UNIVERSITÀ DEGLI STUDI
DI MILANO

SCUOLA DI DOTTORATO IN INFORMATICA
DIPARTIMENTO DI INFORMATICA E COMUNICAZIONE
XXII Ciclo
Settore Scientifico Disciplinare INF/01

Tesi di Dottorato di Ricerca

Testing colour appearance models in complex scene

Cristian Bonanomi

Relatore: Prof. Daniele Marini
Correlatore: Prof. Alessandro Rizzi

Il Direttore della Scuola di Dottorato in Informatica:
Prof. Ernesto Damiani

Anno Accademico 2009/2010

Abstract

The sensation of sight is our primary mechanism to perceive the world around us. However it is not yet perfectly clear how the human visual system works. The images of the world are formed on the retina, captured by sensors and converted in signals sent to the brain. Here the signals are processed and somehow interpreted, thus we are able to see.

A lot of information, hypothesis, hints come from a field of the optical (or visual) illusions. These illusions have led many scientists and researchers to ask themselves why we are not able to interpret in a correct way some particular scenes. The word “interpret” underlines the fact that the brain, and not only the eye, is involved in the process of vision. If our sight worked as a measurement tool, similar to a spectrophotometer, we would not perceive, for example, the simultaneous contrast phenomenon, in which a grey patch placed on a black background appears lighter than an identical coloured patch on a white background. So, why do we perceive the patches as different, while the light that reaches the eyes is the same? In the same way we would not be able to distinguish a white paper seen in a room lit with a red light from a red paper seen under a white light, however humans can do this. These phenomena are called colour appearance phenomena. Simulating the appearance is the objective of a range of computational models called colour appearance models. In this dissertation themes about colour appearance models are addressed. Specific experiments, performed by human observers, aim to evaluate and measure the appearance. Different algorithms are tested in order to compare the results of the computational model with the human sensations about colours. From these data, a new printing pipeline is developed, able to simulate the appearance of advertising billboard in different context.

Acknowledgements

This thesis is dedicated to my wife, Sara. When I was in the dark, her light was shining, indicating the way.

Thanks to my family that always pull for me, *grazie ma' e pa'*. And Silvia, Mirco and nonna.
Thanks to Sara's parents that believe in me.

And then thanks to everybody I met in this coloured path.
To Prof. Marini that four years ago took me in his lab.
To Alessandro, because he taught me the 1% of what he knows. I hope some day to learn also the remaining 99%. And maybe more?
Thanks to the Eidoboys, Alberto, Dario and Tamberlo. When I was in the dark Norway, they came through the fog and bring me in a pub full of drunken people. Wonderful night. And yes, sometimes we spoke also about work.
To Saim, I hope one day you will learn Italian.
To Prof. G. alias Davide, for his help and *bla bla bla* (often useful).
To the guys of the Artificial Intelligence Lab, that allow me to use their microwave oven.
Thanks to Isabella, sometimes I don't understand you.
Thanks to Prof. Borghese, he introduced me to the PhD.
Thanks to Anna *Barolo,* you know why I wrote your surname in italic.
Thanks to the Create, it was the occasion to take my first flight. And thanks to all the friends I met there, too.

Thanks to my referees Dr. Parraman, Prof. Oleari, and Prof. Chambah to their kindness in accepting to read this thesis. The comments after the first draft improve enormously this work, making it more precise.

Finally, a big thanks to all my friends, that keep me alive in these years. I am quite sure that at this point they would say: *grazie Cri*.

1. Introduction

The sensation of sight is our primary mechanism to perceive the world around us. The human visual system (HVS) processing is often mistakenly compared to the acquisition of an image performed by a camera. However, given an example of two identical grey patches that are placed on a different background, a camera will capture them with the same digital values, while our HVS perceives them differently. This effect is called *simultaneous contrast*, described for the first time by the French chemist Michel Eugène Chevreul (Chevreul, 1854 p. 30). From this example is clear that colour cannot be just considered as a physical quantity. It is often asserted that colour is an attribute of visual sensation (Young, 1802).

The images of the world are formed on the retina, captured by sensors and converted into signals which are sent to the brain. Here the signals are processed and interpreted in several stages (Boynton, 1979). It is important, at this point, to clarify the difference between *sensation* and *perception*. These terms are often used interchangeably but they have different meaning. Sensation is related with the stimulation of the organism, while perception involves sensation and past experience, in order to recognize the objects (Optical Society of America, 1953). In (McCann, et al., 1983) is given an example to explain this concept. A picture of a raft photographed in a lake is shown to some observers (figure 1). In the picture, one side of the raft is illuminated by the sun-light and the other is illuminated by the sky-light. Asking to the observers which paint should a painter use to depict an image of the raft, the answer will be white with some yellow for the sunlit side and grey with some blue for the other side. This is related to sensation. However, asking which paint should be used in order to re-paint the raft, the answer will be white for both the faces. This is related to perception. Sensation is connected with *appearance*. Simulating the appearance is the objective of some computational models called *colour appearance models* (CAMs).

Fig. 1: the original image of the raft on the lake. Image courtesy of J. McCann.

A lot of information, hypothesis, hints on the functioning of the human visual system come from the field of optical or visual illusions (like the previously mentioned simultaneous contrast). Illusions led many scientists and researchers to wonder why we are not able to interpret some particular scenes in a correct way. The word interpret underlines how the brain, and not only the eye, is involved in the process of vision. Some of these visual illusions are called colour appearance phenomena (Hunt, 1977), (Fairchild, 2005 p. 111-133). The aim of the CAMs is to simulate these phenomena.

1.1 Aims and objectives

The intent of this dissertation is to contribute the knowledge of themes about colour appearance models. We have already stated that appearance means how colours are perceived by the human being. Is it possible to reproduce and evaluate the appearance? Moreover, is it possible to compute the appearance algorithmically? In chapter 7 we are going to answer to these questions, performing experiments to study simultaneous contrast (Chevreul, 1854 p. 30), (Hering, 1964) and colour constancy (Land, 1977) both psychophysical experiments and as result of CAMs computation. Aim of the experiments is to give a quantitative evaluation of the appearance as perceived by humans, and as calculated by CAMs. However, the appearance of the scene calculated by an algorithm is related to the input image. If we are going to evaluate the appearance of a natural scene, starting from a picture of that scene, taken by a camera, we need to know if the values captured by the camera are truthful. In chapter 6 a range of experiments are performed to investigate the capability of a camera to capture the

luminance of a scene. Different cameras have different in-built features that can affect the final result. When the light reaches the sensor, how does it respond? The light should be capture linearly proportional to the light in the scene, this is called *raw* file, usually 10-14bit per channel. The raw file is then processed to generate the typical 8-bit image. However, the high dynamic range of luminance in a natural scene cannot be capture by a single shop. HDR imaging (chapter 5) is a technique to increase the dynamic range. More pictures are used to create a radiance map of the scene. In chapter 6 we are going to investigate these aspects of the photographic processing, as the resulting image will become later the input of the appearance model.

Many applications could take advantage of being able to predict some behaviours of the human visual system. Knowing in advance the visual rendering of an image in different conditions can avoid problems related to its visualization (Bonanomi, et al., 2009). This can be related to different media on which an image is presented (monitor, paper), different context/environment in which an image is set, or different kind of illumination (as shown in chapter 8). We also used the Random Spray Retinex (RSR) and the Automatic Colour Equalization (ACE) (derived by the Retinex theory and presented in chapter 3) to filter different stereoscopic streams with visible colour differences between right and left frames. Both the algorithms can correct, with different accuracy, colour disparities in the stereo pairs (Gadia, et al., 2010). HDR image tone mapping (or better tone rendering) (Debevec, et al., 1997), (Rizzi, et al., 2007) can be improved using the appearance framework in order to compress the dynamic range of an image and enhance the local contrast. The automatic colour equalization (ACE) it is used enhance the appearance of the software interfaces, in order to improve the usability (Rizzi, et al., 2003). The same algorithm is applied to digitally restore movies after scanning the film. This technique aims to correct colour and dynamic range of old movies that lost their original appearance (Chambah, 2008), (Rizzi, et al., 2006). A similar application of the algorithm is used to process underwater colour images that are characterized by a strong blue colour cast (Semani, et al., 2005). The Retinex theory is widely used by the Langley research centre of the NASA (NASA, 2010). They developed a multi-scale Retinex to improve brightness, contrast and sharpness in images for a lot of application: from aerial imagery (Woodell, et al., 2006) to improve the performance of segmentation algorithms (Rahman, et al., 2004). The appearance model developed by the CIE (presented in chapter 4) is used to segment the head from the rest of the scene in an application which determines the head movement from pictures taken by digital cameras during PET (Positron

Emission tomography) scanning (Gao, et al., 2007). The same model is used for segmentation of traffic sign (Gao, et al., 2008).

1.2 Context

Studying the stimuli as seen inside a context is quite a recent approach with respect to the history of colorimetry. Wyszecki make the distinction between basic and advanced colorimetry (Wyszecki, 1973). Basic colorimetry provides a measurement of the colours observed in isolated viewing condition. The CIE XYZ and the CIELAB colour space are not enough to specify the features of a colour seen in a certain context. However, in everyday life, colours are rarely seen in isolation; rather, they are seen in complex scene, with elaborated surround, and in disparate illumination conditions. Therefore, it originates the needs of an advanced colorimetry, able to specify the features of a colour seen in a complex environment. As explained in the case of simultaneous contrast, two identical grey patches on different background match if measured with the basic colorimetry. On the other hand, specifying the colour of the background, the advanced colorimetry will measure different values for the two stimuli. However, the early models (Hunt, 1982), (Nayatani, et al., 1987) and the CIE models CIECAM97s (CIE, 1998), CIECAM02 (Moroney, et al., 2002), have some limitations. For example the background of a stimulus can be specify only in terms of its luminance and not its chromaticity. Furthermore these models do not consider the spatial properties of an image. Every pixel is considered as an independent stimulus. This is, of course, an extreme simplification of the appearance calculation that make these models unsuitable to work in the field of the imaging. In 2002 Mark Fairchild developed iCAM (image Colour Appearance Model) (Fairchild, et al., 2002), (Fairchild, et al., 2004) in order to predict the colour appearance of an image, and not only the appearance of a single stimulus. This model is a sort of extension of the CIECAM02 to which is strictly connected.

Regardless of the CIE, from the beginning of the 1970s, Land and McCann, developed the Retinex theory (Land, et al., 1971), (Land, 1977). Their aim is still to simulate the appearance, mimicking some behaviours of the human visual system. In this theory every pixel of the scene is recomputed according to the spatial distribution of the other pixels. Starting from the Retinex theory a range of models has been implemented that modify the

colour information in the visual context (Land, et al., 1971), (Rizzi, et al., 2003), (Provenzi, et al., 2007), (Provenzi, et al., 2008)[1].

1.3 Methodology

In (Moroney, et al., 2004) a qualitative comparison between iCAM (Fairchild, et al., 2002) and a version of Retinex (Funt, et al., 2000) is reported. Three high dynamic range (HDR) greyscale images had been filtered using iCAM and Retinex. The computational results has been shown to ten observers, which had to rate the images depending on their preference. For one image the iCAM computation resulted better than Retinex, for another Retinex was preferred to iCAM, and for the third there was no consistent preference.
In (McCann, et al., 2009) two identical 3D scene of painted wooden shapes have been built. One was illuminated uniformly, in order to have a low dynamic range scene (LDR). For the second scene, a directional non uniform illumination was used, to achieve a high dynamic range scene (HDR) [2]. The observers was asked to compare the appearance of the surfaces in the two scene with a test target that had the same reflectances of the painted wooden shapes. Furthermore one of the author painted a watercolour copy of the two scenes. The purpose was to study the colour constancy effect in a 3D scene, with different type of illumination. The results showed that in uniform illumination many samples looked the same. However, in the directional illumination, for the most of the case the samples looked dissimilar.

In this thesis some quantitative tests are performed (chapter 7) to compare iCAM06 (Kuang, et al., 2007) and Retinex (Provenzi, et al., 2007) with the rendering of the real scene performed by some human observers. Three different scenes have been set up, in order to study two different phenomena: simultaneous contrast and colour constancy. The observer were asked to paint in different way a copy of the scene, composed only by the contours. This step it is used to estimate the appearance of the scene as the human visual system perceive it. The digital image in input to the two algorithms have to be a high dynamic range image (HDR image). An HDR image is an image with a dynamic range of luminance greater than a

[1] In literature, the acronym CAM is usually referred to a family of models that include the Hunt, the Nayatani and the CIE models. The iCAM model is sometimes called scene rendering algorithm. The Retinex based models are often referred to as SCA (Spatial Color Algorithms) (Rizzi, et al., 2007). However in this thesis the term CAM is used in general to include all the models that compute the appearance of the colors, regardless to the method adopted.

[2] The dynamic range of a scene is the ratio between the lightest point of a scene and the darkest point of the same scene. LDR is also used to indicate 8-bit per channel images, where 8 bit are enough to capture the luminance of the scene, while HDR is used to indicate images that have more than 8 bit per channel, to capture more dynamic range of luminance.

standard 8-bit image (Debevec, et al., 1997). It is possible to achieve an HDR image using several pictures of the scene taken with different exposure time (Debevec, et al., 1997). In order to create an HDR image as input for the models tested, eleven shots of the scene have been taken. In chapter 7 the results of the experiments on the appearance are presented.

It is important to prepare the data in input to the appearance models. The aim of the HDR technique is to create a radiance map proportional to the radiance values in the scene (Debevec, et al., 1997). However many factors can influence this reconstruction. In chapter 6 a range of experiments are performed in order to understand what factors can affect the results. Building factors (lens, sensor) and software of the camera can modify the pixel values of the picture taken. Furthermore, different software for the creation of the HDR image use different methods to reconstruct the radiance of the scene, leading to different results.

1.4 Thesis organization

The dissertation is organized as follows:

In chapter 2 an overview of the human visual system is given (Fairchild, 2005), (Oleari, 2008). The explanation of the main vision theories (Young, 1802), (von Helmholtz, 1892), (Hering, 1964), (Boynton, 1979), (Zeki, 1993) is provided after a schematic description of the eye. The chapter ends with the depiction of some appearance phenomena (Hunt, 1977), (Fairchild, 2005 p. 111-133).

In chapter3 we review the Land's Retinex theory (Land, 1977), (Land, et al., 1971) and the algorithms resulting from it (Rizzi, et al., 2003), (Provenzi, et al., 2008), (Provenzi, et al., 2007). This kind of algorithms attempt to simulate the human visual system behaviour.

In chapter 4 we focus on the classical CAMs that led CIE to conceive the CIECAM02 (Moroney, et al., 2002), and the iCAM (Fairchild, et al., 2002), (Kuang, et al., 2007), originating from it.

In chapter 5 a background on the raw format and the HDR image acquisition process is exposed (Debevec, et al., 1997), (Reinhard, et al., 2005).

Some experiments are performed in chapter 6 in order to understand if the real dynamic range of a scene can be captured with this kind of image, and what factors can influence the results.

In chapter 7 some experiments are executed to make a quantitative estimation of the appearance. These experiments are inspired by McCann and al. idea (McCann, et al., 2009). Some human observers reproduce the appearance of a scene placed in front of them.

Subsequently the Random Spray Retinex (Provenzi, et al., 2007) and iCAM06 (Kuang, et al., 2007) algorithms are used to predict the appearance of the scene, working on an HDR image. The result of the computation is compared with the reproductions of the observers.

In chapter 8 a working pipeline is implemented for a real case. The objective is to study the appearance, here intended as colour and contrast reproduction, of advertising in indoor and outdoor environments, under different light conditions, and printed on different supports. A CMYK image is converted into a spectral image (Neugebauer, 1937), (Rolleston, et al., 1993), (Zuffi, et al., 2006), in order to add a certain illuminant. After being converted in RGB, the visual appearance of the image in the context is calculated.

