefore, it was decided to compare only a few possible solutions. Additionally, to test the possible influence of the context in which the transition is shown, the transitions were shown both with and without content on the TV.

7.3.1 Method

Stimuli

Two of the four *color pairs* of the first experiment were used: "blue-green" and "magenta-yellow". For each color pair, six color transitions in different *directions* were created: a linear transition in ***CIE L*a*b**** ("Lab"), a linear transition in ***RGB***

("rgb"), and the four transitions from the first experiment with a color difference ΔE_{ab} of about 3 times the average threshold value.

The experiment consisted of three *context conditions*:

1. the light transitions were shown while the display was turned off and there was "no content"
2. the light transitions were shown together with two images having "meaningful content"
3. the light transitions were shown together with two images with a "single color"

The first condition corresponds to an atmosphere creation application using only light sources, while the second and the third correspond to an AmbiLight TV application. In the conditions where images were also shown, two images with a duration of 4 s were shown immediately after each other (see Figure 7.3). During the first 2 s, the LEDs emitted the start color, then the color transition was shown for 4 s, and finally the end color was shown for 2 s. Hence, the transition started 2 s before the second image was shown. The two images that were shown had either meaningful content or were uniformly colored. The average chromaticity of the start and end images was always the same as those of the start and end colors of the LEDs.

Figure 7.3: The image stimuli used in the second experiment.

Participants

Fifteen male and fifteen female participated in the experiment. Their age ranged between 21 and 46 years, with an average of 27 years. All participants had normal color vision, as measured with the Ishihara test.

Procedure

Participants were seated at the couch in front of the display. During the experiment the ambient illumination was turned off. Two color transitions with the same start and end color were shown on the LEDs immediately after each other. After the second transition, all LEDs emitted white light. Participants were asked to indicate which of the two transitions they preferred most. They had the possibility to repeat the trial.

All participants were exposed to the first two *context conditions*, but only half of the participants performed the third condition. For each *condition* and each *color pair*, all possible combinations of the six different transitions were presented. This resulted in 30 trials per condition, which were presented in a random order. Half of the participants started with *condition* "no content" and half of them started with *condition* "meaningful content". *Condition* "single color" was always presented last for the people involved. Participants started with four practice trials before the first condition was presented. The experiment took about half an hour per condition per participant.

7.3.2 Results

The preference responses were analyzed using Thurstone's law of Comparative Judgment case V [101]. This analysis calculates a preference score (z-score) for each evaluated algorithm. The 95% confidence intervals on the difference between preference scores were calculated with the generalized linear model (GLM) procedure described in [81].

Figure 7.4 presents the obtained preference score of the six algorithms per *color pair*, taking into account the data of all participants and the three *conditions*. The difference in preference scores between two algorithms can be interpreted in terms of the percentage of participants that preferred one algorithm above the other, which can be calculated with the cumulative normalized probability function. To facilitate the interpretation, the least preferred algorithm was given a value of zero. The horizontal lines divide the algorithms into groups for which the preference scores are not significantly different $(p < 0.05)$. The GLM analysis revealed a significant effect of *color pair* $(p < 0.001)$. There was no significant difference between the conditions "no content", "meaningful content", and "single color" $(p = 0.24)$.

Figure 7.4 shows that the order of the algorithms was very similar for the two color pairs: "C_{in}" was the least preferred algorithm, whereas the algorithms "L_-", "Lab" and "RGB" were preferred by most of the participants. However, the difference between the least and most preferred algorithms was larger for "blue-green" compared to "magenta-yellow". When "C_{in}" was removed from the data, the effect of color pair disappeared. Apparently, this algorithm was differently evaluated for the two color pairs. This is consistent with the first experiment, where the discrimination threshold was found to depend on color pair only for "C_{in}".

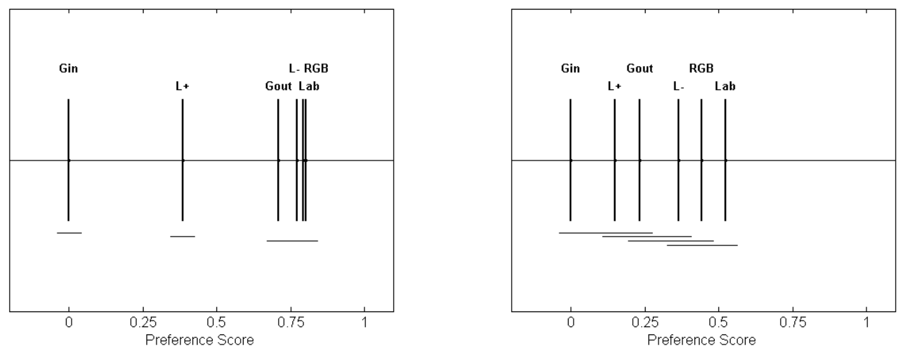

Figure 7.4: Preference scores of the six algorithms for a temporal color transition between (a) "blue-green" and (b) "magenta-yellow". According to Thurstone's model, the percentage of participants that prefer one algorithm above the other is 50%, 69% or 84% for a difference in preference score of 0, 0.5 and 1, respectively.

7.4 Discussion

The ability to discriminate between two temporal color transition was found to be similar for all color pairs, except for the *direction* "C_{in}", where thresholds were lower for the "blue-green" transition. This result might be related to the observation that two colors in different linguistic color categories are more easily distinguished than those in the same category [104]. Indeed, the intermediate color for the blue-green transition with *direction* "C_{in}" has a magenta hue and does not fall into the same color category as blue or green in contrast to the direction "C_{out}", for which the intermediate color has a cyan hue. A second explanation might be that cyan is located between green and blue on the color circle, whereas magenta is located on the opposite site. Therefore, the transition via magenta might be perceived as unnatural and, hence, more salient. Since there was no difference between the *color pairs* "green-magenta" and "magenta-green", it can be concluded that chromatic adaptation of the eye did not play a role in the ability to discriminate between temporal color transitions.

Discrimination thresholds were found to increase with *duration* for lightness variations and to be rather constant for chromatic variations. This result confirms the expectation stated in the introduction, which was based on research on temporal contrast sensitivity [100] and supports the results from Chapter 6. In addition, thresholds for temporal variations in lightness were lower compared to temporal variations in chromaticity. This again supports the intrinsic difference between chromatic and lightness variations as described in Chapter6. Additionally, the perceived size of the light spot on the wall increased with increasing lightness, as is the case for many light sources, which could have made transitions in the lightness-plane even more visible.

Discrimination thresholds ranged between 2.5 and 10.5 ΔE_{ab}. Hence, in the most critical situation, thresholds were comparable to those for spatially separated color patches, which are usually around 1 ΔE_{ab} [32]. However, when the duration increased or the direction of the transition changed, thresholds for temporal color differences were considerably larger compared to spatial color differences. This is consistent with literature on the perception of temporal color differences [103].

The most preferred way to make a temporal transition between two colors was found to be independent on the application, i.e. whether the light effects had to enhance the experience of watching TV or whether they were used for atmosphere creation. A linear transition in ***CIE L*a*b**** was evaluated as one of the best algorithms, which is in agreement with the results of Montag [72]. Interestingly enough, this algorithm was evaluated to be as appealing as a linear transition in ***RGB***. For "magenta-yellow" the maximum color difference between these algorithms was 6 ΔE_{ab} and, hence, hardly visible. However, for "blue-green" the difference of 20 ΔE_{ab} was clearly visible.

The results of this study suggest that it is possible to design a general algorithm for temporal color transitions that are appreciated by human observers, independent of color pair and application. However, additional research is needed including more color pairs to confirm this statement.

7.5 Conclusions

This chapter investigates what is perceptually the most optimal way to create a temporal color transition between two colors. The first experiment measured the ability to distinguish between two temporal color transitions. The reference transition was a linear interpolation between two colors in ***CIE L*a*b****, the test transitions were arcs defined in different planes going through the linear transition. Discrimination thresholds ranged between 2.5 and 10.5 ΔE_{ab}, dependent on the *color pair*, *direction* and *duration* of the transition. In the second experiment, several perceptually different color transitions were evaluated. The most preferred transitions were a linear transition in ***CIE L*a*b**** and a linear transition in ***RGB***. This suggests that appealing temporal color transitions can be created without complicated calculations.

8

Flicker perception in the periphery under mental load

"Maybe your mind is playing tricks
You sense, and suddenly eyes fix
On dancing shadows from behind"

Iron Maiden, Fear of the Dark

Light is seldom the focus of attention in everyday life and usually supports the main activity. This is also largely the case in the application examples in the previous chapters. The previous two chapters studied the temporal sensitivity of the human visual system using complex stimuli that were always presented in the center of the visual field. Due to the supporting nature of lighting, it is also important to study the influence of the position of temporal transitions in the visual field. The first experiment in this chapter studies the visibility of flicker at different eccentricities at a number of base points and in a number of directions in a color space. To ensure that the modulated light remains in peripheral vision, a visual task in central vision was introduced. The second experiment was designed to test the effect of the task on the visibility of the peripheral temporal effects. This chapter is based on [79].

8.1 Introduction

In recent years LEDs have become more frequently used due to several advantages of such systems over conventional lighting systems. The dynamic capabilities of LEDs allow their use in many applications, which range from common indicators and road signs to personalized atmosphere creation. The main challenge in the design and control of dynamic lighting systems is the perceived attractiveness of the temporal light changes they produce. LED lamps are typically controlled in a way that produces chromaticity and brightness changes in discrete steps at discrete time intervals. When the size of these steps is too large the light change is perceived as jerky, stepped, or unsmooth. Flicker, which is defined as the rapid alternation between two colors or brightness levels, is one example of such effects, and from the perspective of perceived attractiveness, it is an especially annoying one. Under some conditions, for example for very high frequencies (above 80 Hz), flicker is not perceived. Instead, a temporal fusion, which means a steady, continuous light is perceived [111]. Previous studies demonstrated that flicker perception is influenced not only by frequency, but by numerous other factors, which include luminance [43], chromaticity irrespective of brightness[56], size of the stimulus [45] and finally the retinal position of the stimulus[102].