In Appendix A the colour appearance terminology is depicted. Terms as lightness, chroma, saturation are defined. In Appendix B there is an overview of the pipeline implementation using NVdia CUDA™, a parallel computing architecture which is able to solve many complex computational problems in a fraction of the time required on a CPU.

2. The human visual system

2.1 Introduction

In order to understand and interact with the world that surround us, the most powerful sense we use is the sight. Our faculty of seeing the objects around us is due to the light that hits the matter and it is reflected or diffused until it reaches our eyes.

The light is an electromagnetic radiation, that is an electromagnetic field propagating in the space with wavelength approximately between 380 and 730 nm. Light of different wavelengths in the visible range stimulates in different ways the receptors set on the retina causing the perception of colours. The nature of the light has been explained by Newton's experiments (Newton, 1704), in which the light, passing through a prism, is decomposed in its constituent spectral components. Therefore the light can be physically characterized by its spectral composition, in the form of a spectral power distribution (SPD) of the energy in function of the wavelengths.

In the following chapter a brief introduction to the eye is presented, summarizing some ideas useful for the purpose of this dissertation. In chapter 2.3 some historical theories, that are the basis of the modern theory of the vision, are presented. Finally, in chapter 2.4, concepts about colour appearance phenomena are introduced.

2.2 A description of the functions of the eye

The structure of the eye is roughly spherical, with a diameter of about 2,4 cm. Located in the frontal part there is an external lens, the cornea, in contact with a transparent liquid called aqueus humour. Behind the aqueus humour the iris controls the diameter of the pupil, that is a hole able to change its diameter from about 2 mm to 8 mm, in order to allow more or less light to reach the retina, the membrane that lies in the bottom of the eye where the photoreceptors are set. In order to direct the light to focus on the retina, the lens, a biconvex structure set behind the iris, is able to change its shape altering the focal distance of the eye. This process is called accommodation (von Helmholtz, 1962). When we look at something an optical image is projected on the retina. An illustration of the eye can be seen in figure 2.

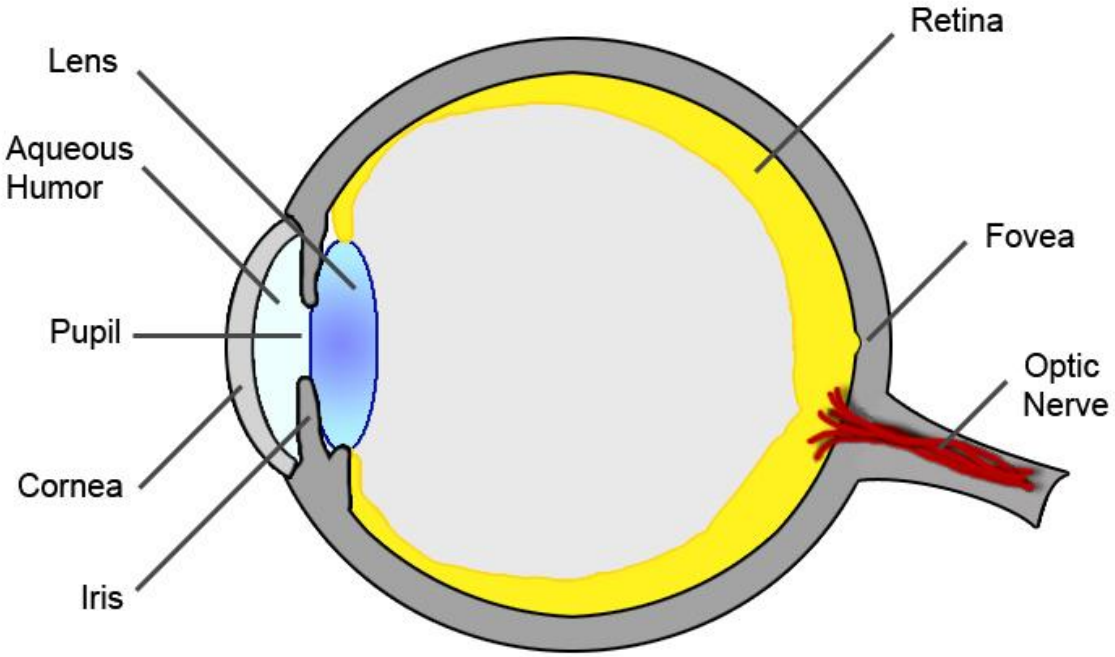


Fig. 2: A schematic diagram of the human eye.

Two types of photosensitive cells, called rods and cones, are located on the retina. These neurons can be considered as part of the brain. Their purpose is to transform the information present in the optical image as chemical and electrical signals that are then transmitted and processed by the brain. The specific process is still not perfectly clear, but is important to be aware that it consists of complicated combinations of the receptor signals. In particular, rods are responsible for low luminance levels vision (scotopic vision), while at high levels they are saturated. In this situation (photopic vision) cones are active, and permit the vision of the colours. Mesopic vision is referred to an intermediate level of illumination, in which both rods and cones are active. In the central part of the retina there is an area called the fovea, in which the best spatial and colour vision is possible. The fovea contains only cones, and the highest visual acuity is possible. Moving to the periphery of the fovea the number of cones decrease while the number of rods increase (figure 3).

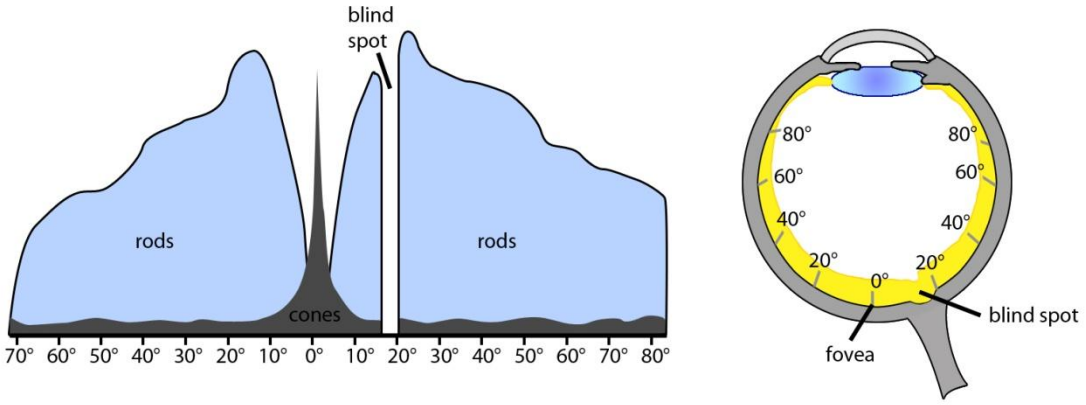


Fig. 3: distribution of cones and rods in the retina.

Cones provide the colour vision and are characterized by three different kind of pigments that absorb the spectral radiation in different percentages. Roughly speaking, we say there are three type of cones sensible to different regions of the spectra: L cones have an absorption curve that cover long and medium wavelength of the visible spectrum, with a peak at around 560 nm, M cones in the medium wavelength, with a peak at 530 nm, and S cones in the short wavelength, with a peak at 420 nm. Note that the spectral responsivities of the M and L cones are partially overlapped, while the S cones curve is separated from the former two. The spectral peak of the rods, that perceive only shades of grey, is at 507 nm.

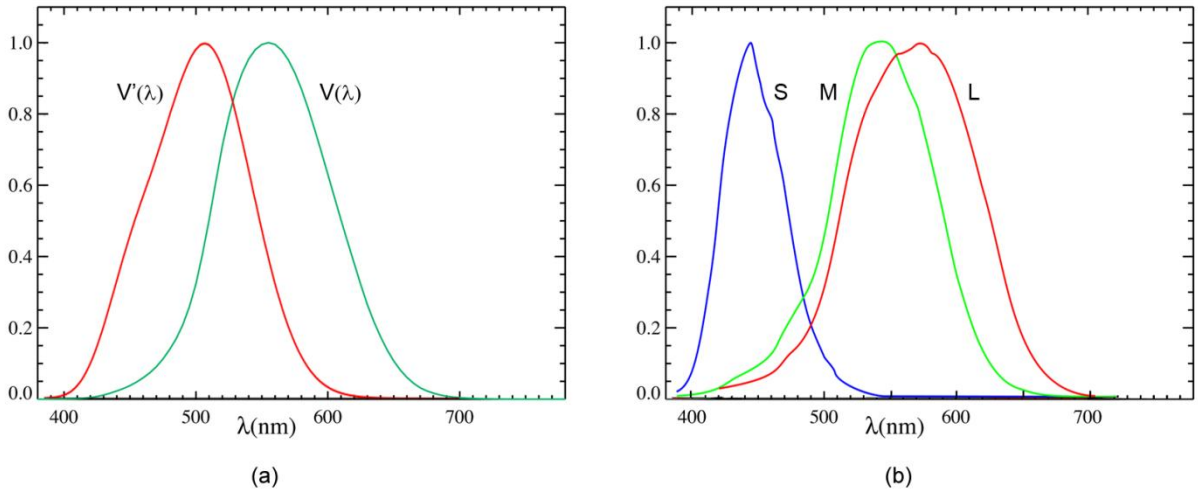


Fig. 4: (a) scotopic (in red) and photopic (in green) functions. (b) Cone responsivities.

In figure 4b it is possible to see the spectral absorption of the cones, normalized to their maxima. However is important to be aware that the three types of cones are present in the retina at different amounts. The relative populations is approximately 12:6:1 for the L:M:S

cones. There are two curves defined from the CIE (Commission Internationale de l'Eclairage) that represent the sensitivity of the eye in photopic (V(λ)) and scotopic (V'(λ)) vision. The latter is identical to the spectral responsivity of the rods, as only one type of these photoreceptors exist. The V(λ) instead, is a weighted combination of the three types of cones, keeping in mind that the S cones are not an important contribution to the lightness (figure 4a) The difference between the peaks of the two curves causes the Purkinje effect (Purkyně, 1825) that is the shift of the spectral sensitivity from the 555 nm to the 507 nm when decreasing the illumination level[3].

2.3 Vision theories

2.3.1 A background to the theories on trichromacy

In the 18th century, Thomas Young (Young, 1802), and later Helmholtz (von Helmholtz, 1892), (von Helmholtz, 1962) hypothesized that three different kind of mechanisms, sensible to three different spectral regions, resided inside the eye. These receptors are sensitive approximately to red, green and blue colours of the spectrum . The excitements of these three neural mechanisms are proportional to the respective stimulations of the radiation reaching the eye. This mechanism can be empirically verified by the fact that it is possible to match the colour of any radiation in the visible spectrum summing the appropriate amount of three primary radiations, that is, three colorimetrically independent light sources. With this term is intended that none of the spectra can be obtained with a mixture of the other two. The colour television is based on this concept (additive synthesis), where the combination of the three primaries takes place and is based on the spatial integration in the eye of the observer. The experimental verification of the presence of three kind of cones in the retina is dated 1964 (Marks, et al., 1964), (Brown, et al., 1964).

2.3.2 A background to the theories on opponent colours

Young-Helmholtz's (Young, 1802) (von Helmholtz, 1892) theory is able to explain the phenomena of additive mixture, but fails to explain some visual effects. For example, looking at a red patch for some time, and moving the gaze to a white surface, like a wall, one will see a greenish patch. This is called afterimage and was illustrated by Goethe at the beginning of 1800s (Goethe, 1840 p. 22).

[3] For example in the sunlight a red flower appears brighter than a blue flower, while in the evening, at low illumination level, the red flower appears darker than the blue one.

Hering (Hering, 1964) noted that certain hues were never perceived to appear together, for example red and green, or yellow and blue. Furthermore he realized that the humans perceived four different hues, corresponding to red, green, blue and yellow. The latter can be produced combining red and green in additive way, but generally yellow is perceived to be a fundamental colour. By explaining these effects in 1872 (Hering, 1964) Hering proposed his opponent colour theory, claiming that there were four elementary colours: red, green, blue and yellow and that the visual system divides the information coming from the eyes in three different channels: red-green, blue-yellow and black-white, the latter relative to brightness.

Hering's theory seems to be antagonist trichromacy, but in recent years the two theories have been combined in order to explain how the human visual system works (Boynton, 1979). Both are corrected but are referred to different stages of the vision. At receptor level we can consider the trichromatic theory as valid, since experimentally three kinds of cones are based in the retina. These are sensitive to certain wavelengths, corresponding roughly to red, green and blue. Therefore the colour is encoded in opponent signals and processed by the brain.

Summarizing in few words: we can say that the path of the visual information, after being acquired and converted as a chemical and electrical signal from the photoreceptors located on the retina, passes through the optical nerve to the lateral geniculate nucleus (LGN), and from here to the visual cortex. Here the process becomes really complex as we explain in the next section.

2.3.3 A background to visual processing of the brain

A further indication of the involvement of the brain in the colour perception derives from Semir Zeki's work (Zeki, et al., 1991), (Zeki, 1993). In this section we describe in broad terms his theory, in order to show the complexity of the way the brain processes the visual signals. From his point of view the brain, even if living in an environment that is always changing, is able to acquire an invariant knowledge about certain constant or physical characteristic of the objects. For example by varying the illumination of a surface, the reflected wavelengths change, but our brain is able to assign to the surface a constant colour. During our daily experience, the objects and the surfaces are seen under different angles and distances, in different contexts and in continually changing illumination conditions, but they maintain their identity, that is their shape, size and colour. This property, known as constancy, is often linked to the colour but exists in other fields. It is clear that the cerebral cortex is not a passive analyzer of the external world but it requires some functioning to discard the ever-

changing information that reaches the brain and attempts to extract the physical constant (colour, form, motion) of the objects. So the interpretation comes together with the sensorial perception.

During the 1970s Zeki led studies on Macaques, by testing the cerebral activities of the prestriate cortex using electrodes. He discovered that there are different visual areas, each specialized in a specific visual task (Zeki, 1978), (Zeki, 1993). Summarizing and simplifying his work on monkeys, he found that the retina is almost uniquely connected with the visual cortex, an area called V1, in a way that adjacent retinal points are connected with adjacent cortical points in V1. Furthermore these are connected with another area that surrounds V1 called V2; the V1 and V2 have a detailed topographic map of the retina. These two areas work as pre-processors, by analyzing the entire visual field and propagating different signals to some other specific areas. Areas V3 and V3A, for example, are composed of cells responsive to lines of specific orientation, but regardless of the colours or their background. The most part of the cells in V4 are wavelength selective. Finally cells in another prestriate cortex, called V5 (or MT) react to the motion. From these experiments Zeki assumes that the brain is partitioned in functional specific areas (Zeki, 1978), (Zeki, 1993).

In 1988 a new measuring device (Positron Emission Tomography or PET) was developed in order to investigate the functions of the brain of the man/woman. Activity in a region of the brain imply an increasing of the local cerebral blood flow through it, that can be detected by the PET, that consists in introducing a known amount of detectable substance into the blood, in order to be able to calculate the amount flowing through any region and detected increases or decreases. Two kinds of experiment has been undertaken by Zeki and his colleagues, one concerned with colour, the other with motion (Zeki, et al., 1991).

In the first experiment the subject looks at a "Land colour Mondrian", a collage of papers with different sizes and colours[4]. In this case has been recorded an activity in two areas of the brain, one in the striate cortex and the cortex surrounding it, that we call human V2, and one centred on the fusiform gyrus. Viewing the same scene, but in shades of grey, the activity in the striate cortex was the same, but that on the fusiform gyrus was much reduced. So, he called that area human V4, similarly to the V4 of the Macaque.