In order to be able to define a distance between a pair of color stimuli defined within a color space, at which alternating light produces smooth and not flickering light, the notion of visibility threshold was introduced [69]. In Chapter 6 it is defined as the largest amplitude for the particular stimuli that produces smooth and not flickering light changes. Spatial just noticeable differences are described by, among others, the ***CIE Luv*** and ***CIE*** $\boldsymbol{L^*a^*b^*}$ color spaces, which demonstrate relatively uniform color differences predictions for spatial vision. However color spaces in the temporal domain are not defined yet.

Recent studies, also part of this thesis, have shown and roughly defined the flicker sensitivity in the central vision; i.e. when fixated directly on the light source. However, LED based lighting systems can also be extended to the peripheral vision, which comprises the region beyond the very center of the eyes' fixation to the limit of about a 100° to each side. It is known that, due to the structure of the human eye, the perception of motion and color changes at the periphery. For example, Abramov et al.[2] showed that the photopic luminosity function (the average visual sensitivity of the human eye to light of different wavelengths), as defined by heterochromatic flicker photometry, is not the same in the periphery of the retina as it is in the fovea. For that reason one particular alternating color pattern that appears to be smooth in central vision may be perceived as flickering in the periphery, and vice versa. Thus, the results from Chapter 6 cannot be assumed in the periphery and a new study is required. The hypotheses formulated in the current research are largely based on the findings in Chapter 6.

It is commonly known that the density of cones in the human retina decreases when moving away from the center of the retina, and therefore color perception is impaired significantly as well. On the other hand, Hansen et al.[42] found that humans are able to perceive some colors at large eccentricities (more than 50°) and cone opponent channels, however sparse, are still present at these angles. Therefore, it is hypothesized that there is an effect of the eccentricity on the lightness, chroma and hue flicker visibility thresholds. Even though chroma and hue flicker visibility should be diminished, it is expected that chromatic flicker is still perceivable at large eccentricities. In Chapter 6 it was demonstrated that lightness changes are more visible than changes in chroma and hue, when the color changes are expressed in the ***CIE*** $\boldsymbol{L^*a^*b^*}$ color space. The experiments in this study are designed to check if this fact is valid for peripheral vision as well. In line with previous findings it is also hypothesized that there exists a frequency for which the visibility threshold has a minimum value.

Lighting is sometimes the main object of direct attention but most frequently it facilitates other activities, which are the main tasks. Carmel et al. [15] demonstrated that including a task and increasing the mental load impairs flicker perception. It happens due to the temporal resolution of conscious perception being determined by the attention availability. Similarly, it is hypothesized that there is an effect of mental load on the flicker sensitivity. Flicker visibility thresholds are expected to be smaller without mental load task [15].

In order to gain a better understanding of the way in which the human visual system processes temporal patterns in the peripheral vision as a consequence of dynamic lighting, two experiments were conducted, the results of which are presented in this chapter.

8.2 Flicker sensitivity thresholds at different eccentricities

The first experiment was designed to measure human sensitivity to chromatic flicker at different visual angles. Flicker stimuli were created by choosing pairs of color in ***CIE*** $\boldsymbol{LCh_{a^*b^*}}$ color space and alternating between them. These pairs were selected along each of the $\boldsymbol{L}$, $\boldsymbol{C}$, and $\boldsymbol{h}$ axes, and thresholds were expressed in ΔE_{ab}, the Euclidean distance in the ***CIE*** $\boldsymbol{L^*a^*b^*}$ color space.

Given the results of previous research on related topics discussed above, the effect of *eccentricity*, *frequency* of the change, *base-color* point and *direction* of change on the visibility threshold was investigated.

8.2.1 Method

Equipment

As a light source, a LED luminaire was used, with three ***RGB*** LEDs at the top and three ***RGB*** LEDs at the bottom of a cylindrical diffuser. Another, flat, diffuser was mounted 20 cm in front of the lamp to reduce sharp edges that may be detected during saccadic eye movements. The light source was 23 cm wide, or 5.3° of visual angle at a distance of 2.5 m. Maximum luminance of the LED luminaire was $400\,\mathrm{cd/m^2}$. The detailed profile of the light source used is presented in Figure 8.1

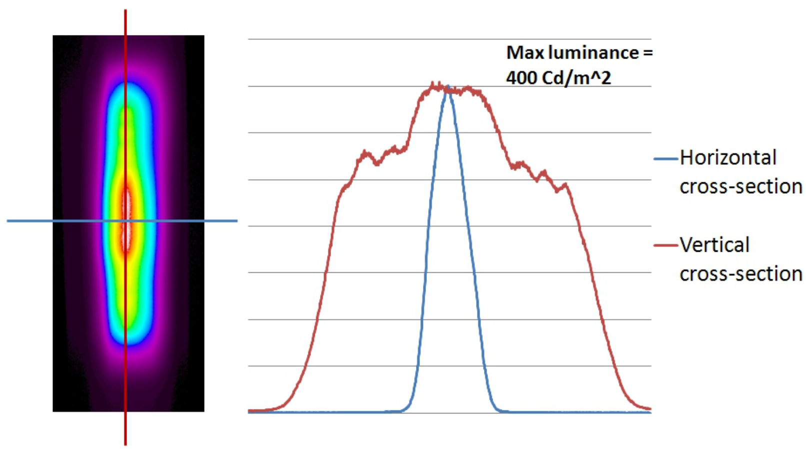

Figure 8.1: False color image showing the luminance distribution along the cross sections

The LEDs were driven using pulse width modulation (PWM) with a driving frequency of 500 Hz and 11 bit levels. The driver accepted ***RGB*** values in the device color space of the LEDs. The target stimuli were defined in $\boldsymbol{CIE\ LCh_{a^*b^*}}$ and transformed via ***CIE XYZ*** to the ***RGB*** device color space using a computer program running on a PC connected to the light units.

In the transformation, the CIE D65 white point was used, being the white point of the display used for the mental task. Care was taken that the requested colors were within the device gamut of the light source.

Stimuli

The stimuli were square-wave discrete modulations (flicker), alternating between $\boldsymbol{CIE\ LCh}_{base} - \frac{S}{2}$ and $\boldsymbol{CIE\ LCh}_{base} + \frac{S}{2}$ every $\frac{1}{f}$ seconds. $\boldsymbol{CIE\ LCh}_{base}$ denotes the *base point* in the $\boldsymbol{CIE\ LCh_{a^*b^*}}$ color space and S is the amplitude of the flicker in one of the *directions* $d \in \boldsymbol{L}, \boldsymbol{C}, \boldsymbol{h}$, and f is the *frequency*. Example stimulus for a flicker change in *direction* d is depicted in Figure 8.2. Each stimulus consisted of 4 s with amplitude S, followed by between 2 and 4 s with amplitude 0. In this way participants had a reference to a non-flickering stimulus, but could not predict the exact time the modulated light conditions started.

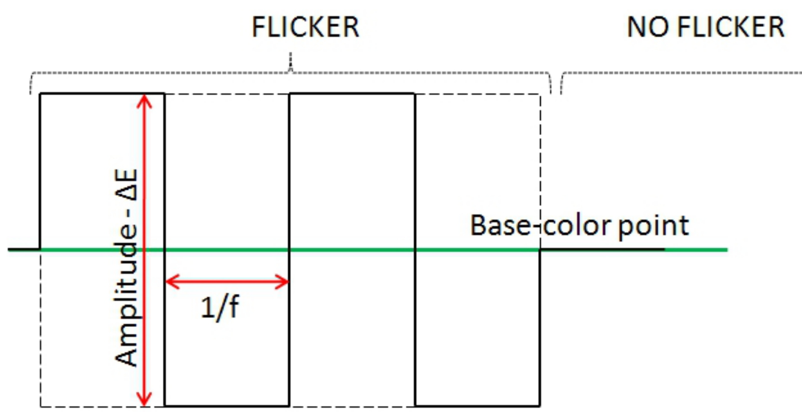

Figure 8.2: Example stimulus used in the experiment

The stimuli were varied in *eccentricity* ("35°", "60°" and "90°"), *frequency* ("5", "10", "20", "40" and "60 Hz"), *base-color point* ("Red", "Green" and "Blue"), and *direction* of change ("Lightness", "Chroma" and "hue"). In the choice of the *eccentricities* various criteria were considered. 35° was chosen because it corresponds to the retinal region of highest flicker sensitivity [102]. 90° was chosen as it matches the most distant retinal region, at which human vision still functions. Finally, 60° was chosen as a middle value between 35° and 90°. *Base-color points* had the following ***CIE LCh***$_{a^*b^*}$ values: "red" $(60, 60, 45)$, "green" $(60, 60, 150)$, and "blue" $(60, 60, 290)$. The three *base-color* points were chosen because these are the desaturated primaries for which the change within ***CIE LCh***$_{a^*b^*}$ allows for large amplitudes along each of the axes. For each *direction* of change different amplitudes of flicker were included in the experiment. A series of pilot tests were conducted to adjust the amplitudes, in order to cover a range around the expected visibility threshold. As such, lightness flicker was implemented at amplitudes of 1,2,4,6,10, and 60 ΔE_{ab}, Chroma flicker at amplitudes of 10, 20, 40, 60, and 80 ΔE_{ab}, and hue flicker at amplitudes of 24, 40, 60, and 80 ΔE_{ab}.

The experiment consisted of three sessions, each of them at different *eccentricity*. The stimuli were presented in three blocks, each representing different *base-color point*, randomized across the subjects. The total number of stimuli including the different levels of amplitude S was 675 (i.e. 225 for each *eccentricity*).