[4] In this experiment a collage of 15 multi-coloured rectangles was used (Zeki, et al., 1991) in order to have a scene with no recognizable objects, to avoid memory mechanisms. This display is called Land colour Mondrian because in the 1970s Edwin Land used a similar collage of papers in order to study the colour constancy phenomenon (Land, et al., 1971). An overview of Land's experiments and theory is given in Chapter 3.

In the second experiment the observer looked at a pattern of stationary and moving black and white squares. When in motion the recorded activity was again in the striate cortex and then in area of the prestriate cortex located more laterally, that was called human V5.
These two experiments show that also in humans there are other visual areas besides the striate cortex, and these areas are specialized in visual motion and colour. It is important to note that in both the experiments the same part of V1, and V2 was active. The conclusion is that motion and colour visual signals reach V1 that spread them to the two specialized areas V4 and V5. Possibly this functional specialization is the approach adopted by the brain in order to gain the necessary knowledge about the constant properties of objects. Further demonstration of this hypothesis derives from the fact that damages to specific cortical areas produce specific visual syndromes. For example damages to area V4 of the brain lead to achromatopsia, a condition in which the patient lose the ability to see the world in colour (Pearlman, et al., 1979), (Damasio, et al., 1980), (Zeki, 1993). However it is not clear how the specialized areas interact in order to produce an unique image. It has been not possible to identify a dominant area that synthesises all the information coming from the different areas. One hypothesis is that all these specialized areas create a network that allows a kind of communication among the cells across the different areas. The synthesis of the visual image is the result of the continuing activity of this network.

2.4 Colour appearance phenomena

The previous sections have summarized a range of theories relating to vision. The objective was not to give a comprehensive explanation of the human visual system, but to demonstrate an understanding of some important issues regarding it. The process of vision starts from the eyes and it is realized in the brain. In section 2.3.2, Hering proposed his opponent colour theory in order to explain visual phenomena such as afterimages. This phenomenon is considered as part of field of optical illusions. In general an optical illusion is something that deceives our mind. We tend to be fascinated by the way they are able to cause a misinterpretation of a visual scene. More importantly optical illusions can reveal important mechanisms of perception. They can be arranged according to: motion, perspective, geometry, colour and luminance (Gregory, 1968), (Robinson, 1972), (Kitaoka, 2010). Some of the latter can be referred to as colour appearance phenomena. The basic colorimetry, as represented by CIE XYZ or CIELAB colour spaces, is not able to explain them. Basic colorimetry in fact provides a basic measurement whether two stimuli with different spectral distributions match

under given illumination conditions (Wyszecki, 1973). The specification of a colour is founded on the interaction of light source, object and observer. Colours are observed in isolated viewing conditions. If two identical stimuli are viewed under particular illumination conditions or in a certain context, the appearance of the stimuli will change. Same stimuli in different complex environments will appear different. Surround and background affect the colour stimulus. To solve this issues, the CIE (Commission Internationale de l'Eclairage) develop the advanced colorimetry (Wyszecki, 1973). The idea is to extend the basic techniques in order to incorporate specification about the appearance. The colour appearance models (CAMs) are therefore developed in order to fit an experimental dataset created with the intent to study a specific colour appearance phenomenon (Hunt, et al., 1975), (Breneman, 1987), (Helson, et al., 1952).

Regardless of the CIE another model rose up from the Retinex theory developed by Land and McCann (Land, 1977) (Land, et al., 1971). The aim of the models both developed by the CIE and developed starting from the Land's theory is to simulate some of the colour appearance phenomena, mimicking the way the human visual system (HVS) operates. Although the objective is the same they are very different (see chapter 3 and chapter 4).

2.4.1 Simultaneous contrast

In the following section some of these phenomena are described. One of the most known is that of simultaneous contrast (Chevreul, 1854 p. 30), (Hering, 1964), (Albers, 1963). Two identical grey patches presented on different backgrounds appear to be different. The patches on the white background appears darker and the patches on the black background appears lighter (figure 5) This effect, that follows the opponent theory of vision, take place also in the chromatic version. In fact, a green background induces the stimulus to have a reddish component and vice versa, as a yellow background induces a bluish component and vice versa (figure 6).

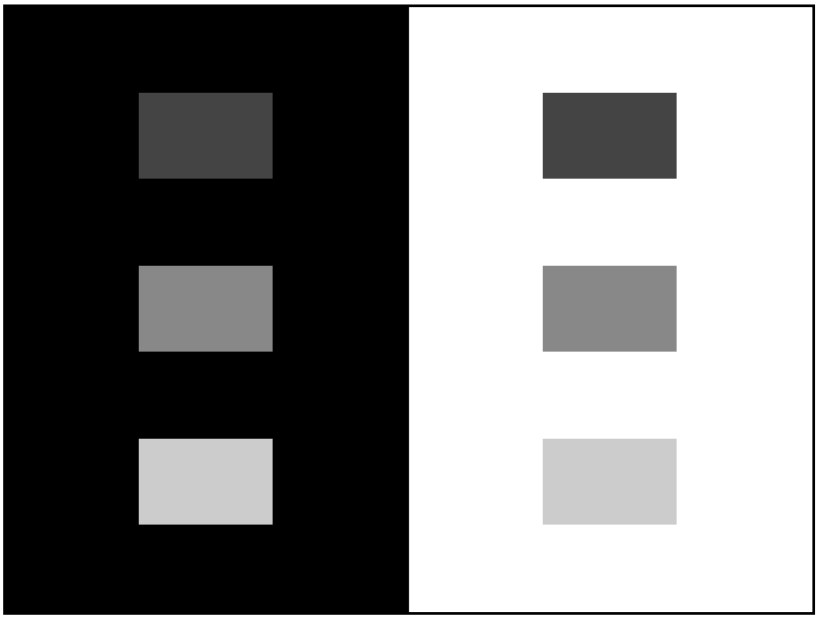


Fig. 5: Examples of simultaneous contrast. They grey patches on the right side are identical to the corresponding patches to the left, but they appear darker.

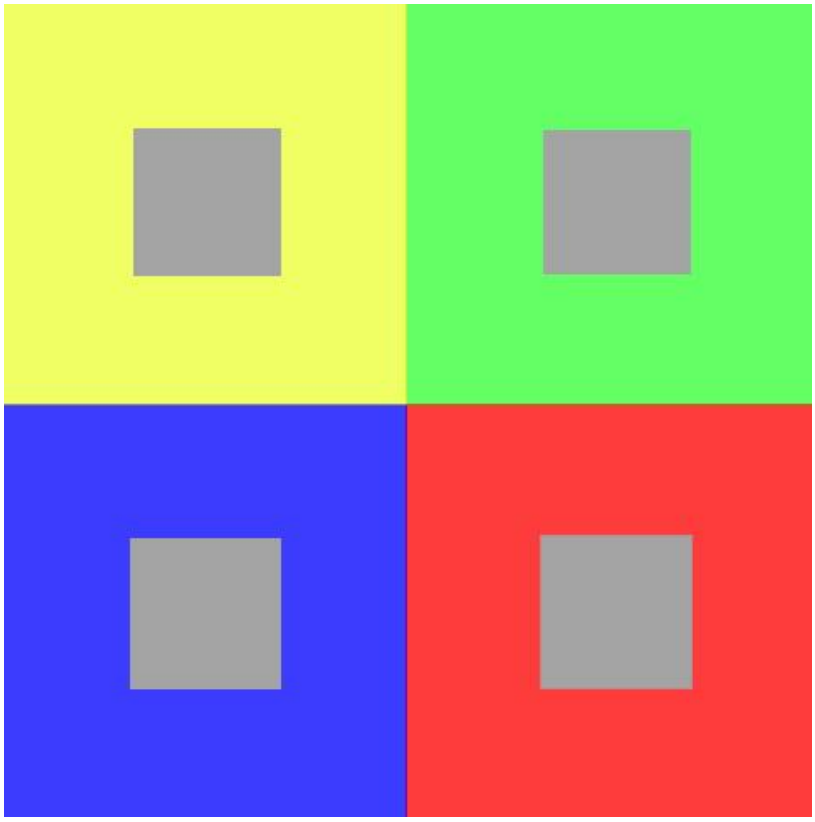


Fig. 6: Simultaneous contrast on chromatic background.

During the 1960-70s Josef Albers undertook experiments on the interaction of colours. In figure 7 there is a digital reproduction of the cover of the revised edition of his book (Albers, 1963). In the image the simultaneous contrast effect takes place, in fact the two squares have the same colours, but they appear different.

Fig. 7: digital reproduction of the simultaneous contrast described by Albers.

The effect can be increased using a background with a different geometry (Somers, et al., 1997). In figure 8, called snake illusion (Somers, et al., 1997), the rotated squares are all of the same grey, but they appear different. It is possible to see how the geometry can influence the appearance. Spatial and colour variables are closely related.

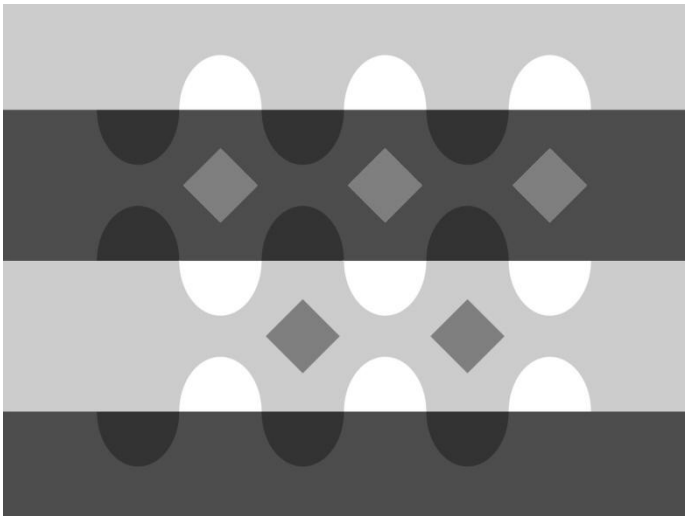


Fig. 8: Optical illusion showing how the geometry can influence the appearance.

In figure 9 it is possible to observe the chromatic version created by Akiyoshi Kitaoka (Kitaoka, 2010).

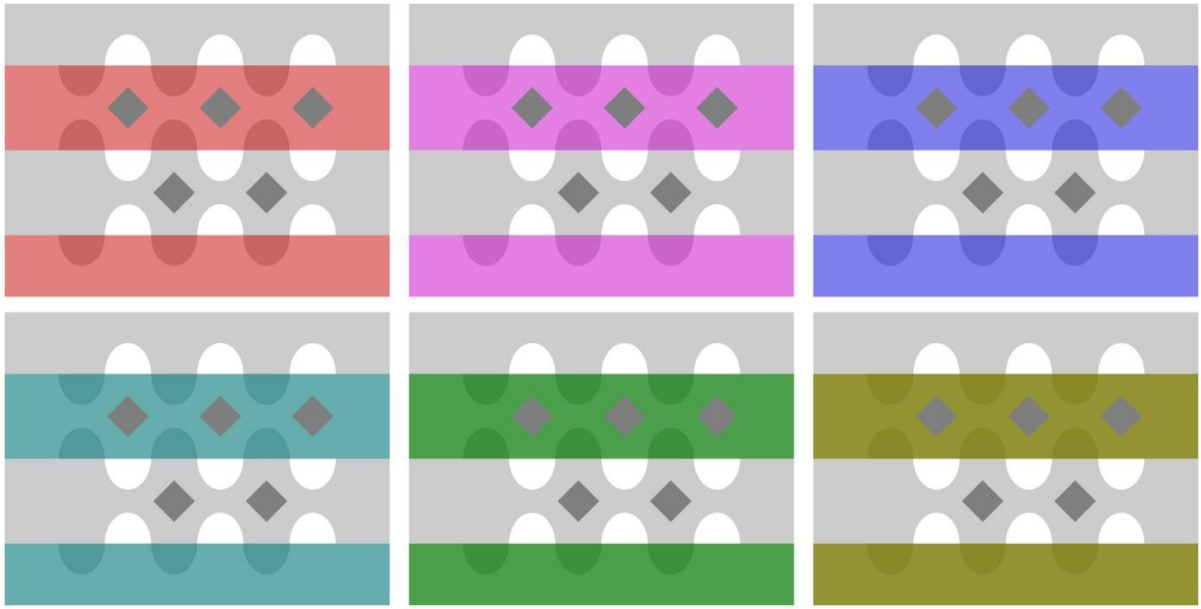


Fig. 9: chromatic version of the snake illusion.

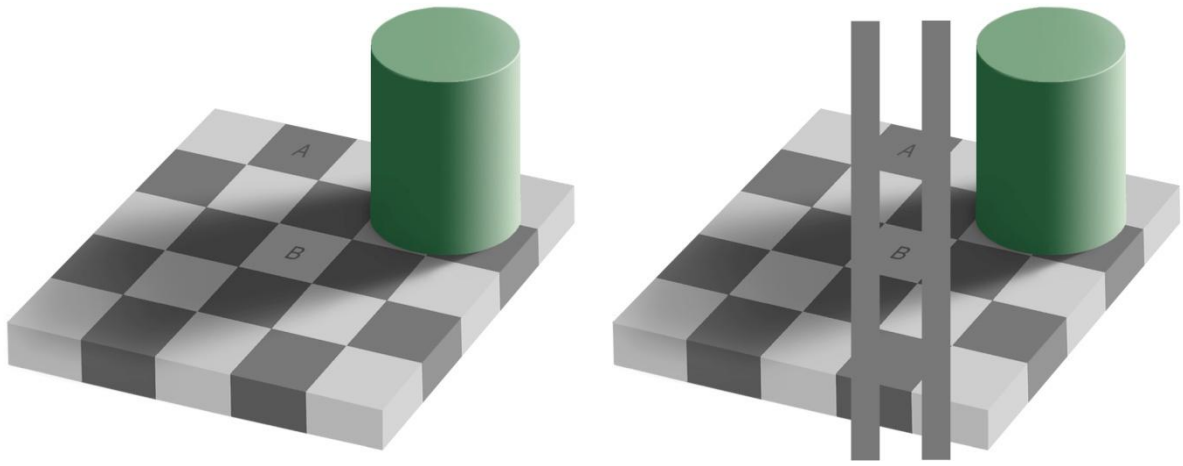


Fig. 10: Adelson's illusion.

Figure 10 is called Edward H. Adelson's "Checker–shadow illusion" (Adelson, 1995). The two squares denoted with the A and B letters are of the same grey but A appears darker than B. In order to explain this phenomenon we have to consider some aspects of the human visual system (Adelson, 1995), (Somers, et al., 1997). The first aspect relates to local contrast: B is light if compared to its neighbours, as A is dark because of the light surround. Furthermore the shadow of the cylinder has a soft edge, whereas the checks have sharper edges. The visual system partially ignores a light gradient, in order to determine the colour of a surface.

2.4.2 Afterimages and Daw experiment

Looking at a red patch for some time, and moving the gaze to a white surface, like a wall, one will see a greenish patch. This is called afterimage and was illustrated by Goethe at the beginning of 1800s (1840 p. 22). Although the debate about the explanation of afterimages is still active, they are likely to be a retinal phenomenon: cones adapt to the stimulation and become saturated, loosing sensitivity for that colour. When looking to a white surface the saturated cones send to the brain a weaker signal than the others, and thus the opposite colour appears (Geisler, 1978). In figure 12 it is possible to see this effect.

An interesting illusion based on the same mechanism is visible in figure 11: fixing the centred point in the upper side of the image for 20/30 seconds and moving the gaze to the centre point of the lower image, the picture of Venice will become as uniformly coloured.

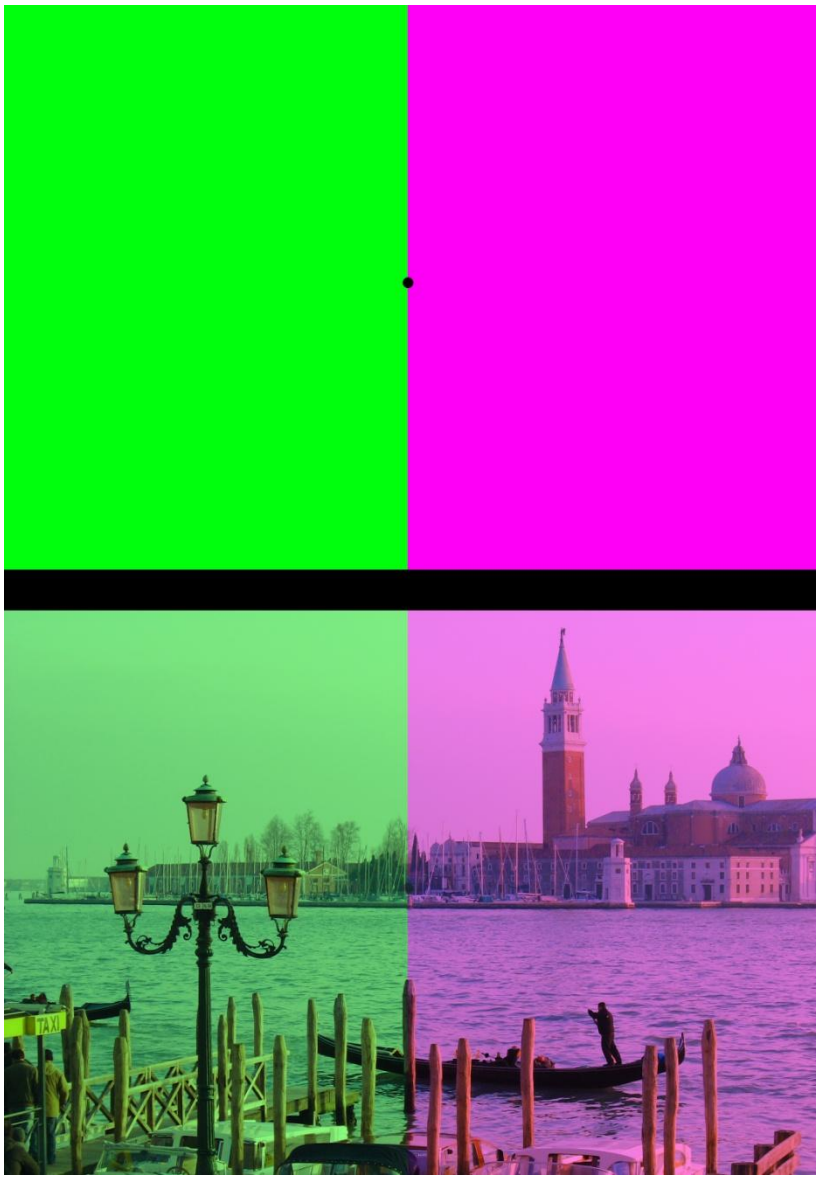


Fig. 11: an illusion based on the after image.

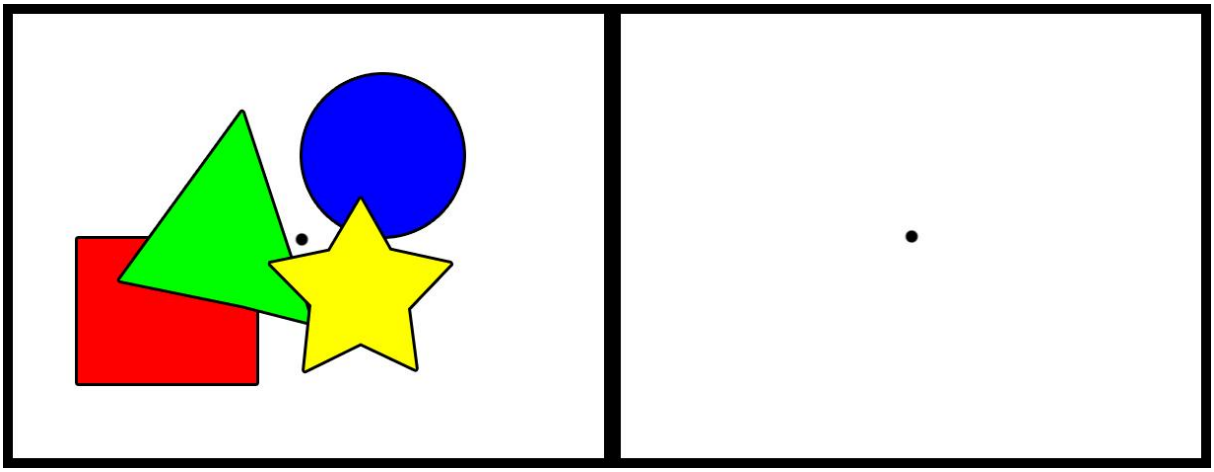


Fig. 12: classic afterimage: fix the centre dot on the left side for 20/30 seconds, and then move the gaze to the dot on the right side: the same figures appear, but as opposite colours.

In 1962 Nigel Daw realized an experiment that underlined how spatial mechanisms of edge preservation interacts with afterimages. In fact, the effect of an after image is stronger if the image is aligned to a monochromatic image with the same contours and that disappears when there is an edge mismatch (Daw, 1962). In figure 13 the same shape of figure 12 are used, but the right blank side is substituted by the only edge of the figures of the left part.

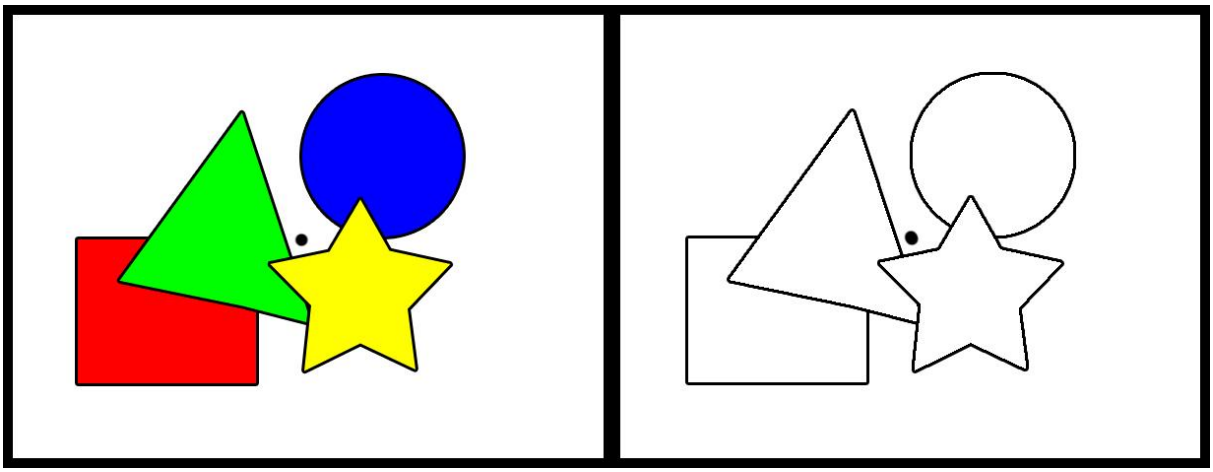


Fig. 13: the Daw experiment.

Fixing at the centre dot of the left part, and then moving the eyes to the centre dot of the right part composed by the contours, gives a stronger afterimage. Moving the gaze out from the centre dot makes the after image disappears but looking back again at the centre dot of the scene composed by the contours, the afterimage appears again.

2.4.3 Colour constancy

Another important phenomenon relates to colour constancy (Land, et al., 1971), (Land, 1977). For example a piece of white paper that is illuminated by the a tungsten-filament lamp, will appear white even if it reflects a different amount of energy with a yellow cast. Taking a picture of that paper with a day-light film will actually reveal the yellow as dominant, but our visual system is able to perceive the paper as white. This ability oh the human visual system to perceive the colour of the objects almost the same despite changing of the illumination condition is called colour constancy. In the following images (figure 14) the phenomenon is simulated in this way: the pillow on the left in figure 14(a) has the same RGB values in figure 14(b) but they appear different. Figures 14(a) and 14(b) represent the same scene, but the second is seen under *natural* light yet. They appear to be the same scene.

Fig. 14: images that represent the colour constancy effect.

2.4.4 Crispening

Crispening is an increasing in perceived colour difference if the background of the stimuli is similar to the colour of the stimuli themselves (figure 15) (Semmelroth, 1970).

Fig. 15: The crispening effect: A1, A2, A3 have the same RGB values. B1, B2, B3 have the same (different from A1, A2, A3) RGB values. However the colour difference between A2 and B2 is perceived greater than the difference between A1 and B1, or A3 and B3.

The effects presented above demonstrate the most obvious appearance phenomena, but there are other less known effects which are sometimes taken in account to develop a colour appearance models. A CAM, in fact, is developed in order to simulate one or more of these phenomena (Fairchild, 2005 p. 111-133). It doesn't exist until now a single CAM able to predict all the colour appearance phenomena. A range of other effects can be found in the book "Color Appearance Models" (Fairchild, 2005 p. 113-133). A table listing CAMs and the effects they are able to predict is provided in the same text (Fairchild, 2005 p. 281).

2.5 Conclusions

Colour research has a long history. In this chapter we have summarized some steps in the colour research field. The most famous theories, originated in the 18th century, are the trichomacy theory, by Young-Helmholtz, and Hering's colour-opponent theory. They seems

to be antagonist each other but in recent years the two theories have been combined in the direction of explaining how the human visual system works: they are set to refer to different stages of the vision. At the end of 1900 Zeki showed the complexity of the way the brain processes the visual signals. He underlined how the brain is able to acquire an invariant knowledge about certain constant characteristics of the objects. From his studies on Macaques, he discovered that different areas of the brain respond to specific vision stimuli, as shape, motion and colour. His theory is confirmed by the fact that damages to specific areas of the brain produce specific visual syndromes.

Finally a range of colour appearance phenomena and colour illusion are presented. Besides the fact that they are a fascinating subject of study, visual illusions are the key to understand the visual system. In chapter 7, two specific colour appearance phenomena, simultaneous contrast and colour constancy, are used to test two colour appearance models. The purpose of the two algorithms considered, the Random Spray Retinex and iCAM06, is to simulate the behaviour of the human visual system in these situations.

3. Retinex based models

3.1 A description of Land's Retinex theory of colour vision

Since the 50s Edwin Land undertook experiments in order to study the colour constancy phenomenon. In (Land, 1959) he took two pictures of coloured scene, composed by some fruits, using a black and white film. The first picture was taken placing in front of the camera a green filter, for the second a red filter was used. Therefore he projected the two black and white transparencies (slides) by superposition. The picture taken with the red filter was projected with a red filter, the picture taken with the green filter was projected with a neutral filter. Surprisingly enough the image displayed on the screen showed many colours, not only white, red and pink as it would be expected (a scheme of the setup of the experiment is shown in figure 16). In these image peppers appear to be green, strawberry red, lemon and bananas yellow.

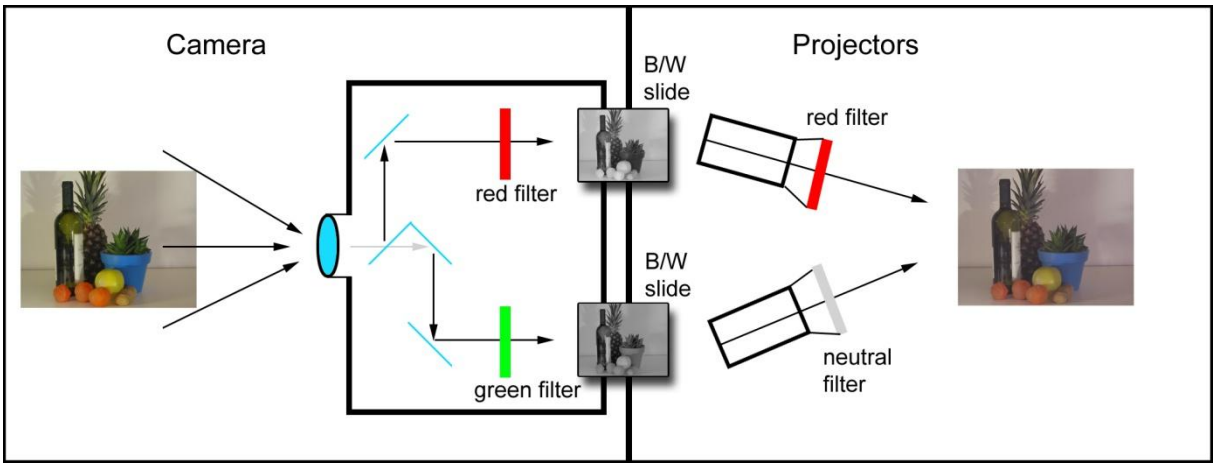


Fig. 16: set up of the Land experiment.

For more than 20 years Land persisted in his study, remarking how these colours appear outright and therefore were not due to some adaptation of the eye. In his 1977 paper (Land, 1977) he asserted that if a daylight film is used to take a picture of a scene illuminated by a tungsten-filament lamp, the photo will show a reddish cast. Differently from the picture, we can vary the illumination of a scene without almost perceiving the changing of the objects colour. The human visual system has evolved to see the world in unchanging colours,

independently from the illumination conditions. He formulated his theory explaining that the perception of a colour does not depend on the wavelengths reflected from that colour. In order to demonstrate his theory, he designed some experiments using collage of shaded grey or coloured papers, that has been called Mondrian. In the first experiment he placed a black piece of paper on one side of a collage, and a white one, on the other side, with many other pieces of grey shaded between this two. It is possible to position a light source in a way that from the black surface more light is reflected than from the white one, but the black still appears black, and the white appears white. In the second experiment, Land and McCann created two big coloured Mondrian, of about one hundred of coloured patches. Both were illuminated with the projectors with a narrow band pass filter, with peaks in 670, 540 and 450 nm. A separate variable transformed is used to control the amount of light coming out from each projector. A photometer was used to measure the reflected radiation of any point on the Mondrian. Suppose to illuminate the left Mondrian with the three projectors in order that the colours appear "natural". Then, one projector at time, the reflected energy of one particular patch, for example white, is measured with the photometer. Turning off the projectors illuminating the left collage, the projectors at the right are adjusted separately in a way that one test patch on the right collage reflects the same three triplet of energy than the white patch on the left. Enlightening now both the Mondrian, the left patch is white while at the right side, the test patch is green (figure 17). Nevertheless both the surface reflect the same triplet of energy. It is possible to illuminate other areas, with different colours, obtaining the same result. This lead to the conclusion that the sensation of colour is not correlated to the product of reflectance by illumination. The first response of the visual system is the absorption of the light from the receptors in the retina, but the processing of the image take place somewhere between the retina and the cortex, so the term Retinex.