The mental load task was a simple game created in Adobe Flash. It was displayed on a screen positioned 2.5 m in front of the participants, who controlled it with a computer mouse placed on a table in front of them. Participants had to move one of four balls in such a way that it did not touch/hit the remaining three balls which moved according to general physics laws. The ball controlled by the participant was colored middle gray, while the other three balls were black. When the player ball touched any of the other balls, the red ”HIT” text appeared in the left corner of the screen. This specific mental load task was chosen because it forces the participant to fixate on the screen, but still allows for flicker detection. Moreover, it enforces two types of eye movements: saccadic from one ball to the other and continuous ones following a single ball.

Participants

Participants were 10 female and 20 male Philips Research employees, aged between 23 and 45 years, with a mean age of 29 years. All subjects had normal or corrected to normal vision. They did not wear glasses (as the flickering light in the distant periphery could be perceived reflected from the glasses). They all had normal color vision, tested with Ishihara Test of Color Deficiency. Before the start of the experiment, participants were asked if they, or a member of their family, suffered from epileptic attacks, migraines or any other known condition for which flickering light could have negative consequences.

Procedure

In an otherwise darkened room, the LED luminaire was placed at different eccentric positions and it faced the participant. Subjects were seated at a distance of 2.5 m from the light source. To ensure fixed position of subjects' eyes and the constant distant to the light source, their chin was placed on the chin rest mounted on the table in front of them. They were directly facing the display with the mental load task.

Before the start of the session the participants received verbal instructions. They were instructed to fixate on the display and carry on the mental load task for the duration of the entire session. They were asked to press the space bar on the keyboard placed on the table in front of them every time they saw the flickering. A brief trial session was implemented so that the participants got used to the stimuli before starting the experiment.

8.2.2 Results

In order to obtain the visibility threshold, a psychometric curve was estimated for every stimulus by fitting a logistic curve to the percentage of people indicating flicker at each amplitude. The visibility threshold was then defined as the amplitude at which the psychometric curve reached 0.5, representing the amplitude at which 50% of the participants detected flicker. To compute 95% confidence intervals for the thresholds, a bootstrap procedure [27] was used. The number of resamples was 1000.

The median threshold amplitudes were analyzed using ANOVA (Analysis of Variance) with *eccentricity*, *base-color point*, *direction* of change and *frequency* as fixed factors. Based on the distribution of the median threshold amplitudes as computed by the bootstrap procedure, a new set of 30 threshold values for each factor combination was created and used in the fit of the model. This was done to add the variance information into the ANOVA procedure. The model fitted consisted of all main effects and all two way interaction effects. In several high-eccentricity, high-frequency cases, reliable thresholds were not found. These points are absent from the plots.

A significant main effect was found for all fixed factors, namely *eccentricity* ($F(2,40) = 46.151,\ p < 0.001,\ r^2 = 0.775$), *base-color point* ($F(2,40) = 116.813,\ p < 0.001,\ r^2 = 0.891$), *frequency* ($F(3,40) = 118.199,\ p < 0.001,\ r^2 = 0.892$), and *direction* ($F(2,40) = 456.459,\ p < 0.001,\ r^2 = 0.967$). A post-hoc test on the factor *direction* revealed that the "lightness" visibility threshold is significantly lower then the "chroma" and "hue" visibility thresholds. This is in line with the findings in Chapter 6 for central vision.

All interaction effects with *frequency* were found significant, namely with *eccentricity* ($F(6,40) = 9.336,\ p < 0.001,\ r^2 = 0.466$) , *base-color point* ($F(6,40) = 145.335,\ p < 0.001,\ r^2 = 0.91$), and *direction* ($F(6,40) = 42.453,\ p < 0.001,\ r^2 = 0.761$). Additionally, a significant interaction was found between *eccentricity* and *direction* ($F(4,40) = 12.490,\ p < 0.001,\ r^2 = 0.526$). Due to the large main effect of *direction* and the interaction effects, the data was subdivided and three separate models for the three directions were fitted. The most important findings from the analysis of the separate models are given below.

Effect of eccentricity on lightness flicker

Eccentricity was found to have a significant effect on peripheral "lightness" flicker perception ($F(2,11) = 450.473,\ p < 0.001,\ r^2 = 0.968$), as shown in Figure 8.3. A post-hoc test revealed that the visibility threshold for 90° is significantly different from both 35° and 60°. Additionally, a significant two way interaction effect between *eccentricity* and *frequency* was found ($F(6,11) = 397.899,\ p < 0.001,\ r^2 = 0.964$).

Effect of eccentricity on chroma and hue flicker

Eccentricity was found to have a significant effect on "chroma" ($F(2,4) = 38.899,\ p < 0.001$) and "hue" ($F(2,5) = 855.525,\ p < 0.001$) flicker perception, as illustrated in

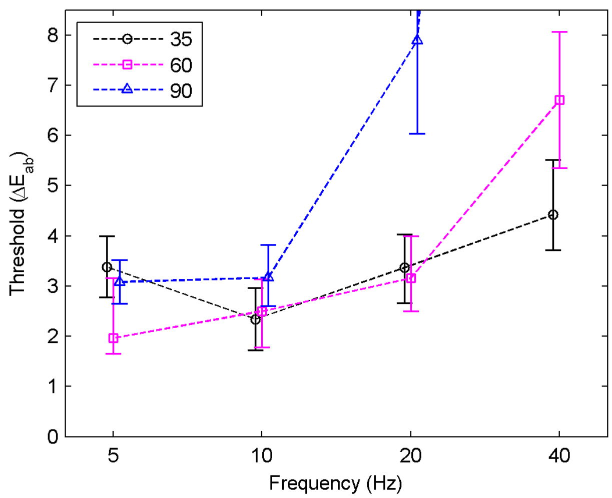

Figure 8.3: Lightness flicker thresholds averaged across *base-color points* for the 3 *eccentricities* tested. For the 90° and 20 Hz, the upper error bar reaches 13 ΔE_{ab}

Figures 8.4 and 8.5. Post-hoc tests revealed that all of the tested visual angles significantly differ from each other; the further into the periphery, the larger the visibility threshold. This is in line with what is expected based on the retinal structure and the study of Hansen et al. (2009) [42].

Independence of the thresholds for lightness flicker and dependence of the thresholds for chroma and hue flicker on the chromaticity of the base-color point

No main effect of *base-color* point ("Red", "Green", and "Blue") was found for peripheral lightness flicker. Contrary to "lightness" flicker, it was found that *base-color* point has a significant effect on "hue" $(F(2,5) = 3238.62,\ p < 0.001)$ and "chroma" $(F(2,4) = 1778.05,\ p < 0.001)$ peripheral flicker. For "chroma" flicker, it was found that "Green" has higher thresholds than "Blue" and "Red". In fact, visibility thresholds for "Green" are so high for all of the eccentric angles, that they can be neglected. In case of "hue" flicker, all of the base color points were found to significantly differ from one another. Post-hoc test revealed that humans are most sensitive to "Blue" and least sensitive to "Red" hue flicker.

8.3 Effect of mental load on flicker perception

One of the assumptions in the first experiment was that the introduction of an additional task changes the sensitivity to peripheral flicker. Experiment 2 was designed to test this

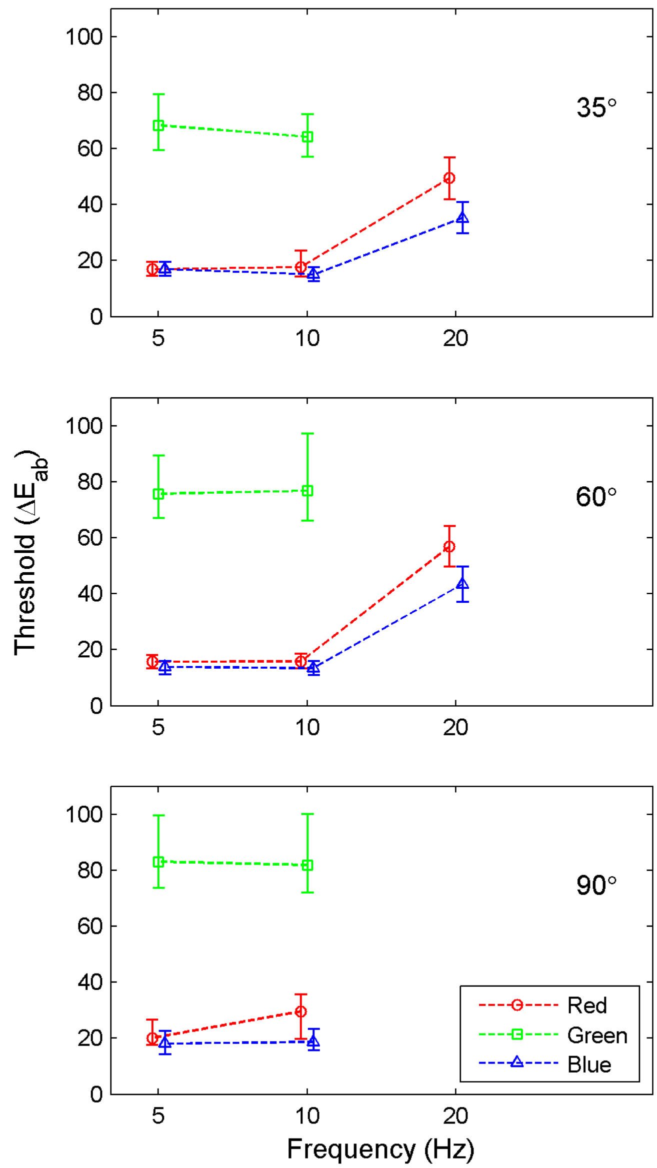

Figure 8.4: Thresholds for “Chroma” changes expressed in ΔE_{ab} as a function of frequency and base color point for different eccentricities

assumption and instead of the mental load task a fixation cross was displayed on the screen. To limit the duration of the experiment, the results for only one visual angle

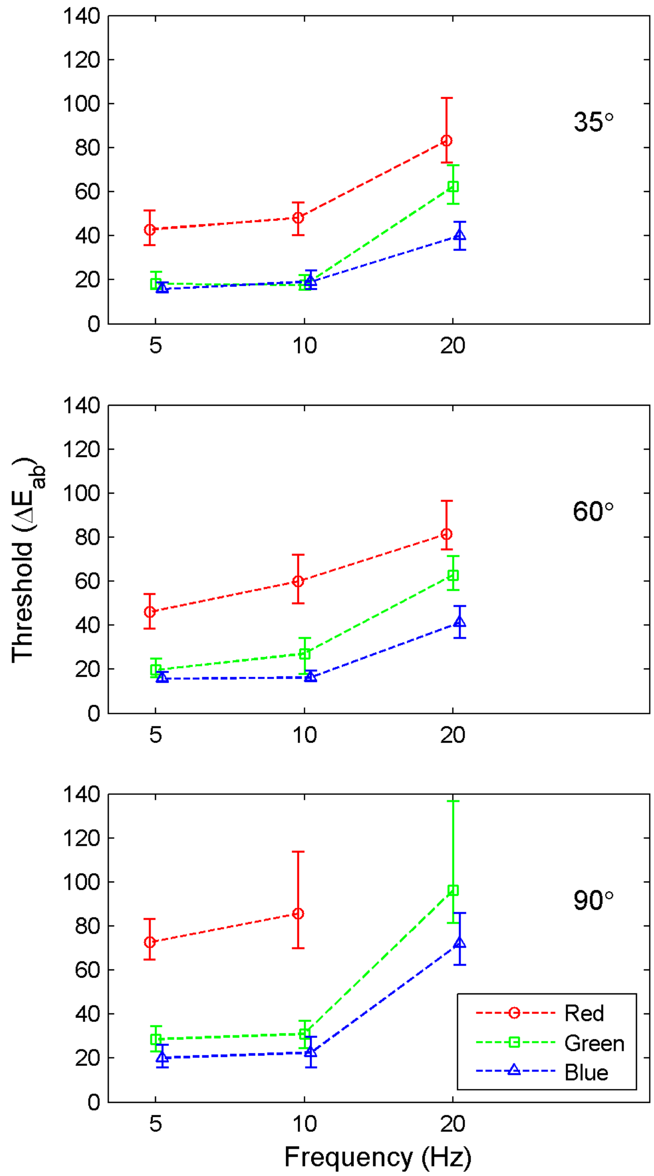

Figure 8.5: Thresholds for hue changes expressed in ΔE_{ab} as a function of frequency and base color point for different eccentricities

with and without a task were compared. The 35° visual angle was selected because the thresholds for 35° were the most accurate.

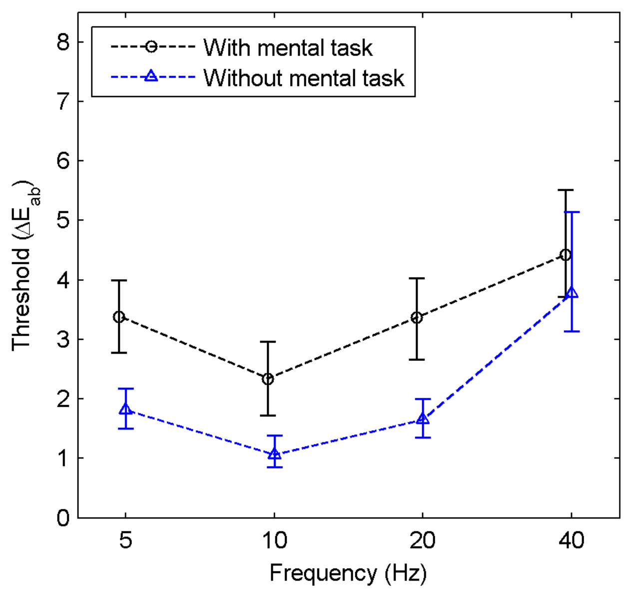

Figure 8.6: Effect of mental task on Lightness flicker perception

8.3.1 Method

Participants

Seven female and thirteen male Philips Research employees, aged between 23 and 41 years, with a mean age of 27 years, participated in the experiment. They fulfilled the same conditions as in the first experiment. The two sets of participants had partial overlap.

Stimuli

A subset of the stimuli from the first experiment was used. Only the 35° visual angle was used as well as only two *directions*, "lightness" and "chroma". The same *frequencies* and *base-color points* were used.

Procedure

The same experimental setup and procedure as in the first experiment were used.

8.3.2 Results

Effect of mental load on lightness and chroma flicker

Two ANOVA models for the two *directions* were fitted separately with *mental load*, *frequency*, and *base-color point* as fixed factors. All main effects and two-way interactions were included in the models. It was found that *mental load* significantly impairs flicker perception for both “lightness” $(F(1,6) = 341.347,\ p < 0.001)$ and “chroma” $(F(1,3) = 187.640,\ p < 0.01)$. Thresholds for “lightness” flicker with mental load were 2.09 times higher than without mental load for 5, 10, and 20Hz, and for 40Hz this factor was 1.2. Figure 8.6 depicts the effect of the *mental load* on “lightness” flicker perception. The main effect of the *base-color point* was found to be significant $(F(2,6) = 8.919,\ p < 0.05)$. In case of chroma flicker, the performance was impaired under mental load mostly for “Green”. The mental load used in the current study was relatively low. If the mental load is increased it is expected that flicker thresholds would increase even further [15].

8.4 Conclusions

Results from two experiments were presented. They support most of the previous findings on flicker visibility thresholds and demonstrate new effects relevant to the control of dynamic, LEDs based systems. It was shown that when using ***CIE LCh***$_{a^*b^*}$ color space, which is almost uniform for spatial differences, the peripheral flicker sensitivity is much higher when the change is in lightness compared to chroma or hue. Furthermore, chromaticity independence of lightness flicker sensitivity was shown, contrary to the chroma and hue flicker sensitivity. Chroma and hue flicker gets impaired with increasing eccentricity. Lightness flicker sensitivity increases for 5 Hz and 10 Hz frequencies but decreases for higher frequencies when eccentricity increases. Finally, mental load was verified to have an impairing effect on flicker perception in the periphery for both lightness and chroma flicker.

9

Conclusions and future work

In this thesis six studies on selected topics in the area of Ambient Intelligent Lighting (AIL) are presented.

Solid state lighting systems are revolutionizing the lighting industry. Beyond their traditional advantages of long lifetime, high efficiency, and environmentally friendly materials, solid state lighting systems enable the creation of spatially, spectrally and temporally complex light effects, with the sources of light embedded in the environment. With the newly found freedom the users can switch their focus from the confines of the technology to the expression of their needs, regardless of the technology. Identifying those needs, creating an effective language to communicate them to the system, and translating them to control signals that will fulfill them is the core differentiator of Ambient Intelligent Lighting (AIL) systems.

To facilitate the development of AIL systems, it is useful to treat solid state lighting as a digital medium. As with any other medium, the basic challenges can be subdivided in: creation and capture, storage and communication, reproduction, and measuring and controlling the quality of the content. Due to the fact that AIL is a broad and emerging area, the two basic challenges of new content creation and quality control were selected for further exploration in this thesis.

9.1 Light effect creation

Creation of new content for an AIL system is the problem studied in the first part of the thesis. The related studies are based on the insight that the traditional control driven creation is unsuitable for solid state lighting systems due to their increased capabilities and complexity. Instead of the traditional system centric *control driven* approach to lighting design, Chapters 3, 4, and 5 present examples of user centric *effect driven* light design with increasing complexity. The path to full spatio-temporal light effects is subdivided in its basic components, the selection of colors, the control of the spatial distribution of colors and the addition of temporal effects.

In the introduction, the AIL model is introduced as a way to select the language, or level of abstraction, that users can facilitate to communicate their desired effect to the system. The requirements for the solutions presented in the first part of the thesis were that the selected interaction language should be understandable and natural to the users but also translatable to device settings to render the desired effect. The solutions presented should be seen only as examples of usage of the basic features/ideas in every chapter. When developing AIL systems, any subset of these features can be used, independent of the chapter they are presented in. An example of such a system is sketched at the end of the section.

As a first step towards creating complex light effects in the environment, in Chapter 3 the problem of selecting a set of colors is studied. In this chapter an unsupervised method to relate terms with relevant images and colors is presented. The results show that meaningful colors for a term can be computed using large image repositories and image processing. Furthermore, a measure of suitability of a term for color extraction based on KL divergence is presented. The method is tested in two ways. First, using color names as terms enables the comparison of the computed set of colors to hue angle boundaries for high chroma colors based on psychophysical research. Second, to extend the evaluation to more general terms, a comparison between the computed set of colors based on equivalent terms in two different languages, English and Finnish, is used.

A prototype system based on the method is presented where the method is applied to song lyrics. In order to identify terms within the text that may be associated with images and colors, noun phrases are selected using a Part of Speech tagger. Large image repositories are queried with these terms and the results are gathered in a set of images per term. The images that are best ranked by the search engine are displayed on a screen, while the extracted representative colors are rendered on controllable lighting devices in the living room, synchronized with the song, resulting in a full multimedia experience.

The basic ideas in Chapter 3 are the usage of an auxiliary medium to define light effects and the use of simple statistics on large noisy datasets. The example auxiliary media used in the chapter are written language and digital images. Even though they are used together in the chapter, it is easy to imagine systems that use only one of them. For example, a pre-made database of colors and terms can be used as a basic implementation of the feature of selecting a set of colors based on a term. Similarly, any set of images, for example a user provided photo album or even a single image can be used to extract a set of representative colors. Furthermore, the feature extracted is not limited to a set of colors. The spatial distribution of those colors, as an example, can also be extracted from the set of images, as well as the temporal changes in case of digital video.