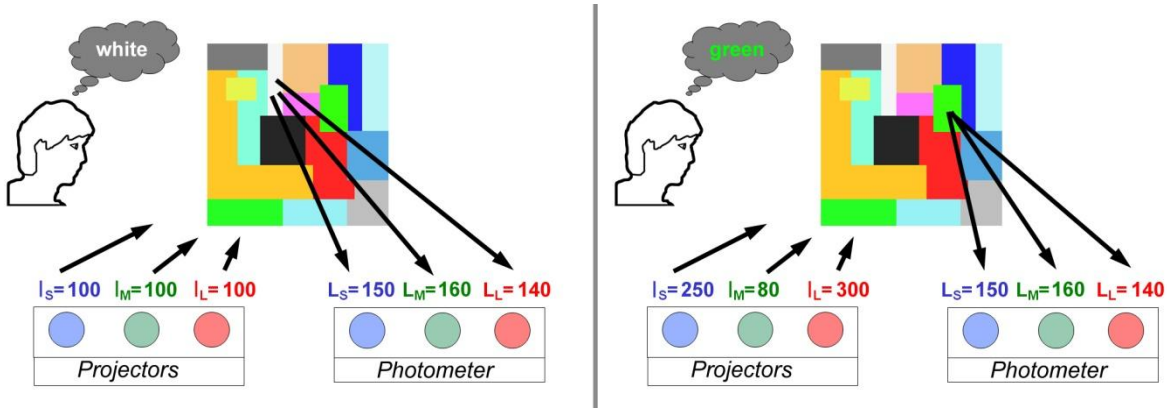


Fig. 17: the Mondrian experiment (the values do not have physical meaning)

In his experiments he noticed that illuminating the collage with one projector and increasing or decreasing the intensity of light the different patches keep a constant lightness scale, but changing projector many pieces of papers will occupy a different position in the lightness scale. That means that, for example, a red surface appears light under the long-wave illumination but very dark under the short wave, while a blue paper will appear light under the short wave illumination and dark under the long-wave. So the three kind of cones form three different images of lightness. These are not mixed but are compared. This comparison give us the sensation of colour.

Land underlined also the fact that the edges play a fundamental role in the definition of the objects and surfaces in a scene. Illuminating a white paper from one side, no colour discontinuity is observed. The luminance of two points on the paper, will be different, but as the two points are taken closer one to the other the luminance will tend to approach, and their ratio tends to one. If the points are taken on two different coloured patches the ratio of the luminances will tend to the ratio of the two reflectances. With this procedure it is possible to detect an edge and eliminate the effect of a non uniform illumination. For example imagine four points (A, B, C, D) taken on a surface split in two parts: one grey (with reflectance R = 30), near a white one (with reflectance R = 100) illuminated with a source of light placed obliquely. The luminance (L) reaching the eye (that is the quantity we can measure when taking a picture) is: $L = R * I * \cos\theta$, where θ is the angle between the direction of the light and the normal to the surface (θ is renamed α, β, γ depending on the point on the surface we are considering, see figure 18). Consider the cosine of α, β, γ as 0.7, 0.8, 0.9 respectively.
The ratio between two points A and B, that is two points on the same surface, will be:

$$Ratio_{AB} = \frac{30 * I * 0.7}{30 * I * 0.8} = 0.875 \qquad (3.1)$$

The ratio between B and C, that is two points near the edge (we can assume the cosine of the angle is the same), will be:

$$Ratio_{BC} = \frac{30 * I * 0.8}{100 * I * 0.8} = 0.3 \qquad (3.2)$$

From equations 3.1 and 3.2 it is clear that, taken two points that are near on the same surface, the ratio of their luminance will be close to 1, and the ratio between two points near an edge will be equal to the ratio of their reflectance, independently from the illumination I.

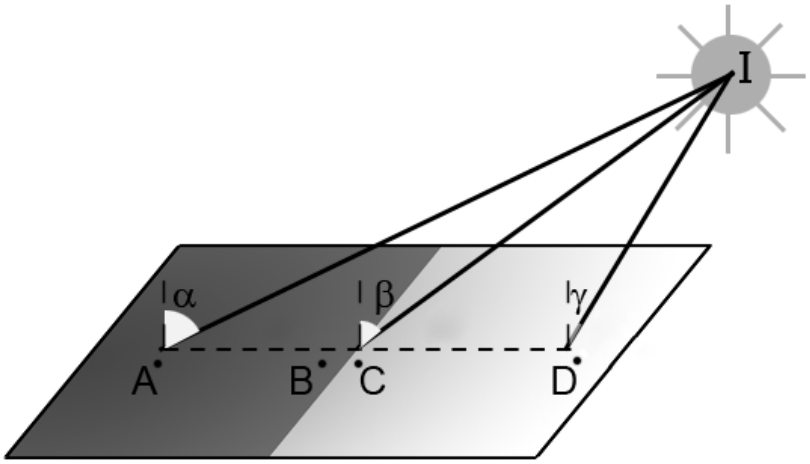


Fig. 18: θ is the angle between the direction of the light and the normal to the surface. In this figure it is renamed α, β, γ depending on the points (A, B/C and D) on the surface.

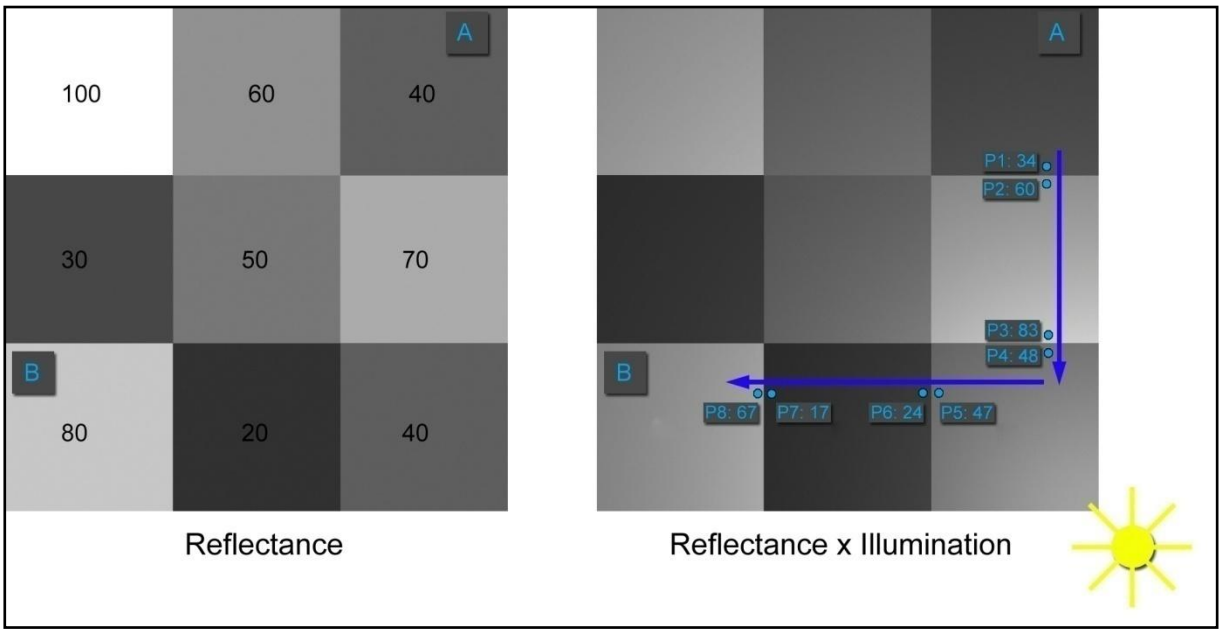
Processing the entire image in terms of ratios of luminance implies obtaining the ratio of reflectance at the edge between adjacent areas. In figure 19 (a) it is possible to see the reflectance of nine patches. In figure 19 (b) a non uniform illumination is applied to the image. The ratio between the luminance of two points (in the non uniform illuminated image) P1/P2 = 34/60 = 0.5666 rounds the ratio between the reflectance of the surface A (where P1 lies) and the reflectance of the near surface (where P2 lies): 40/70 = 0.571.

Consider now the surfaces A and B, the ratio between the reflectance of the two areas is 40/80=0.5.

Multiplying the ratio of luminances between the points on a path from A to B, that lie near the edge (figure 19 (b)), it is possible to approach the ratio of reflectance between A and B, regardless of the distribution of illumination:

$$\frac{P1}{P2}\ x\ \frac{P3}{P4}\ x\ \frac{P5}{P6}\ x\ \frac{P7}{P8} = \frac{34}{60}\ x\ \frac{83}{48}\ x\ \frac{47}{24}\ x\ \frac{17}{67} = 0.4868 \cong 0.5 = \frac{reflectance\ A}{reflectance\ B} \qquad (3.3)$$

Therefore it is possible to obtain the ratio of reflectances between any two areas in a scene multiplying the luminances ratio of all the boundaries between the first and the last area.



(a) (b)

Fig. 19: (a) reflectance of an image composed by nine patches. (b) one possible path between the surface A and B. It is shown the luminance (RxI) of every point on the path situated near the edge of two adjacent surfaces.

Many arbitrary path are used to compute sequential products, that can begin anywhere. So the first value in any sequence is set to 100 %, and therefore as soon as the path reaches an area whose reflectance is higher than that of the starting area, the sequential product becomes greater than one. This new area is assumed to have a reflectance of 100% and the sequence should start again. With this technique (called *reset mechanism*) it is possible to find the highest reflectance in the path.

The white sensation will be given by any surface placed at the top of the lightness scale by all three Retinex system. Once the white is obtained, the other colours are computed accordingly, as showed in the next sections.

3.2 Retinex algorithms

Edwin Land's theory led to the development of some different algorithms that try to simulate somehow the behaviour of human visual system. The term spatial colour refers to a family of algorithms that recompute the colour of each pixel arrays according to the spatial distribution of pixels values in the scene. Spatial distribution describes how the colour appearance of an area is related to the context, that is the colour distribution around that area.

From the Retinex theory some points evolve. Often this theory is used as a starting point to develop new computational models. The algorithms that belong to this family attempt to model colour functions of the human visual system. In particular their task is to reproduce the colour constancy behaviour. From the Retinex point of view, human vision is based on three retinal cortical systems, each independently processing the low, middle and high frequencies of the visible spectrum. Each system processes a separate image determining on the relative lightness values of the various areas of the scene. Every region of the image is compared with an adjacent area and the lightest area is considered as the local white. Retinex models compute the lightness of each pixel in the image using data collected from other surrounding pixels. These models incorporate the concept of locality that considers the data near to the computed pixel and not an averaged value of the whole data of the input image. There are many algorithms based on the Retinex theory, all sharing some common points: the computation is undertaken independently on different cone responses, the computation is mainly based on a comparison made by difference or ratios, relating to the concept of spatial colour.

3.2.1 Random Spray Retinex

In order to achieve the locality of colour perception (the fact that colour of an area is influenced by the chromatic content of the areas surrounding it), the first implementations of the Retinex algorithms were based on paths scanning an image. As has been shown in section 3.1, the Retinex model processes the entire image in terms of ratio of luminance along a certain path. The essence of the Retinex algorithm is the *reset mechanism* that is responsible for the white patch behaviour (that means searching for the lightest patch to use as a white reference). If, during the computation of the ratios on the path, the cumulated product is greater than 1, the *reset mechanism* set it equal to 1, since nothing can be whiter than white. Using paths leads to some problems: strong dependency on paths' geometry, high computational cost, and noise sampling. Edoardo Provenzi et al. (2007) proposed a method consisting in replacing paths with a cloud of points (called random spray) across the image (figure 20). He observed, in fact, that the pixel which caused the reset mechanism, is considered as the local white reference. Therefore the problem is reduced to find the pixel with maximum intensity value for every path, and this is independent from the geometry of the paths.

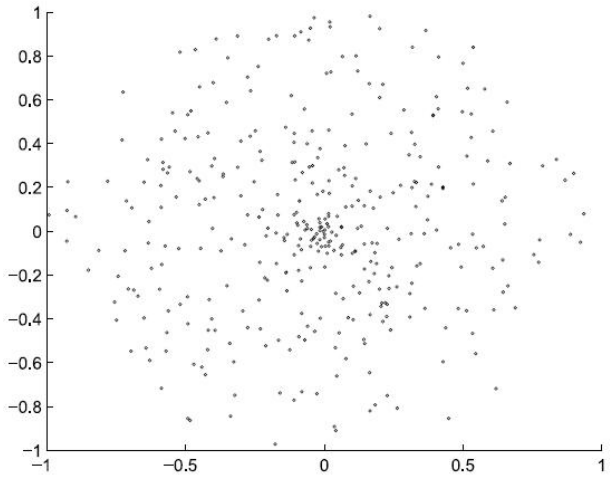


Fig. 20: Example of a spray. Note that the density sampling is higher near the target pixel (placed in 0,0).

Provenzi investigated the possibility of using areas with a density sample depending on the distance from the pixel he is considering (called *target*) (figure 20). The result is the Random Spray Retinex (RSR). In order to avoid noise, the original algorithm implemented with paths used N paths terminating all on the target pixel. In RSR the N paths are replaced with N sprays, therefore a final average between the spray contributions is needed. For every pixel of the image the new value of the pixel is calculated (independently for the three channels: red, green and blue) by the formula:

$$L_c^{new}(t) = \frac{1}{N}\sum_{k=1}^{N}\frac{L_c(t)}{max_{c,k}} \tag{3.4}$$

Where $L_c^{new}(t)$ is the final value of the target pixel t, for the channel c (red, green or blue), $L_c(t)$ is the initial value of the target pixel and $max_{c,k}$ is the maximum value for the channel c found inside the k-th spray.

3.2.2 Automatic Colour Equalization

ACE (Automatic Colour Equalization) is another spatial algorithm model proposed by Gatta et al. (2002) It reproduces a grey world behaviour which assumes that the average colour in the scene is the middle grey, simulating the mechanism of lightness constancy. This mechanism allows the human visual system to adapt to different lightness levels in a scene. The model was developed starting from low level mechanisms of colour vision. The locality is achieved using a weighting function. This function attempts to balance local and global

behaviour of the model. The colour sensation of a pixel target is more influenced by neighbouring. Unlike the previously described Retinex algorithms, ACE is based on the difference, instead of ratio, as comparison method (Gatta, et al., 2002), (Rizzi, et al., 2003). In ACE the white patch behaviour is achieved by modifying the difference between pixel values, with a non linear function, that computes the local lightness of pixel *p* respect to pixel *j*.
The image in input is normalized between 0 and 1. Thereafter ACE is performed in two steps:

- Recompute the value of every pixel *p* according to the values of the pixel of the whole image. This gives as result an intermediate image R for every channel *c*.
- Map the image R in an image O displayable by the monitor.

The formula that implements the first step is:

$$R_c(p) = \frac{\sum_{\forall j \neq p} \frac{r(I(p) - I(j))}{d(p,j)}}{\sum_{\forall j \neq p} \frac{1}{d(p,j)}} \tag{3.5}$$

Where *p* is the pixel we are considering, *I(p) – I(j)* is the difference that account the way to compare pixels. *d* is a distance function which calculates the distance between *p* and *j* (for example *d* calculates the Euclidean distance between the two pixels). *r* is the function that determines the relative appearance between two pixels *i* and *j*, in a range between -1 and 1.
In particular:

- $r(I(p) - I(j))$ = -1 means that *i* is perceived black respect to *j*
- $r(I(p) - I(j))$ = 1 means that *i* is perceived white respect to *j*

Choosing a threshold *thr* the function is calculated as follows:

$$r\big(I(p) - I(j)\big) = \begin{cases} -1 & if \quad I(p) - I(j) \leq -thr \\ (I(p) - I(j))/thr & if \quad -thr \leq I(p) - I(j) \leq thr \\ +1 & if \quad I(p) - I(j) \geq -thr \end{cases} \tag{3.6}$$

The graph of the function is shown in figure 21.

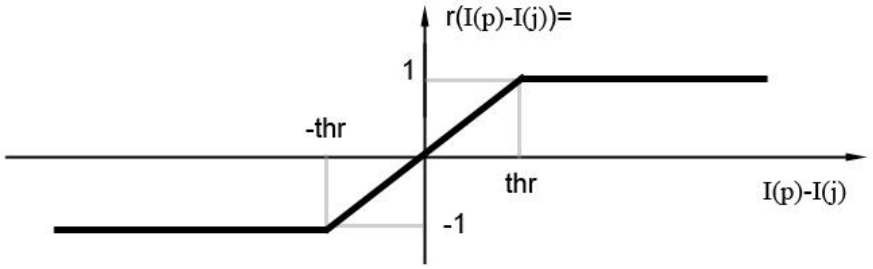


Fig. 21: graph of the r function.

It is now necessary to map the values just obtained into a range displayable by the monitor (step 2). The simplest method consist in linearly scaling the range of computed values into the dynamic range of the output. If we consider an image with 8-bit per channel c, therefore the red, green, blue channels have a range between 0 and 2^8-1=255 values, the linear mapping is provided by the following equation:

$$s_c = \frac{255}{Max_c - min_c} \tag{3.7}$$

$$new\_value = round(s_c * (old\_value - min_c)) \tag{3.8}$$

Where Max_c and min_c are the maximum and minimum values for the intermediate image.
With this formula the minimum value is mapped as 0, and the maximum is mapped as 255.
An alternative technique to achieve the grey world property is given by the following formula:

$$s_c = \frac{127}{Max_c} \tag{3.9}$$

$$new\_value = round(127 + s_c * old\_value) \tag{3.10}$$

In this way the dynamic of the final image is centred around the middle grey, but some values can result negative. In this case the values lower than zero are forced to zero.
In its original form, ACE has a algorithmic complexity of the order of $O(N^2)$, where N is number of pixel of the image. To reduce the computational time a faster technique has been

implemented, based on Local Look Up Table, that gives similar results but in a fraction of time (Gatta, et al., 2006).

3.2.3 RACE

In 2008 Provenzi and Gatta fused RSR and ACE in order to produce an improved model, by taking the better features of the two algorithms. The result is RACE (Provenzi, et al., 2008). The idea started observing that the RSR and ACE output images exhibit mutually distinct defects. RSR shows good saturation properties and tends to keep the original photographic key of the image in most cases, but it has poor detail recovering, especially in dark zones. ACE is able to put in evidence details, particularly in dark regions, but it tends to corrupt the photographic key. Both RSR and ACE are able to remove colour cast, with individual properties: in particular RSR reveal a good saturation level, ACE, thanks to the presence of the slope function, increases the percentage of used dynamic range. The fusion of RSR and ACE into RACE allows combining the positive features of both algorithms resulting in an improved model with respect to the original ones. As the two algorithms implement two different way to compare pixel, one should be chosen. The authors decide to use the spray formulation, which requires fewer computations.

3.3 Conclusions

In this chapter we presented the Land's Retinex theory. The starting point of his theory is the statement that our vision system has evolved to perceive in the more stable way as possible the colour of the objects. This ability is called colour constancy. Since the 1950s, Edwin Land and the colleague John McCann undertook several experiments in order to study colour vision. In many of their experiments they used collages of shaded grey or coloured paper patches, called Mondrian, to study the characteristics of the human visual system. This configuration of colours is important to avoid memory effects. The Retinex theory explains that the perception of a colour does not depend on the wavelengths reflected from that colour, but from the comparison with the neighbour colours. Finally an overview on the algorithms based on the Retinex theory is given: Random Spray Retinex (RSR), Automatic colour equalization (ACE), and Race, the fusion between RSR and ACE in order to produce an improved model.

4. CIE Colour appearance model

4.1 Introduction

As mentioned in chapter 1 some colour appearance phenomena cannot be explained using only basic colorimetry. If two identical stimuli are viewed under particular illumination conditions or in a certain contest, the appearance of the stimuli will change. The term colour appearance is referred to the study of how a human observer perceives a certain colour stimulus. Many factors affect the appearance of a stimulus: the illumination conditions, the size and shape of the stimulus, the background etc.

In this chapter we will examine the appearance as intended by the CIE (Hunt, 1977), (TC1-34, 1997), (CIE, 1998), (Fairchild, 2005). Advanced colorimetry is needed in order to integrate concepts that the basic colorimetry is unable to explain.

In the previous chapter the Retinex theory and algorithms have been presented; calculating the appearance meant to simulate how the human visual system perceives a certain scene. For the CIE a colour appearance model is a model able to predict at least the relative colour-appearance attributes of lightness, chroma, and hue (TC1-34, 1997). The colour appearance attributes are examined in Appendix A. In section 4.2 an explanation about the viewing conditions is given. In 4.3 it is discussed about chromatic adaptation transforms, used to predict the colour appearance under disparate illumination conditions. In 4.4 basic concepts about CIE colour appearance models are presented. Finally in 4.5 and 4.6 the CIECAM02 and iCAM models are presented.

4.2 Viewing conditions

As stated in the introduction, the colour appearance of a stimulus is strongly related with the surrounding area. Two stimuli, with the same tristimulus values but different surround, will appear different. In order to use a colour appearance model it is necessary to define the viewing field, that is the environment in which a stimulus is viewed. The colour appearance model implemented by Hunt (Hunt, 1982) divides the viewing field in four components: stimulus, proximal field, background and surround (figure 22). The other models use the same set or a subset of these components. The stimulus is the target intended for measurement. In

1931, the standard colorimetric observer was used a stimulus of approximately 2° in size to estimate the colour matching function. Therefore a stimulus with the same size is taken. Around the stimulus there is the proximal field, extending outside the stimulus for about 2°. Only the Hunt model uses the proximal field, as it is often hard to specify it precisely. Normally it is considered the area around the stimulus, extending for about 10° from the edge of the stimulus: the background. In order to model the simultaneous contrast, it is necessary to specify the background. The area that is outside the background is the surround, which can be considered as the environment in which the stimulus is viewed. Usually it is categorized by a few situations: dark, dim and average. Everything outside the stimulus, therefore the set of proximal field, background and surround, is called the adapting field.

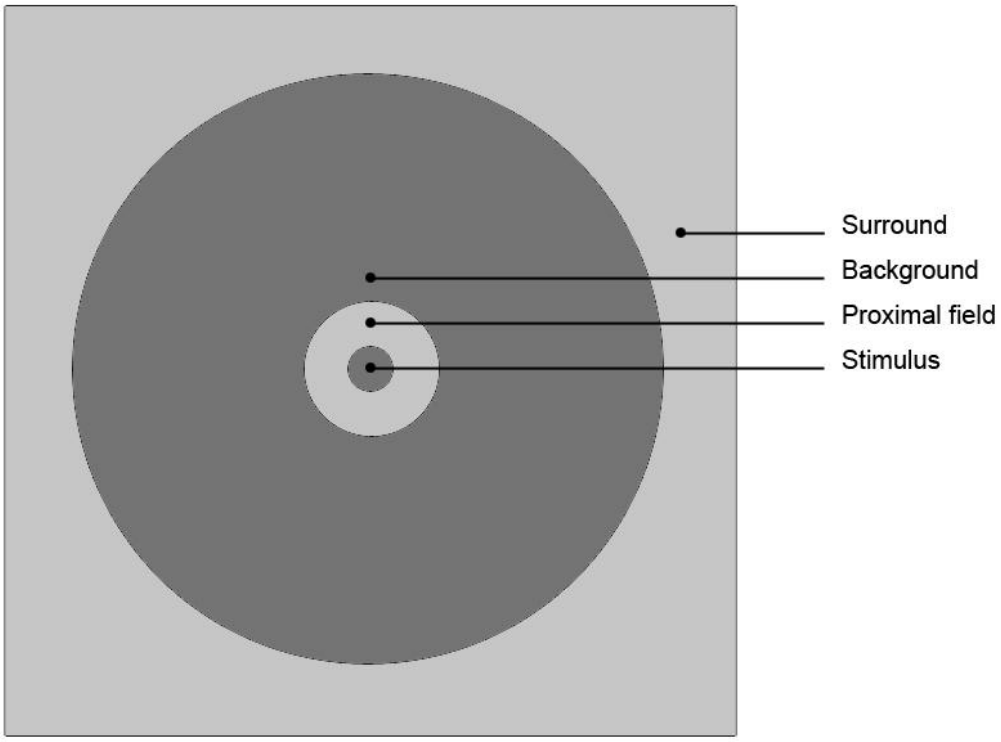


Fig. 22: The four components of the viewing field.

4.3 Adaptation and Chromatic Adaptation Transform

It has been already highlighted in chapter 2 that the human visual system is able to perceive the colours of the objects even under a change in the illumination condition. This ability is called colour constancy (Land, 1977). One explanation to this behaviour is known as *adaptation*, that is the capability to adjust the general sensitivity in order to preserve the appearance of the objects even if some conditions change. For example a traditional camera with a daylight film (or a digital version with a white balance set on D65 light) will take

pictures with a yellow cast if used in a room lit by an incandescent lamp. The human visual system however is able to reduce this effect and perceiving little change in the colours of objects. This is called chromatic adaptation (von Kries, 1902).

Light adaptation, instead, is a physiological mechanism and it is the decrease in visual sensitivity caused by the increasing of the level of illumination. Leaving a movie theatre in the day light, the visual system is dazzled, but after ten seconds, we adapt to see normally again. This mechanism makes the observer less sensitive to a stimulus when its intensity increases. The inverse phenomenon is the dark adaptation, which is when an the adjustment occurs if the illumination level is reduced, so the visual sensitivity is increased. Dark adaptation takes place at a slower rate (in order of minutes) than light adaptation.

Comparing these examples it is straightforward to note that our visual system does not act as a physical tool to measure the radiation but aims to obtaining meaningful information in order to perceive the invariant information about certain characteristics of the objects.

There are many physiological mechanisms that enable visual adaptation. One is the dilatation and constriction of the pupil, that allows more or less light to reach the retina. A second mechanism is the transition from cones to rods and vice versa. As described in chapter 2, cones are less sensitive and are active only with high levels of illumination, while rods respond to lower levels of illumination. The independent sensitivity change of the cones could explain chromatic adaptation. This mechanism is sometimes called receptor gain control. This gain control reduces the electrochemical signal produced by the photoreceptor when there are many photons, and increase the signal when there are few photons

In order to investigate and implement this phenomenon in a colour model some data are needed. Through a multiplicity of experiments, a variety of corresponding-colour data have been obtained (Hunt, 1952), (Helson, et al., 1952), (MacAdam, 1961), (Wright, 1981), (Mori, et al., 1991), (Fairchild, 1991), (Luo, et al., 1991a), (Luo, et al., 1991b), (Hunt, et al., 1994). The definition is straightforward: when two different stimuli, viewed under different viewing conditions, match in appearance, represent a pair of corresponding colours. Of course even the respective viewing conditions must be specified in order to build the data-set (figure 23).

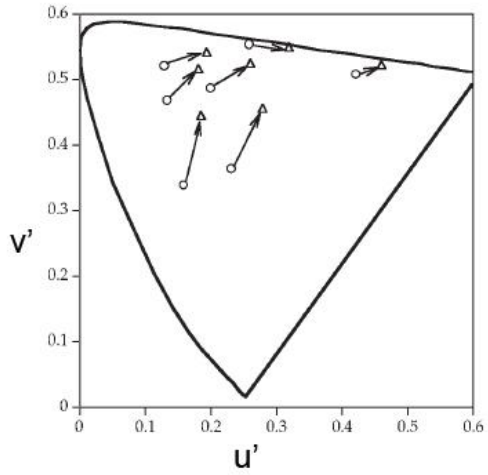


Fig. 23: Example of corresponding colours. Circle could represents the colours under the D65 illumination, triangles the corresponding colours under the A illumination.

A range of experimental methods are used: haloscopic matching, memory match and magnitude estimation.

- Haloscopic: refers to those experiments in which one eye is adapted to one viewing condition and the other one is adapted to a different viewing condition. Then two stimuli are presented separately to the eyes in order to verify if there is matching (Mori, et al., 1991).
- Memory matching: the observer looks to a stimulus in a viewing condition, then adjust another stimulus with different viewing condition until they match (Helson, et al., 1952), (Wright, 1981), (Fairchild, 1991).
- Magnitude estimation is an experiment that assigns a value to various type of appearance attributes as lightness, chroma, hue etc (Luo, et al., 1991a), (Luo, et al., 1991b), (Hunt, et al., 1994).

Once a colour appearance model is implemented this data can be used to compare the model prediction with the visual result.

The first step towards the building of a CAM is the chromatic adaptation transform (CAT).

A CAT is able to transform the tristimulus values observed under certain viewing conditions to matching tristimulus values observed in a second set of viewing conditions. Therefore a CAT can be used to make predictions of corresponding colours, but does not provide estimation for the appearance attributes like lightness, chroma, hue etc. In the literature there are many type of chromatic adaptation transform (von Kries, 1902), (Lam, 1985), (Luo, et al., 1998), (Li, et al., 2002) but they all follow the same idea that can be summarized in the following steps:

- Transform the CIE tristimulus values (XYZ_1) into a cone response domain (LMS_1).

- Predict the adapted cone signals (LMS_a) based on the initial viewing conditions (VC_1).
- Determinate the cone excitation (LMS_2) according to the second set of viewing conditions (VC_2).
- Back to tristimuls values (XYZ_2).

LMS stands for Long, Medium and Short wavelength, referred to cone excitations. Normally the tristimulus values XYZ are converted in terms of cone excitations LMS, to better model the chromatic adaptation. This can be done using a 3x3 matrix, as exampled in the Hunt-Pointer-Estevez matrix (Fairchild, 2005 p. 177, 206, 211).
The basis for the formulation of the chromatic adaptation transform were presented by von Kries (von Kries, 1902). He didn't explicitly write the equations to build a CAT but proposed a concept that today is still valid. From these ideas arise the von Kries model. He suggested that the cones adapted themselves independently.
The interpretation of the von Kries ideas can be transposed in the following equations:

$$L_a = k_L L \tag{4.1}$$
$$M_a = k_M M \tag{4.2}$$
$$S_a = k_S S \tag{4.3}$$

L, M, S represent the initial cone responses to a certain stimulus, while L_a, M_a, S_a are referred to the post adaptation cone responses. To obtain the adapted cone signals, L, M, S are multiplied by k_L, k_M, k_S, which are the independent gain controls. These coefficients are taken to be the inverse of the maximum L, M, S cone responses in the scene, that typically is the white of the scene.

$$k\_L = 1 \,/\, L\_max \tag{4.4}$$
$$k\_M = 1 \,/\, M\_max \tag{4.5}$$
$$k\_S = 1 \,/\, S\_max \tag{4.6}$$

It is now straightforward to transform the XYZ tristimulus values from a viewing condition to another one, calculating therefore the corresponding-colours. In order to reach this objective it is necessary to calculate the post-adaptation signals for the first and the second conditions, and make them equal. For the L cone:

$$L_a = k_{L1} * L_1$$
$$L_a = k_{L2} * L_2$$

$$k_{L1} * L_1 = k_{L1} * L_1 \quad \rightarrow \quad \frac{1}{L_{1\_max}} * L_1 = \frac{1}{L_{2\_max}} * L_2 \quad \rightarrow \quad L_2 = \frac{L_1}{L_{1\_max}} * L_{2\_max} \qquad (4.7)$$

Therefore, the formula to calculate corresponding colours can be written in matricial notation:

$$\begin{bmatrix} X_2 \\ Y_2 \\ Z_2 \end{bmatrix} = M^{-1} \begin{bmatrix} L_{MAX2} & 0.0 & 0.0 \\ 0.0 & M_{MAX2} & 0.0 \\ 0.0 & 0.0 & S_{MAX2} \end{bmatrix} \begin{bmatrix} 1/L_{MAX1} & 0.0 & 0.0 \\ 0.0 & 1/M_{MAX1} & 0.0 \\ 0.0 & 0.0 & 1/S_{MAX1} \end{bmatrix} M \begin{bmatrix} X_1 \\ Y_1 \\ Z_1 \end{bmatrix} \qquad (4.8)$$

Where M is the Hunt-Pointer-Estevez matrix and M^{-1} its inversion.