Simple statistics on large noisy datasets produced by querying Internet search engines are used as a way to capture knowledge. This, contrary to static knowledge representations, enables the system to be agnostic to the language of the user and responsive to changes of meaning of a term over time. The language independence is not used in the example application as the English language is required due to the Part of Speech tagger. This can be, however, easily generalized using the measure of suitability presented. Instead of the terms that should carry the strongest links to color information, as is the case when using noun phrases and proper nouns, terms with the highest computed suitability measure can be selected.

Once the desired set of colors for a light effect is selected, the user needs to define the spatial arrangement of these colors. In Chapter 4, an interactive lighting design and control approach is introduced. It is based on large LED-based lighting systems that use visible light communication. The visible light communication technology enables the on-line estimation of individual illumination contributions, using invisible light source identifiers. Algorithms that allow light rendering for single and multiple target positions in such a system are presented. These algorithms enable the control of chromaticity, luminance, and the spatial distribution.

The method was implemented in two test systems. A limited physical system was created to demonstrate the feasibility of the proposed method. A simulation of a larger system was used to demonstrate the usability and the improved user interaction of the proposed method.

The technological base of the solution in this chapter is visible light communication. The defining feature of the chapter, real time interactive light design, however, does not depend on the availability of visible light communication hardware. An alternative way of estimating individual illumination contributions, is for example dark room calibration. In this case, instead of moving through the space, the points of interest are selected on a screen. Using smart-phone or tablet cameras, this alternative implementation emerges as being simpler to implement and, arguably, easier to control.

Light in nature is always dynamic at some time scale, indicating the need of introducing dynamic light in artificial light systems. In Chapter 5 a new method of dynamic light effect generation using stochastic models is presented. Similar to dynamic lighting scenes in nature, the resulting light effects are unpredictable, yet recognizable. A method to learn the stochastic models from a video source of a natural scene is shown. The method extracts the representative colors from the video and subsequently learns the typical transitions between the colors. After the model has been learned, the rendering of the effects has low memory and processing requirements, making it suitable for implementation even on embedded platforms. The recognition of the produced light effects was tested for three automatically created models and a hand crafted one using a large user base. The results show the suitability of the method for dynamic atmosphere creation, but also a high appreciation of the produced light effects. Finally, possible extensions of the light effect to multiple dependent light sources and to interactivity are proposed.

The basic distinguishing feature in Chapter 5 is the introduction of stochastic behavior in the lighting system. In the solution presented the feature is coupled to another feature also used in Chapter 3, the translation of information from a different medium to light. The addition of stochastic behavior to lighting systems can be used in a much broader context. Some possibilities, like spatio-temporal effects and interaction are discussed in the possible extensions part of the chapter. Stochastic behavior can be also be used for static light effects. Research in the area of atmosphere perception shows that the set of device settings that produce a desired atmosphere does not consist of a single point, but rather is a subspace in the parameter space of device settings. This means that by a random sampling of the solution subspace, the system can produce the same desired atmosphere that looks different per device, adding unexpected and organic behavior to the system.

One advantage of doing a thesis in an industrial environment is the ability to transfer the results of research into products easier and faster. An example of this is the Philips Naturelle Candlelights product, which directly use the results of Chapter 5. Due to its simplicity and the fact that the separate lights are independent and self controlled, this solution is easy to implement. The implementation of the results of the other chapters usually needs additional hardware and software. This disadvantage is recently significantly alleviated due to the soaring use of smart-phones and mobile applications. The always on, fully connected and camera enabled phones and tablets bring forward the possibility of easy control, creation and sharing of light effects. Recently introduced Internet/light connection bridges enable direct control of light sources and open the possibility to create all the applications envisioned in this thesis and much more.

As an example of a mobile application, we can envision the use of interactive lighting design based on dark room calibration. The calibration is simply done by using the camera of the mobile phone and automatically switching on individual lights in the system and capturing their effect on the environment. In a manner even closer to painting with light, the user can select colors from a palette made by using algorithms from Chapter 3 and paint their spatial distribution directly on the images of the room. Using the algorithms in Chapter 5, we can go one step further and add dynamic brushes, learned from a video found on-line or captured with the mobile camera. Using the envisioned system, adding a reflection of a fireplace on a wall in the room would be as simple as touching the image of the wall on the mobile device.

Even though this thesis gives examples of abstractions that can be used to create AIL content, the basic question of the most suitable interaction language remains open. The specifics of lighting and the complications of working with unrelated colors increase the challenge. Most of the models used in contemporary color science are built with imaging in mind and treat illumination like a constant, taken from a limited set of standard possibilities. Contrary to these, Color Appearance Models, equivalent Lighting Appearance Models should account for temporal changes in illumination and both global and local adaptation in spatially complex illumination conditions. As the appearance of the illumination is inseparable from the illuminated environment, defining a standard environment to test the models that can be generalized to an arbitrary environment further complicates the development of such models. These facts, together with the limited capabilities of traditional lighting sources, result in the development of Lighting Appearance Models still being in its infancy. The lack of basic model makes it hard to use abstractions even at a level of *perceptual effects*, let alone at the *overall atmosphere* level. Looking forward, defining the set of light effect primitives and the hardware building blocks to render them is and will be one of the main challenges in the development of AIL systems.

9.2 Temporal quality of light

The second part of the thesis concerns the perceived quality of the produced light effects. From the multitude of light quality parameters discussed in the introduction, the temporal light quality was selected for further exploration due to the particularities of LEDs.

As mentioned before, the temporal response of LEDs to changes in driving current is much faster than the response of classical light sources. This fact produces new temporal artifacts, but also explains the lack of models to predict the visibility and acceptability of such artifacts. The emphasis of the chapters in the second part is on two temporal artifacts, luminance and chromatic flicker and perceived unsmooth light transitions. At a high level the chapters answer the questions: “How fast can a transition between two colors be?”; “What path through color space constitutes a pleasant temporal transition?”; and “How does the sensitivity to temporal artifacts change over visual angle and with the addition of a mental task?”

Chapter 6 explores the sensitivity of the human visual system to unsmooth light transitions. The results show the relation between the perception of unsmooth light transitions and the perception of flicker and give a contribution towards building a model of visibility of unsmooth lighting transitions and chromatic flicker. Using a spatially nearly uniform color space, $\boldsymbol{CIE\ LCh_{a^*b^*}}$, the results show that the sensitivity to temporal changes in the lightness direction is higher than in the chroma and hue directions. The size of the effects justify the use of a specific model of perceived smoothness for control of dynamic lighting instead of the use of the traditional control in a device dependent color space.

Even though a difference in the absolute thresholds and the frequency dependence between smoothness thresholds and flicker visibility thresholds was found, the same *base color* and *direction of change* effects were demonstrated. The correlation of the results for different types of dynamic transitions enables building a model for perceived smoothness based on flicker visibility thresholds. The results of the experiments can be used to determine system requirements and devise temporal control algorithms which take into account the properties of the human visual system. The results for the dependence of the speed on frequency give a straightforward algorithm to determine the speed of a transition between two points given the frequency of the lighting system.

The results show two effects with interesting implications for future work. First, the sensitivity to changes in lightness is shown not to depend on the chromaticity of the *base color point*, simplifying the development of a future model. Second, the sensitivity for hue and chroma transitions do depend on the chromaticity of the *base color point*, demonstrating the need for a chromaticity dependency. The results presented in the chapter represent a starting point in the development of a full temporal smoothness and flicker visibility model.

To satisfy the smoothness requirements in a dynamic lighting system, new color points between fixed endpoints have to be introduced. To study this problem, Chapter 7 deals with the preferred temporal path between two colors in a color space. Through two experiments, the perceptually most optimal way to create a temporal color transition between two colors is investigated. The first experiment measures the ability to distinguish between two temporal color transitions. Discrimination thresholds ranged between 2.5 and $10.5\Delta E_{ab}$, and are significantly higher than the threshold for smoothness or flicker visibility. This is expected, as the temporal separation between the middle points in the transitions is much higher, in the order of seconds in contrast to flicker, where the stimuli are presented in direct succession. In the second experiment, several perceptually different color transitions were evaluated. The most preferred transitions are a linear transition in $\boldsymbol{CIE\ L^*a^*b^*}$ and a linear transition in $\boldsymbol{RGB}$. This suggests that appealing temporal color transitions can be created without complicated calculations.

There are a number of improvements possible for the studies presented in Chapter 7. Due to the large size of the space of possible transitions, a simple, but limited, parametrization was selected. A second limitation to the applicability of the results is the limited number of color pairs that constituted the beginning and ending point of the transitions. For both of the above, a study of existing dynamic light solutions, both natural and man made, can be used as inspiration for the search.

Lighting is seldom the focus of attention and it typically supports other activities. To this extent, Chapter 8 studies the difference between the temporal sensitivity of central and peripheral vision. Results from the experiments show that when using the ***CIE*** $\boldsymbol{LCh_{a^*b^*}}$ color space, which is almost uniform for spatial differences, the peripheral flicker sensitivity is much higher when the change is in lightness compared to chroma or hue, similar to sensitivity in central vision. Furthermore, the chromaticity independence of lightness flicker is also shown. Chroma and hue flicker gets impaired with increasing eccentricity. Lightness flicker sensitivity increases for $5Hz$ and $10Hz$ frequencies but decreases for higher frequencies when eccentricity increases.

Apart from shifting the focus away from the dynamic light, adding a task to a user introduces mental load. The final experiment of the chapter shows that this mental load has an impairing effect on flicker perception in the periphery.

The second part of the thesis demonstrates the usefulness and limitation of the use of spatial color spaces for predicting temporal quality of light. The approach in this thesis was application driven and the established models are empirical. However, using the theoretical framework of linear system analysis as used in model of luminance flicker has a promising prospect as an approach for modeling the effects described in these chapters.

Looking forward, there is a wealth of undescribed temporal artifacts present in LED lighting beyond the ones studied in this thesis. The interaction between modulated lighting and moving objects and eye movements for example is an especially interesting field of study as it impacts the perceived quality of a large number of light systems.

Similar to the effect creation, the freedom that LEDs give over traditional lighting systems can be used to optimize the light quality on any of the parameters discussed earlier. Interestingly, for different parameters, the main driving force behind the improvements is different. For example, due to the embedding of the light sources in the environment, the aesthetics of the light emitting surface become more important, but also are very different from the standard. The ability to use a large number of light sources distributed in the environment enables the optimization of object appearance properties. The unifying factor for all of the above is the presence of an observer, making perception research more important. Though important, perception is of secondary interest for the development of improvements in the light designed to influence biological and chemical processes which are and will be mainly driven by understanding those processes.

The wish of the author is that this collection of studies is only the first of many concerning AIL systems, enabling solid state lighting, and lighting in general to fulfill its full potential.

Bibliography

[1] E. Aarts, R. Harwig, and M. Schuurmans. *The invisible future: the seamless integration of technology into everyday life*, chapter Ambient intelligence, pages 235–250. McGraw-Hill, Inc., New York, NY, USA, 2002.

[2] I. Abramov and J. Gordon. Color vision in the peripheral retina. I. Spectral sensitivity. *Journal of the Optical Society of America*, 67(2):195–201, 1977.

[3] R. Adams and C. Atman. Cognitive processes in iterative design behavior. *Frontiers in Education, Annual*, 1:11A6/13–11A6/18vol.1, 1999.

[4] O. M. E. Alva, P. A. B. Martínez, L. J. M. Cueva, C. T. H. Sagástegui, and P. B. López. Comparison of methods and existing tools for the measurement of usability in the web. In *Proceedings of the 2003 international conference on Web engineering*, ICWE'03, pages 386–389, Berlin, Heidelberg, 2003. Springer-Verlag.

[5] amBX UK Ltd. `http://www.ambx.com/`. Accessed: May 2012.

[6] C. Bartleson. Memory colors of familiar objects. *Journal of the Optical Society of America*, 50(1):73–77, 1960.

[7] B. Berlin and P. Kay. *Basic Color Terms: Their Universality and Evolution*. Berkeley: University of California, 1969.

[8] R. Bowen. Latencies for chromatic and achromatic visual mechanisms. *Vision research*, 21(10):1457–1466, 1981.

[9] P. Boyce. *Human factors in lighting*. Applied Science Publishers, 1981.

[10] P. Boyce. *Contemporary ergonomics 2003*. Taylor & Francis Group, 2003.

[11] E. Brill. A simple rule-based part-of-speech tagger. In *Proceedings of the third Conference on Applied Natural Language Processing (ANLP'92)*, pages 152–155, Trento, Italy, 1992.

[12] S. Brin and L. Page. The anatomy of a large-scale hypertextual web search engine. *Computer Networks and ISDN Systems*, 30(1–7):107–117, 1998.

[13] J. Brox. *Brilliant: The Evolution of Artificial Light*. Houghton Mifflin Harcourt, 2010.

[14] J. Bullough, Y. Akashi, C. Fay, and M. Figueiro. Impact of surrounding illumination on visual fatigue and eyestrain while viewing television. *Journal of Applied Sciences*, 6(8):1664–1670, 2006.

[15] D. Carmel, P. Saker, G. Rees, and N. Lavie. Perceptual load modulates conscious flicker perception. *Journal of Vision*, 7(14):14, 2007.

[16] J. Chaves. *Introduction to nonimaging optics*. Optical science and engineering. CRC Press, 2008.

[17] K. Chen, S. Gao, Y. Zhu, and Q. Sun. Popular song and lyrics synchronization and its application to music information retrieval. In *Proceedings of the Thirteenth Annual Multimedia Computing and Networking 2006*, San Jose, CA, January 2006.

[18] J. Chuang, M. Stone, and P. Hanrahan. A probabilistic model of the categorical association between colors. In *Proc. of the Color Imaging Conference*, pages 6–11, 2008.

[19] CIE. Publication no 17.4, international lighting vocabulary. Technical report, Central Bureau of the Commission Internationale de L'clairage, Vienna, Austria, 1987.

[20] CIE. Publication no 15.3, colorimetry, third edition. Technical report, Central Bureau of the Commission Internationale de L'clairage, Vienna, Austria, 2004.

[21] D. Comaniciu and P. Meer. Robust analysis of feature spaces: Color image segmentation. In *IEEE Conf. on Comp. Vis. and Pattern Recognition*, pages 750–755, 1997.

[22] S. Coren, L. Ward, and J. Enns. *Sensation and Perception*. Harcourt Brace College Publishers, 1994.

[23] H. De Lange. Research into the dynamic nature of the human fovea-cortex systems with intermittent and modulated light: I. attenuation characteristics with white and colored light. *Journal of the Optical Society of America*, 48(11):777–784, 1958.

[24] H. De Lange. Research into the dynamic nature of the human fovea-cortex systems with intermittent and modulated light: Ii. phase shift in brightness and delay in color perception. *Journal of the Optical Society of America*, 48(11):784–789, 1958.

[25] DIALux software. `http://www.dial.de/DIAL/en/dialux.html`, Accessed October 2012.

[26] Digital Addressable Lighting Interface Activity Group, ZVEI Division Luminaires, Stresemannallee 19, D-60596 Frankfurt am Main, Germany. *DALI Manual*.

[27] B. Efron and R. Tibshirani. *An Introduction to the Bootstrap*. Monographs on Statistics and Applied Probability. Chapman & Hall, 1993.

[28] P. G. Engeldrum. *Psychometric Scaling: A Toolkit for Imaging Systems Development*. Imcotekpress, 2000.

[29] Entertainment Technology. *USITT DMX512A Asynchronous Serial Digital Data Transmission Standard for Controlling Lighting Equipment and Accessories*.

[30] European Commission. EU LED Quality Charter. `http://iet.jrc.ec.europa.eu/energyefficiency/residential-lighting/european-led-quality-charter`. Accessed: October 2012.

[31] European Commission. Green paper - lighting the future: Accelerating the deployment of innovative lighting technologies. Technical report, European Commission, 2011.

[32] M. D. Fairchild. *Color Appearance Models, Second Edition*. Wiley-IS&T Series in Imaging Science and Technology, Chichester, UK, 2005.

[33] S. Farahani. *ZigBee Wireless Networks and Transceivers*. Newnes, Newton, MA, USA, 2008.

[34] G. D. Finlayson and P. M. Morovic. Human visual processing: Beyond 3 sensors. In *IEE International Conference on Visual Information Engineering*, volume 1-7, pages 1–7, April 2005. Awarded best paper prize.

[35] D. M. Franz. Markov chains as tool for jazz improvisation analysis. Master's thesis, Faculty of the Virginia Polytechnic, Institute and State University, 1998.

[36] M. S. Freedman, R. J. Lucas, B. Soni, M. Schantz, M. Munoz, Z. David-Gray, and R. Foster. Regulation of Mammalian Circadian Behavior by Non-rod, Non-cone Ocular Photoreceptors. *Science*, 284:502, Apr. 1999.

[37] K. Fukunaga and L. Hostetler. The estimation of the gradient of a density function, with applications in pattern recognition. *Information Theory, IEEE Transactions on*, 21(1):32–40, 1975.

[38] D. Gamerman and H. Lopes. *Markov Chain Monte Carlo: Stochastic Simulation for Bayesian Inference*. Texts in Statistical Science Series. Taylor & Francis Group, 2006.

[39] G. Geleijnse and J. Korst. Efficient lyrics extraction from the web. In *Proceedings of the Seventh International Conference on Music Information Retrieval (ISMIR'06)*, pages 371 – 372, Victoria, Canada, October 2006.

[40] G. Geleijnse, D. Sekulovski, J. Korst, S. Pauws, B. Kater, and F. Vignoli. Enriching music with synchronized lyrics, images and colored lights. In *Proceedings of the 1st international conference on Ambient media and systems*, Ambi-Sys '08, pages 1:1–1:8, ICST, Brussels, Belgium, Belgium, 2008. ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering).

[41] J. J. Gooley, I. Ho Mien, M. A. St. Hilaire, S.-C. Yeo, E. C.-P. Chua, E. van Reen, C. J. Hanley, J. T. Hull, C. A. Czeisler, and S. W. Lockley. Melanopsin and rod-cone photoreceptors play different roles in mediating pupillary light responses during exposure to continuous light in humans. *The Journal of Neuroscience*, 32(41):14242–14253, 2012.

[42] T. Hansen, L. Pracejus, and K. Gegenfurtner. Color perception in the intermediate periphery of the visual field. *Journal of Vision*, 9(4):26, 2009.

[43] W. Hart Jr. The temporal responsiveness of vision. *Adler's physiology of the eye, clinical application*, pages 429–457, 1987.

[44] S. Hattar, H. Liao, M. Takao, D. M. Berson, and K. Yau. Melanopsin-Containing Retinal Ganglion Cells: Architecture, Projections, and Intrinsic Photosensitivity. *Science*, 295:1065–1070, Feb. 2002.

[45] S. Hecht. Rods, cones, and the chemical basis of vision. *Physiological Reviews*, 17(2):239, 1937.

[46] L. A. Hiller and L. M. Isaacson. *Experimental Music-Composition with an Electronic Computer*. 1959.

[47] J. Hui and D. Culler. Ipv6 in low-power wireless networks. *Proceedings of the IEEE*, 98(11):1865 –1878, 2010.

[48] S. Ishihara. *Ishihara's tests for colour deficiency*. Hodder & Stoughton, 2003.

[49] D. Iskander, Y. Wang, M.-Y. Kan, and H. Li. Syllabic level automatic synchronization of music signals and text lyrics. In *MULTIMEDIA '06: Proceedings of the 14th annual ACM international conference on Multimedia*, pages 659–662, Santa Barbara, CA, 2006.

[50] B. Johnstone. *Brilliant!: Shuji Nakamura and the Revolution in Lighting Technology*. Prometheus Books, 2007.

[51] M. J. Jones and J. M. Rehg. Statistical color models with application to skin detection. *Int. J. of Computer Vision*, 46(1):81–96, 2002.