4.4 The CIE Colour Appearance Models

CIE Technical Committee 1-34 (TC 1-34) defined a colour appearance model as a model that includes predictors of at least the relative colour-appearance attributes of lightness, chroma, and hue (TC1-34, 1997). The aim of TC1-34 was to test the performance of various existing models. In 1997 the CIECAM97s (CIE, 1998) was designed as a hybrid of the best elements of many different previous models, including Hunt (Hunt, 1982), Nayatani (Nayatani, et al., 1987), RLAB (Fairchild, 1996) and LLAB (Luo, et al., 1996) models. CIECAM97 was recommended by TC1-34 to be used as a standard. Even if the final "s" stands for "simple", CIECAM97s is a rather complicated model (and CIECAM97c was never formulated).
Improvements to this model led to a new recommended model: the CIECAM02.
All the models cited share approximately the same steps:

- Input data: tristimulus values for the stimulus, data concerning the viewing environment, background, surround, light source.
- A transformation from XYZ to cone response, to better model the physiological processes of the HVS.
- A chromatic adaptation transform is performed. Von Kries is an example, but other are possible.
- Some operations in order to model the opponent theory of colour vision.

- Output: appearance attributes including lightness, hue and chroma at least.

In the following the formulation of CIECAM02 is reported.

4.5 CIECAM 02

4.5.1 Input data

In order to use this model some input data about the test stimulus and the condition in which is viewed are required.

The stimulus is specified as tristimulus values XYZ, like the tristimulus values of the white point X_w Y_w Z_w.

Other data regards the context in which the stimulus is viewed:

- Y_b is the luminance of the background,
- L_A is the adapting field luminance, expressed in cd/m^2, often taken to be 20% of the luminance of a white object (L_w) present in the scene, in order to incorporate the gray world assumption.
- A categorical surround, that can be dark, dim or average. Based on the surround some constant have to be chosen (table 1).

	c	N_c	F
Average surround	0.69	1.0	1.0
Dim surround	0.59	0.95	0.9
Dark surround	0.525	0.8	0.8

Tab. 1: Constants based on the surround.

The grey world assumption states that the average colour in a real scene is a middle grey. The 20% (approximation of 18%) comes from the fact that a grey paper reflecting the 18% of the light that it receives it is perceived as “middle”, that is equidistant between black and white. This is incorporated in the equation to find the Lightness in the CIELAB colour space:

$$L^* = 116\left(\frac{18}{100}\right)^{1/3} - 16 = 49.5 \tag{4.9}$$

4.5.2 The model

In order to better model some physiological processes of the visual system, the first objective is to transform the tristimulus values (XYZ) to the cone responses (RGB). In CIECAM02 this is performed using the $\mathbf{M}_{CAT02.}$

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = M_{CAT02} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \tag{4.10}$$

$$M_{CAT02} = \begin{bmatrix} 0.7328 & 0.4296 & -0.1624 \\ -0.7036 & 1.6975 & 0.0061 \\ 0.0030 & 0.0136 & 0.9834 \end{bmatrix} \tag{4.11}$$

Similar equations are used to calculate $R_WG_WB_W$ from $X_WY_WZ_W$.

At this point it is possible to perform the chromatic adaptation transform (CAT), that is the nucleus of a CAM. By doing this, it is possible to choose the degree of adaptation which is described as D. D can be set as equal to 1 if complete adaptation takes place (in this kind of models this is referred to as *discount the illuminant*), or 0 for no adaptation. Otherwise it can be calculated with the following formula:

$$D = F\left[1 - \left(\frac{1}{3.6}\right) e^{\left(\frac{-(L_a+42)}{92}\right)}\right] \tag{4.12}$$

L_A is the luminance of the adapting field, while F has to be chosen according to the surround, and can assume value of 1, 0.9, 0.8 for average, dim or dark surround respectively.

The following figure 24 illustrates a graph showing D at Dark, Dim and Average surround.

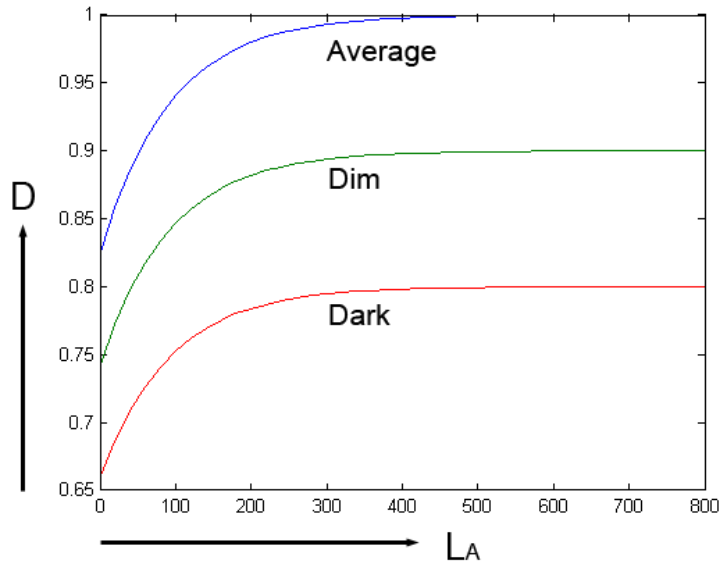


Fig. 24: Graph of D, for dark, dim, average surround

From equation (3.12) it is clear that D can assumes minimum (for L_A=0 cd/m^2) and maximum (for L_A>760 cd/m^2) values according with F as shown in table 2.

	F	Min D	Max D
Average surround	1.0	0.8240	1.0
Dim surround	0.9	0.7416	0.9
Dark surround	0.8	0.6592	0.8

Tab. 2: Minimum and maximum values for D.

Now it is possible to calculate the chromatic adaptation transform:

$$R_C = \left[\left(\frac{100D}{R_W}\right) + (1 - D)\right] R \tag{4.13}$$

$$G_C = \left[\left(\frac{100D}{G_W}\right) + (1 - D)\right] G \tag{4.14}$$

$$B_C = \left[\left(\frac{100D}{B_W}\right) + (1 - D)\right] B \tag{4.15}$$

It is possible to see that for D=0 $\rightarrow R_C = R$ (the same for G and B), indicating that the colour remains the same, and no adaptation occurs.

The next step is to apply a compression to this data. The relationship between the stimulus and the cones' responses is non-linear. If the intensity of the stimulus is too low, the cones do not respond in a significant way, however, if the stimulus is too high the responses reach a

maximum level. This is done in another space, called Hunt-Pointer-Estevez, that can better represent cone responses. From R_C, G_C, B_C we return to the tristimulus values using the inverse of the $\mathbf{M}_{\text{CAT02}}$ and the apply the Hunt-Pointer-Estevez matrix M_{HPE}:

$$\begin{bmatrix} R' \\ G' \\ B' \end{bmatrix} = M_{HPE} {M^{-1}}_{CAT02} \begin{bmatrix} R_C \\ G_C \\ B_C \end{bmatrix} \tag{4.16}$$

$$M_{HPE} = \begin{bmatrix} 0.38971 & 0.68898 & -0.07868 \\ -0.22981 & 1.18340 & 0.04641 \\ 0.00000 & 0.00000 & 1.00000 \end{bmatrix} \tag{4.17}$$

It is now possible to calculate the post-adaptation non-linearity using:

$${R'}_a = \frac{400\left({F_L R'}/{100}\right)^{0.42}}{27.13 + \left({F_L R'}/{100}\right)^{0.42}} + 0.1 \tag{4.18}$$

$${G'}_a = \frac{400\left({F_L G'}/{100}\right)^{0.42}}{27.13 + \left({F_L G'}/{100}\right)^{0.42}} + 0.1 \tag{4.19}$$

$${B'}_a = \frac{400\left({F_L B'}/{100}\right)^{0.42}}{27.13 + \left({F_L B'}/{100}\right)^{0.42}} + 0.1 \tag{4.20}$$

Where F_L is a adaptation factor to the luminance level depending on L_A.

$$k = \frac{1}{(5L_A + 1)} \tag{4.21}$$

$$F_L = 0.2k^4(5L_A) + 0.1(1 - k^4)^2(5L_A)^{1/3} \tag{4.22}$$

The response of the cones depends on the intensity of the stimulus and on the state of adaptation of the eye. F_L is used to model the adaptation to different illumination levels.

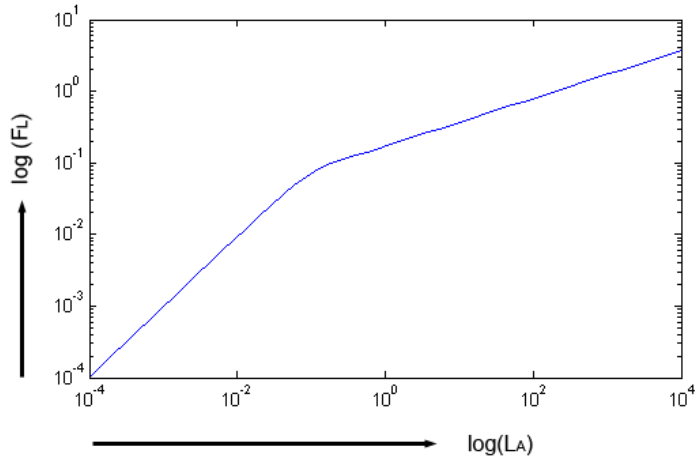


Fig. 25: Graph showing the relation between L_A and F_L

At scotopic level F_L and L_A are proportional (in figure 25, the first part of the graph, curve at 45°), this mean there is no compensation. At photopic level there is partial compensation (full compensation occurs only for a constant F_L).

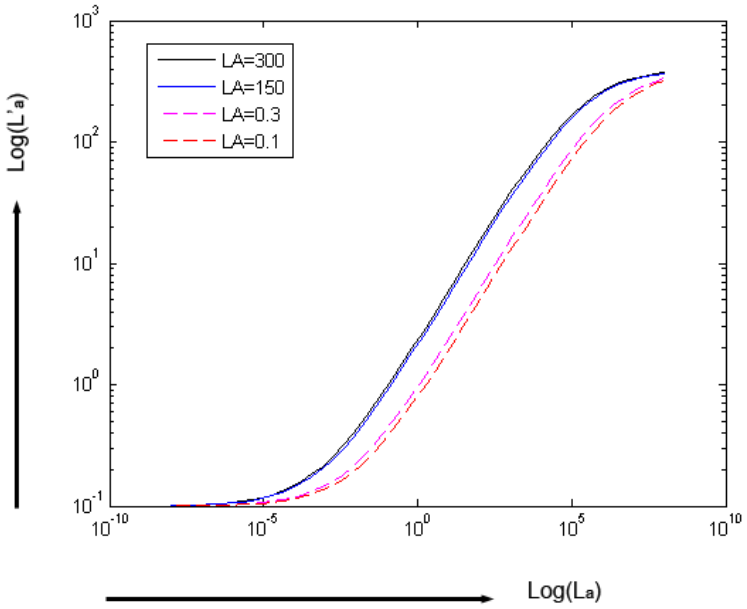


Fig. 26: Curves showing the tone compression

From the curve in figure 26 it is possible to see that at scotopic level, for small changes in L_A [5](from 0.1 to 0.3 cd/m^2) the (dashed) curves have significant change, meaning that there is no adaptation (in sense of compensation). At photopic level, however, even a great increase in L_A (from 150 to 300 cd/m^2) has a small incidence on the curves (solid)

4.5.3 Output

Finally is possible to calculate the appearance terms. Almost all the equations are derived empirically in order to fit various colour appearance data. These transformations are used (even if sometimes slightly modified) also in the iCAM model (described later in this chapter) which is more interesting for our purpose. Therefore the equations presented below are a list. Further information can be found in (Hunt, 1995), (Moroney, et al., 2002), (Fairchild, 2005).

Starting from the opponent-type responses:

$$a = R'_a - 12G'_a/11 + B'_a/11 \tag{4.23}$$

$$b = \left(1/9\right)(R'_a + G'_a - 2B'_a) \tag{4.24}$$

[5] L_A is referred to the adapting luminance, while L_a is referred to the L (Long) cone after adaptation (M_a, S_a have the similar form as the L_a shown in the graph).

And from some other constants depending on viewing condition: the induction factors N_{bb} and N_{cb} and the base exponential non-linearity z:

$$n = \frac{Y_b}{Y_w} \tag{4.25}$$

$$N_{bb} = N_{cb} = 0.725\left({}^{1}/_{n}\right)^{0.2} \tag{4.26}$$

$$z = 1.48 + \sqrt{n} \tag{4.27}$$

Hue is calculated as an angle, expressed in degrees (0°-360°):

$$h = \tan^{-1}\left(\frac{b}{a}\right) \tag{4.28}$$

The achromatic response is a weighted summation on the previously computed cone response, multiply with the brightness induction factor:

$$A = [2R'_a + G'_a + (1/20)B'_a - 0.305]N_{bb} \tag{4.29}$$

In the same way is calculated A_W for the white.

It is now possible to calculate the lightness J:

$$J = 100\left(\frac{A}{A_W}\right)^{cz} \tag{4.30}$$

Where c is chosen from table 1 according to the surround and z is calculated from equation (4.27)

Q is the brightness:

$$Q = (4/C)\left(\sqrt{(J/100)}\right)(A_W + 4){F_L}^{0.25} \tag{4.31}$$

In order to calculate the Chroma C it is necessary to compute an eccentricity factor e_t:

$$e_t = {}^1/_4 \left[\cos\left(h\left(\frac{\pi}{180}\right) + 2\right) + 3.8\right] \tag{4.32}$$

And a temporary quantity t:

$$t = \frac{(50000/13) N_c N_{cb} e_t \sqrt{(a^2 + b^2)}}{R'_a + G'_a + (21/20) B'_a} \tag{4.33}$$

$$C = t^{0.9} \sqrt{(J/100)} (1.64 - 0.29^n)^{0.73} \tag{4.34}$$

M describes the colourfulness, and is related to chroma. The scale factor ${F_L}^{0.25}$ indicates that M increases with the adapting luminance.

$$M = C {F_L}^{0.25} \tag{4.35}$$

Finally the saturation is calculated:

$$s = 100 \sqrt{(M/Q)} \tag{4.36}$$

4.6 iCAM

As just described a colour appearance model attempts to estimate the appearance of a single stimulus observed on a certain uniform background, usually specified simply by its luminance. This kind of model is not useful to study the appearance of an image. In order to perform this task spatial properties of the human visual system has to be taken in account. Fairchild and Johnson (Fairchild, et al., 2002), (Fairchild, et al., 2004) undertook studies on a new kind of model, known as image appearance model (iCAM). These kind of models work on images, in which every pixel is considered within a context. Furthermore, while CAMs predict attributes like lightness, brightness, chroma, colourfulness and hue, iCAMs can predict attributes related to the quality of the image for example sharpness, graininess, contrast and resolution. In 2006 a model called iCAM06 was developed as tone mapping

operator for HDR images (Kuang, et al., 2007). In the next section the framework of the iCAM06 is described. In chapter 7 some experiments are presented in order to compare iCAM06 as a tone mapping operator based on the appearance with a Retinex based algorithm.

4.6.1 iCAM06 framework

The actual input for iCAM06 is an HDR image[6], that should be specified in terms of CIE XYZ tristimulus values. Typically an HDR image is represented as a floating point RGB image linear to absolute luminance, therefore it is necessary to convert these values in XYZ values, using specific camera characterization. Usually camera characteristics are unknown. An example of transformation is from the sRGB space, through the following matrix:

$$M_{sRGB} = \begin{bmatrix} 0.4124 & 0.2127 & 0.0193 \\ 0.3576 & 0.7152 & 0.1192 \\ 0.1805 & 0.0722 & 0.9504 \end{bmatrix} \tag{4.37}$$

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = M_{sRGB} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \tag{4.38}$$

At this point the image is separated in two images, called a base layer and a detail layer. This is done through an edge-preserving filter, called bilateral filter (Durand, et al., 2002), that practically, working in the spatial domain, blurs the image, but retains sharp edges. The output is the *base layer*, containing only large-scale variations. The detail layer is generated subtracting the base layer from the original image. In this way the next steps can be applied only on the base layer, to preserve the details of the image.