[52] B. Kater. Music based light effects. Master's thesis, Eindhoven University of Technology, 2005.

[53] P. Kay and C. K. McDaniel. The linguistic significance of the meaning of basic color terms. *Language*, 54(3):610–646, 1978.

[54] D. H. Kelly. Effects of sharp edges in a flickering field. *Journal of the Optical Society of America*, 49(7):730–732, 1959.

[55] D. H. Kelly. Visual responses to time-dependent stimuli: I. amplitude sensitivity measurements. *Journal of the Optical Society of America*, 51(4):422–429, 1961.

[56] D. H. Kelly. Spatio-temporal frequency characteristics of color-vision mechanisms. *J. Opt. Soc. Am*, 64:983–990, 1974.

[57] D. H. Kelly. Luminous and chromatic flickering patterns have opposite effects. *Science*, 188(4186):371–372, 1975.

[58] P. Knees, M. Schedl, and G. Widmer. Multiple Lyrics Alignment: Automatic Retrieval of Song Lyrics. In *Proceedings of 6th International Conference on Music Information Retrieval (ISMIR'05)*, pages 564–569, London, UK, September 2005.

[59] R. Kuehni. *Color Space and Its Divisions: Color Order from Antiquity to the Present*. Wiley-Interscience, 2003.

[60] R. G. Kuehni. Focal color variability and unique hue stimulus variability. *Journal of Cognition and Culture*, 5:409–426(18), October 2005.

[61] T. S. Kuhn. *The structure of scientific revolutions / Thomas S. Kuhn*. University of Chicago Press, Chicago, IL :, 3rd ed. edition, 1996.

[62] R. Likert. A technique for the measurement of attitudes. *Archives of Psychology*, 140:1–55, 1932.

[63] B. J. Lindbloom. `http://brucelindbloom.com/`. Accessed: May 2012.

[64] A. Lindner, N. Bonnier, and S. Ssstrunk. What is the color of chocolate? - extracting color values of semantic expressions. In *Proc. IS&T 6th European Conference on Colour in Graphics, Imaging, and Vision (CGIV 2012).*, pages 355 – 361, 2012.

[65] A. Lindner, B. Z. Li, N. Bonnier, and S. Ssstrunk. A Large-Scale Multi-Lingual Color Thesaurus. In *Proc. IS&T/SID 20th Color and Imaging Conference (CIC)*, pages 30–35, 2012. Received the MERL Best Student Paper Award. For more information, code, and a link to the Thesaurus, see http://ivrg.epfl.ch/research/ColorThesaurus.html.

[66] J.-P. M. Linnartz, L. Feri, H. Yang, S. B. Colak, and T. C. W. Schenk. Communications and sensing of illumination contributions in a power led lighting system. In *proc. IEEE ICC, May 2008*, pages 5396–5400, 2008.

[67] J.-P. M. Linnartz, L. Feri, H. Yang, S. B. Colak, and T. C. W. Schenk. Code division-based sensing of illumination contributions in solid-state lighting systems. *IEEE Trans. on Sign. Proc.*, 2009.

[68] M. Luckiesh. *Torch of civilization: the story of man's conquest of darkness*. G. P. Putnam's Sons, 1940.

[69] D. MacAdam. Visual sensitivities to color differences in daylight. *Journal of the Optical Society of America*, 32:247–274, 1942.

[70] J. M. Martínez, editor. *MPEG-7 Overview ISO N6828*. October 2004.

[71] S. Marzano and E. Aarts. *The New Everyday View on Ambient Intelligence*. Uitgeverij 010 Publishers, 2003.

[72] E. Montag. The color between two others. In *Eleventh Color Imaging Conference, Scotsdale, Arizona*, pages 294–300, 2003.

[73] D. C. Montgomery. *Design and Analysis of Experiments, 5th Ed.*

[74] N. Moroney. Unconstrained web-based color naming experiment. In *Color Imaging VIII: Processing, Hardcopy, and Applications*, 2003.

[75] A. Munsell and R. Farnum. *A Color Notation: An Illustrated System Defining All Colors and Their Relations 1941*. Kessinger Publishing, 2004.

[76] A. A. Mury, S. C. Pont, and J. J. Koenderink. Representing the light field in finite three-dimensional spaces from sparse discrete samples. *Appl. Opt.*, 48(3):450–457, Jan 2009.

[77] S. Muthu, F. J. P. Schuurmans, and M. D. Pashley. Red, green, and blue leds for white light illumination. *IEEE J. Select. Topics Quantum Electron.*, 8(2):333–338, 2002.

[78] J. Nielsen. *Usability Engineering*. Morgan Kaufmann Series in Interactive Technologies. Acad. Press, 1994.

[79] M. Perz, D. Sekulovski, and M. Murdoch. Chromatic flicker perception in human peripheral vision under mental load. In *Eighteenth Color Imaging Conference: Color Science and Engineering Systems, Technologies, and Applications, San Antonio, Texas*, pages 33–37, 2010.

[80] C. Poynton. *Digital Video and HDTV Algorithms and Interfaces*. Morgan Kaufmann Publishers, 2003.

[81] R. Rajae-Joordens and J. Engel. Paired comparisons in visual perception studies using small sample sizes. *Displays*, 26:1–7, 205.

[82] B. Salters and M. Krijn. Color reproduction for led-based lighting systems. In *Proceedings of the SPIE*, page 6338. SPIE, 2006.

[83] A. Sarkar, M. Fairchild, and C. Salvaggio. Integrated daylight harvesting and occupancy detection using digital imaging. In *proc. Sensors, Cameras, and Systems for Industrial/Scientific Applications 2008*, 2008.

[84] M. Schedl, P. Knees, T. Pohle, and G. Widmer. Towards Automatic Retrieval of Album Covers. In *Proceedings of the 28th European Conference on Information Retrieval (ECIR'06)*, London, UK, April 2006.

[85] E. F. Schubert and J. K. Kim. Solid-State Light Sources Getting Smart. *Science*, 308(5726):1274–1278, 2005.

[86] D. Sekulovski, R. Clout, B. Kater, and T. Overbeek. Creation and rendering of stochastic dynamic light effects. In *proc of Seventeenth Color Imaging Conference, Albuquerque, New Mexico; November 2009*, pages 129–132, 2009.

[87] D. Sekulovski, L. Feri, T. C. W. Schenk, and J.-P. M. Linnartz. Interactive Light Design Using Coded Light. In *proc of Lux Europa 2009, September*, 2009.

[88] D. Sekulovski, G. Geleijnse, B. Kater, J. Korst, S. Pauws, and R. Clout. Enriching text with images and colored light. In *Proceedings of Multimedia Content Access: Algorithms and Systems II 2008*, Electronic Imaging 2008 '08, pages 682004.1–682004.15. SPIE, 2008.

[89] D. Sekulovski, I. Vogels, M. van Beurden, and R. Clout. Smoothness and flicker perception of temporal color transitions. *CIC 2007*, 2007.

[90] D. Sekulovski, I. Vogels, M. van Beurden, and R. Clout. Effect of Frequency on the Sensitivity to Temporal Color Transitions. In *CIE 2009*, 2009.

[91] P. Seuntiens, I. Vogels, and A. van Keersop. Visual experience of 3d-tv with pixelated ambilight. In *Proc Presence 200710th Ann. Int. Workshop on Presence, Barcelona, Spain*, pages 339–344, 2007.

[92] S. Shady, D. MacLeod, H. S. Fisher, and J. Y. Liang. Adaptation from invisible luminance and chromatic flicker. *J. Vis.*, 2(10):68–68, 12 2002.

[93] D. A. Shamma, B. Pardo, and K. J. Hammond. Musicstory: a personalized music video creator. In *MULTIMEDIA '05: Proceedings of the 13th annual ACM international conference on Multimedia*, pages 563–566, Singapore, 2005.

[94] G. Sharma, editor. *Digital Color Imaging Handbook*. CRC Press, Boca Raton, FL, 2003.

[95] R. Simpson. *Lighting control-technology and applications*. Focal Press, 2003.

[96] M. Stokes, M. Anderson, S. Chandrasekar, and R. Motta. A standard default color space for the internet — srgb. `http://www.w3.org/Graphics/Color/sRGB`. Accessed: May 2012.

[97] D. G. Stork and D. S. Falk. Temporal impulse responses from flicker sensitivities. *Journal of the Optical Society of America A*, 4(6):1130–, 1987.

[98] J. Stribling, M. Krohn, and D. Aguayo. Scigen. `http://pdos.csail.mit.edu/scigen/`, Accessed June 2009.

[99] W. Swanson, T. Ueno, V. Smith, and J. Pokorny. Temporal modulation sensitivity and pulse detection thresholds for chromatic and luminance perturbations. *Journal of the Optical Society of America A*, 4(10):1992–2005, 1987.

[100] W. Swanson, T. Ueno, V. Smith, and J. Pokorny. Temporal modulation sensitivity and pulse-detection thresholds for chromatic and luminance perturbations. *J. Opt. Soc. Am. A*, 4:1992–2005, 1987.

[101] L. Thurstone. A law of comparative judgment. *Psychological Review*, 34:273–286, 1927.

[102] C. Tyler and R. Hamer. Analysis of visual modulation sensitivity. IV. Validity of the Ferry-Porter law. *Journal of the Optical Society of America A*, 7(4):743–758, 1990.

[103] K. Uchikawa and I. Mitsuo. Temporal deterioration of wavelength discrimination with successive comparison method. *Vision Research*, 21:591–595, 1981.

[104] K. Uchikawa and H. Shinoda. Influence of basic color categories on color memory discrimination. *Color Research and Application*, 21 (6):430–439, 1996.

[105] US Department of Energy. CALiPER Program. `http://www1.eere.energy.gov/buildings/ssl/caliper.html`. Accessed: October 2012.

[106] M. van Doorn. *End-user programming of ambient narratives for smart retail environment*. PhD thesis, Technische Universiteit Eindhoven, 2009.

[107] V. Vezhnevets, V. Sazonov, and A. Andreeva. A survey on pixel-based skin color detection techniques. In *Proc. Graphicon-2003*, pages 85–92, Moscow, Russia, 2003.

[108] I. Vogels. Atmosphere metrics. In F. Toolenaar, J. H. D. M. Westerink, M. Ouwerkerk, T. J. M. Overbeek, W. F. Pasveer, and B. Ruyter, editors, *Probing Experience*, volume 8 of *Philips Research*, pages 25–41. Springer Netherlands, 2008.

[109] I. Vogels, D. Sekulovski, and B. Rijs. Discrimination and preference of temporal color transitions. In *Proc. of the 15th IS&T CIC Conference, Albuquerque, USA*, pages 118–121, 2007.

[110] I. Vogels, D. Sekulovski, and B. Rijs. How to create appealing temporal color transitions? *Journal of the SID*, 17(1):23–28, 2009.

[111] A. Walberg. *Light Vision Color*. John Wiley & Sons, Ltd, 2005.

[112] Y. Wang, M.-Y. Kan, T. L. Nwe, A. Shenoy, and J. Yin. Lyrically: automatic synchronization of acoustic musical signals and textual lyrics. In *MULTIMEDIA '04: Proceedings of the 12th annual ACM international conference on Multimedia*, pages 212–219, New York, NY, 2004.

[113] M. Weiser. The computer for the twenty-first century. *Scientific American*, pages 94–100, September 1991.

[114] G. Wyszecki and W. S. Styles. *Color Science: Concepts and Methods, Quantitative Data and Formulae (2nd ed.)*. John Wiley and Sons, Inc., New York,, 1982.

[115] A. Yamada, M. Pickering, S. Jeannin, L. Cieplinski, J. Ohm, and M. Kim. Mpeg-7 visual part of experimentation model version 9.0. n3914, 2001.

[116] H. Yang, J. Bergmans, and T. C. W. Schenk. A filter bank approach for led illumination sensing based on frequency division multiplexing. In *proc. of International Conference on Acoustics, Speech, and Signal Processing (ICASSP), April 2009*, pages 3189–3192, 2009.

Acknowledgments

Carrying out a PhD in an industrial environment can be a daunting process. The short lived projects and constant changes in the program require a lot of support to be able to build a consistent body of work. I am grateful to everyone that provided that support and contributed to the work presented.

I would first like to thank my first promotor, Emile Aarts for the always inspirational discussions and the motivation he provided during the whole process. I was always sure to come out of a meeting with him with a dozen of new ideas and insights, both scientific as well as business related. I learned about the existence of uniform color spaces and the intricacies of psychophysics studies from my second promotor, Ingrid Heynderickx. I want to also thank her for the selfless amount of time she spent on the early drafts of the thesis. The countless hours I spent working together with my copromotor Ramon Clout will remain one of the biggest prizes of my PhD work. His guidance and motivation made the everyday activities a breeze and most importantly, a lot of fun. I would like to thank the members of my reading committee, Prof. Eggen, Prof. Luo and Prof. de Ridder for their valuable insights and support. I would also like to thank Prof. Luo who introduced me to the world of CIE and helped me extend my network tremendously.

Even though not an official part of my supervisory team, Ingrid Vogels contributed hugely both to my understanding of human vision and experimental design and analysis as well as the enjoyment of everyday work. Jan Korst and Gijs Gelijnse contributed their knowledge of "Googling" techniques used the system presented in Chapter 3. Together with Steffen Pauws and Fabio Vignolli they also contributed to the audio segmentation and audio-lyrics alignment and the implementation of the demonstrator. I am grateful to Lorenzo Feri and Tim Schenk for the collaboration on the algorithms in Chapter 4 and introducing me to the inspirational world of coded light. I was lucky to work with very talented master students during the studies in the second part of my thesis. I am grateful to Maurice, Bartjan and Gosia for their hard work and dedication.

Two people deserve special credit for the existence of this thesis. Zharko Aleksovski is guilty for my arrival in the Netherlands as well as always being there for me. I would like to thank Srinivas Gutta for the confidence he had in me and the chance to start a PhD at Philips Research. I started my adventure in the Media Interaction group in the Intelligent Algorithms cluster together with Fabio, Gijs, Herman, Kero, Steffen, Igor, Jan, Verus, and Warner. I spend a lot of memorable moments with my fellow

PhDs and officemates Bram, Greg, Herman and Pieter, Marco, Maria, Marjolein, Peter Lutz, and Theresa both at work and after. I enjoyed a lot of discussions on color, computation, and religion with the 4th floor bunch: Bart, Padre Marco, Mauro, and Prathna. After joining the Visual Experiences group I had the privilege to work in the Visual Perception cluster with Arnold, Gosia, Marco, Mariska, Mike, Pieter, Ronald, Roos, Stefan, and Ulrich. The friendly atmosphere, the level of expertise and drive to constant improvement in this small group is seldom found. I am very grateful to Toon Holtslag and Ad Leenaars for the understanding and motivation during the final period of my thesis work.

My PhD life was made more pleasant by my friends Afra, Barbara, Bart, Bahman, Dani, Giovanni, Janto, Jason, Jettie, Kavitha, Kero, Prathna, Lorenzo, Noelie, Marco, Meredith, Shalu, Stefan, Thiru, and Tsvete. I would like to thank the members of the Macedonian and Serbian diaspora in The Netherlands Ana, Caka, Dragan, Jasen, Meri, Natasa, Sonja, and Zvezdan for making me feel like home.

A special thanks to Mama, Buba and Cico and the rest of my family. Their support over the years enabled me to pursue the thing I always found exciting and fun, science.

Finally, I would like to thank my adventure buddy and best friend Biljana for her endless understanding, support and love.

Studies in Ambient Intelligent Lighting

Summary

The revolution in lighting we are arguably experiencing is led by technical developments in the area of solid state lighting technology. The improved lifetime, efficiency and environmentally friendly raw materials make LEDs the main contender for the light source of the future. The core of the change is, however, not in the basic technology, but in the way users interact with it and the way the quality of the produced effect on the environment is judged. With the new found freedom the users can switch their focus from the confines of the technology to the expression of their needs, regardless of the details of the lighting system. Identifying the user needs, creating an effective language to communicate them to the system, and translating them to control signals that fulfill them, as well as defining the means to measure the quality of the produced result are the topic of study of a new multidisciplinary area of study, Ambient Intelligent Lighting.

This thesis describes a series of studies in the field of Ambient Intelligent Lighting, divided in two parts. The first part of the thesis demonstrates how, by adopting a user centric design philosophy, the traditional control paradigms can be superseded by novel, so-called effect driven controls.

Chapter 3 describes an algorithm that, using statistical methods and image processing, generates a set of colors based on a term or set of terms. The algorithm uses Internet image search engines (Google Images, Flickr) to acquire a set of images that represent a term and subsequently extracts representative colors from the set. Additionally, an estimate of the quality of the extracted set of colors is computed. Based on the algorithm, a system that automatically enriches music with lyrics based images and lighting was built and is described.

Chapter 4 proposes a novel effect driven control algorithm, enabling users easy, natural and system agnostic means to create a spatial light distribution. By using an emerging technology, visible light communication, and an intuitive effect definition, a real time interactive light design system was developed. Usability studies on a virtual prototype of the system demonstrated the perceived ease of use and increased efficiency of an effect driven approach.

In chapter 5, using stochastic models, natural temporal light transitions are modeled and reproduced. Based on an example video of a natural light effect, a Markov model of the transitions between colors of a single light source representing the effect is learned. The model is a compact, easy to reproduce, and as the user studies show, recognizable representation of the original light effect.

The second part of the thesis studies the perceived quality of one of the unique capabilities of LEDs, chromatic temporal transitions. Using psychophysical methods, existing spatial models of human color vision were found to be unsuitable for predicting the visibility of temporal artifacts caused by the digital controls. The chapters in this part demonstrate new perceptual effects and make the first steps towards building a temporal model of human color vision.

In chapter 6 the perception of smoothness of digital light transitions is studied. The studies presented demonstrate the dependence of the visibility of digital steps in a temporal transition on the frequency of change, chromaticity, intensity and direction of change of the transition. Furthermore, a clear link between the visibility of digital steps and flicker visibility is demonstrated. Finally, a new, exponential law for the dependence of the threshold speed of smooth transitions on the changing frequency is hypothesized and proven in subsequent experiments.

Chapter 7 studies the discrimination and preference of different color transitions between two colors. Due to memory effects, the discrimination threshold for complete transitions was shown to be larger than the discrimination threshold for two single colors. Two linear transitions in different color spaces were shown to be significantly preferred over a set of other, curved, transitions.

Chapter 8 studies chromatic and achromatic flicker visibility in the periphery. A complex change of both the absolute visibility thresholds for different frequencies, as well as the critical flicker frequency is observed. Finally, an increase in the absolute visibility thresholds caused by an addition of a mental task in central vision is demonstrated.

Curriculum Vitae

Dragan Sekulovski was born 30.08.1979 in Skopje, R. Macedonia. After finishing secondary education in the State High School “Rade Jovcevski Korcagin” in Skopje in 1998, he enrolled at the University “St. Cyril and Methodius” in Skopje. There he studied Computer Science at the Faculty of Natural Sciences and Mathematics. In 2004 he did his diploma internship at Philips Research Europe in Eindhoven. The work in the diploma thesis entitled “Video Algorithms for Living Light” concentrated on the extraction of color information from video and subsequent reproduction with light. In 2005 he started his PhD research as a *van der Pol Junior* in the Media Interaction group at Philips Research. This thesis is the result of that research. Since 2008 he works at the Visual Experiences group at Philips Research on topics related to visual perception of displays and lighting. He has over 30 publications in international conference proceedings and journals, and 19 patent applications.