The first step is the chromatic adaptation. This is similar to that of the CIECAM02, but with a different meaning. From the original image a blurred version of the image itself is calculated. We can term this as the *white layer* (X_W, Y_W, Z_W). The size of blur can be modified.

The CAT is performed in a cone response space (RGB). To convert from XYZ to RGB is used the M_{CAT02} matrix (equation 4.11)

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = M_{CAT02} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \tag{4.39}$$

[6] A range of experiments on the HDR imaging are presented in chapter 6, in order to verify if this kind of images are able to capture the luminance of a real scene.

Similarly R_W, G_W, B_W can be calculated starting from the *white layer* (X_W, Y_W, Z_W)
The CAT is performed in this way:

$$R_C = \left[\left(R_{D65}\frac{D}{R_W}\right) + (1 - D)\right]R \tag{4.40}$$

$$G_C = \left[\left(G_{D65}\frac{D}{G_W}\right) + (1 - D)\right]G \tag{4.41}$$

$$B_C = \left[\left(B_{D65}\frac{D}{B_W}\right) + (1 - D)\right]B \tag{4.42}$$

The main difference with the CIECAM02 (equations 4.13, 4.14, 4.15) model is that in the iCAM, to spatially modulate the adaptation white point, a low-pass version (blurred) of the image itself is used. According to the concept of locality, the white point (R_W, G_W, B_W) is therefore different at each pixel location. Furthermore, the CAT converts the global white point to D65 illuminant, which is useful for working in IPT space. IPT is an opponent colour space, where I indicates light-dark correlate, P the red-green and T the yellow-blue.

The degree of adaptation is calculated in the same way (equation 4.12):

$$D = F\left[1 - \left(\frac{1}{3.6}\right)e^{\left(\frac{-(L_A - 42)}{92}\right)}\right] \tag{4.43}$$

However in this case L_A is not simply the 20% of L_w (white object in the scene), but is a matrix of values calculated as the 20% of the Y_W (Y channel of the blurred image).
The second step is to apply a non linear compression, simulating the photoreceptors responses. In iCAM this is done combining cone and rods responses. Once again it is used a different "cone space". First go back in XYZ inverting M_{CAT02}, second apply the Hunt-Pointer-Estevez matrix (M_{HPE}, equation 4.17):

$$\begin{bmatrix} R' \\ G' \\ B' \end{bmatrix} = M_{HPE}{M_{CAT02}}^{-1}\begin{bmatrix} R_C \\ G_C \\ B_C \end{bmatrix} \tag{4.44}$$

And the cone responses are calculated:

$$R'_a = \frac{400\left(\frac{F_L R'}{Y_W}\right)^p}{27.13 + \left(\frac{F_L R'}{Y_W}\right)^p} + 0.1 \tag{4.45}$$

$$G'_a = \frac{400\left(\frac{F_L G'}{Y_W}\right)^p}{27.13 + \left(\frac{F_L G'}{Y_W}\right)^p} + 0.1 \tag{4.46}$$

$$B'_a = \frac{400\left(\frac{F_L B'}{Y_W}\right)^p}{27.13 + \left(\frac{F_L B'}{Y_W}\right)^p} + 0.1 \tag{4.47}$$

The F_L is calculated in this way:

$$F_L = 0.2k^4(5L_A) + 0.1(1 - k^4)^2(5L_A)^{1/3} \tag{4.48}$$

The form is equal to that of CIECAM02, but, as in equation 4.22, L_A has a different meaning. Similarly to the Hunt model (Hunt, 1982) iCAM06 includes functions to also simulate the rods' responses:

$$A_S = 3.05B_S\left[\frac{400(F_{LS}S/S_W)^p}{27.13 + (F_{LS}S/S_W)^p}\right] + 0.3 \tag{4.49}$$

Where the parameters are obtained from:

$$L_{AS} = 2.26L_A \tag{4.50}$$

$$j = \frac{0.00001}{[(5L_{AS}/2.26) + 0.00001]} \tag{4.51}$$

$$B_S = \frac{0.5}{\left\{1 + 0.3\left[\left(\frac{5L_{AS}}{2.26}\right)\left(\frac{S}{S_W}\right)\right]^{0.3}\right\}} + \frac{0.5}{\left\{1 + 5\left[\frac{5L_{AS}}{2.26}\right]\right\}} \tag{4.52}$$

$$F_{LS} = 3800j^2\left(\frac{5L_{AS}}{2.26}\right) + 0.2(1 - j^2)^4\left(\frac{5L_{AS}}{2.26}\right)^{1/6} \tag{4.53}$$

Summing the cones and the rods responses, the final tone compression response is obtained.

$$RGB_{TC} = RGB'_a + A_S \tag{4.54}$$

This image is then combined with the detail layer. The obtained image is converted in the IPT colour space. In order to perform this operation it is necessary before to transform the image in another LMS cone response, using the $M_H{}^{D65}$ matrix:

$$\begin{bmatrix} L \\ M \\ S \end{bmatrix} = M_H{}^{D65} \begin{bmatrix} X_C \\ Y_C \\ Z_C \end{bmatrix} \tag{4.55}$$

$$M_H{}^{D65} = \begin{bmatrix} 0.4002 & 0.7075 & -0.0807 \\ -0.2280 & 1.1500 & 0.0612 \\ 0.0000 & 0.0000 & 0.9184 \end{bmatrix} \tag{4.56}$$

These LMS are then compressed and the result is converted in the IPT space:

$$L' = L^{0.43} \tag{4.57}$$

$$M' = M^{0.43} \tag{4.58}$$

$$S' = S^{0.43} \tag{4.59}$$

$$\begin{bmatrix} I \\ P \\ T \end{bmatrix} = M_{IPT} \begin{bmatrix} L' \\ M' \\ S' \end{bmatrix} \tag{4.60}$$

$$M_{IPT} = \begin{bmatrix} 0.4000 & 0.4000 & 0.2000 \\ 4.4550 & -4.8510 & 0.3960 \\ 0.8056 & 0.3572 & -1.1628 \end{bmatrix} \tag{4.61}$$

For details see Ebner and Fairchild work (Ebner, et al., 1998).
In order to simulate the Stevens effect (increasing the luminance lead to an increasing of the perceived local contrast, (Stevens, et al., 1963)), a power-function adjustment is applied:

$$Details_a = Details^{(F_L+0.8)^{0.25}} \tag{4.61}$$

To simulate the Hunt effect (increasing the luminance lead to an increasing of the perceived colourfulness) P and T are enhanced:

$$P = P\left[(F_L+1)^{0.2}\left(\frac{1.29C^2-0.27C+0.42}{C^2-0.31C+0.42}\right)\right] \tag{4.62}$$

$$T = T\left[(F_L+1)^{0.2}\left(\frac{1.29C^2-0.27C+0.42}{C^2-0.31C+0.42}\right)\right] \tag{4.63}$$

Where C indicates the Chroma (calculated before the P and T improvement)

$$C = \sqrt{P^2+T^2} \tag{4.64}$$

4.6.2 Output

In order to display the image obtained, the IPT image is transformed back in XYZ image, and to the device dependent RGB image.
For completeness, as for the CIECAM02 model, the equations to calculate the appearance attributes are given below:

$$\text{Lightness: } J = I \tag{4.65}$$

$$\text{Chroma: } C = \sqrt{P^2+T^2} \tag{4.66}$$

$$\text{Hue angle: } h = \tan^{-1}\left(\frac{P}{T}\right) \tag{4.67}$$

$$\text{Brightness: } Q = \sqrt[4]{F_L}\,J \tag{4.68}$$

Colourfulness: $M = \sqrt[4]{F_L}\, C$ (4.69)

4.7 Conclusions

In this chapter the CIE colour appearance models are presented. The needs of these models originate from the fact that specific colour phenomena cannot be explained using only basic colorimetry. According to the CIE, a colour appearance model is a model able to predict at least the relative colour-appearance attributes of lightness, chroma, and hue. In order to predict a colour appearance phenomena, a definition of the viewing conditions and its partition is necessary (chapter 4.2). When changing the illumination condition, a chromatic adaptation transform (CAT) is applied (chapter 4.3). The CAT is the essence of all the CAMs (4.4). In chapter 4.5 it is presented the CIECAM02, which aim is to predict the colour appearance attributes in different situations. However the CIECAM02 allows to estimate the appearance of a single stimulus, observed on a certain uniform background. To study the appearance of a complex image an image appearance model (iCAM) is necessary. In chapter 4.6, the iCAM06 model is presented, which attempts to simulate the appearance of an high dynamic range image. Looking at these two models it is observed that they perform chromatic adaptation transform and tone compression in two different cones space. In fact the traditional von Kries transform uses the CAT02 matrix, while for the luminance adaptation it is used the HPE matrix. These matrix fit the psychophysical data better than using one matrix. If the human visual system is ever to be fully decoded it will maybe possible to discover an *ideal* colour space working for both the adaptation. However passing from XYZ space to this "ideal" colour space could be a non linear operation, and a simple matrix is unlikely to work.
The models presented in the chapter have a remarkable complexity in respect to the algorithms derived from the Retinex theory. In fact, the CIE models attempt to mimic every single step of the human visual system, splitting the vision process in several stages, in contrast to the Retinex model that treat the entire process of vision as a whole.

5. Raw and HDR imaging

The word *Photography* derived from Greek and it is the union of two words: *phos* (light) and *graphis* (writing), literally "writing with light". In other words, the photographic process involves the capture and fixing of an image, by a projection, on a device (analog or digital) that is sensitive to the light.

When taking a picture, the light comes into the objective due to the convergence from one or more lens through the diaphragm (a variable hole), in order to be projected finally on a portion of light-sensitive material, i.e. the film, or the sensor concerning digital devices. Thanks to a particular mechanism, the shutter, it is possible to decide, how long the diaphragm can be held opened (exposure time). The union of these two values, opening/exposure time, defines the exposure value of the shot, or rather the quantity of light that reaches the sensor. The sensor is a two dimensional array that collects the photons and converts them in electronic signals . Each element of the array contributes for one pixel to the final image. The just reordered data are called *raw*, because no processing has been done on them. In theory, the data collected by the raw are linearly proportional to the light of the real scene. At this point the camera performs a raw conversion, to transform these data in a typical 8-bit depth image, usually a jpeg.

The Camera Response Curve is a function that correlates to the quantity of light incoming from a point of the scene and the digit value of that pixel that is captured in the camera. In analog cameras, this function is described as film density. In the digital devices the sensor simulates the behaviour of the film. This function is quite linear for mid values and logarithmic for the extreme values (figure 27). The reason of this behaviour is straightforward: for low exposure values there are not enough photons to excite the sensor. At high exposure values there are too many photons that saturate the sensor. Each sensor has its own response curve.

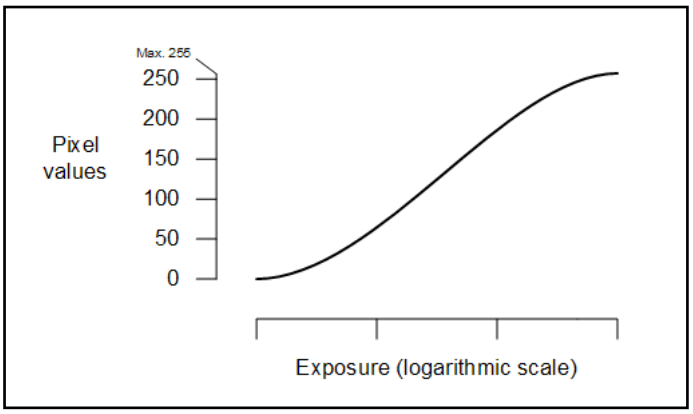


Fig. 27: Shape of a typical camera response curve for a digital camera.

It is impossible to capture the wide range of light that is present in the natural world with a traditional picture. Moreover *media* used to display a picture (monitor, CRT or LCD, or a sheet of paper) are unable to reproduce the light of a real scene. In order to capture the wide range of light in the natural world, HDR (High Dynamic Range) algorithms are used. These algorithms merge different shots of the same scene, this is accomplished by capturing the scene using different exposure time, in a single HDR image, which contains information about luminance values of the scene. The HDR image just obtained has a number of bit greater than 8 per channel, therefore, to display the image on a traditional monitor, is necessary to re-map the values of the HDR image in a typical 8 bit per channel image. With 8 bit per channel it is not possible to record all the range of luminance present in the natural world, therefore these images are called *low dynamic range* (LDR) image. The operation of mapping a HDR image in a LDR image is called *Tone Mapping*. Varying the parameters of the tone mapping it is possible to obtain different goals:

- Produce an image that is aesthetically pleasing
- Lighten the underexposed zones of a picture
- Simulate some perceptive mechanisms, for example mimicking the behaviour of the human visual system.

In chapter 5.1 it is given a background on the raw file, how the data are memorized and then converted in a 8 bit image. In chapter 5.2 the HDR imaging is described. The tone mapping is used to transform the HDR image in a 8 bit image (section 5.2.1). In section 5.2.2 an overview on the existing approaches to create HDR image is given.

5.1 A background about the Raw format

When the light enters in the camera objective, it reaches the sensor composed of photo-sensible material. The sensor produces a charge that is directly proportional to the amount of light that strikes it. This process gives rise to the *raw* file, that therefore it is a greyscale image. To produce a colour image from the raw file a colour filter array (CFA) is used. Each element of the array is covered by a colour filter, so that each element captures only red, green, or blue light. The typical filter used by the camera has in Bayer pattern (figure 28).

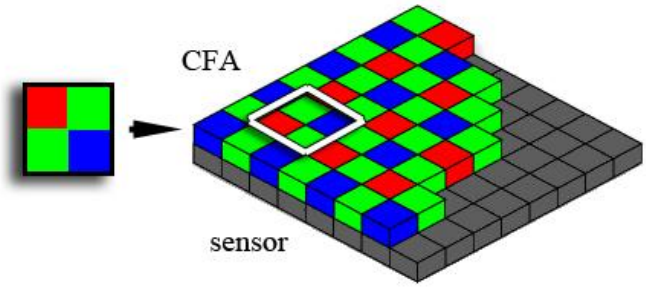


Fig. 28: Bayer pattern for the CFA, that cover the sensor.

As the human eye is most sensible to green light, in a portion of four filters, two of them are green. Each element of the sensor captures just one colour. The red-filtered elements produce a greyscale value proportional to the amount of red light that reaches the sensor, and the same for the green and blue filtered elements. The raw converter transforms the greyscale raw capture into a colour image by interpolating the colour information of every pixel from its neighbours. This process is called demosaicing. Many other operations are involved in the raw conversion, for example:

- White balance. The raw file is not affected by the white balance setting of the camera. These data are recorded as metadata by the camera, and applied when generating the 8 bit image (jpeg).
- Gamma correction: the sensor capture the data linearly to the light of the scene and therefore a gamma correction is applied to the 8-bit image.
- Compression: usually a raw file is 10-14 bit depth, and it is compressed to 8-bit depth when converted in jpeg file.
- Noise reduction and anti-aliasing: to avoid artefacts most raw converters perform some edge detection and anti-aliasing to reconstruct noisy data.

Therefore a raw file contains the unprocessed image data from a digital camera sensor. Unfortunately, the raw files are not standardized, and each camera manufacturer has its own file format (table 3) that is usually not documented[7]. When a camera that supports raw is sold, a software to convert the specific raw file is usually provided. However, many software able to read and convert specific raw files, are available both from vendors that bought the license for a particular technology and from programmers that reverse-engineered specific raw formats.

Canon	CRW/CR2
Nikon	NEF
Minolta	MRW
Olympus	ORF
Sony	SRF
Kodak	DCR
Sigma	X3F

Tab. 3: Table with the most common raw file for different manufacturers.

5.2 High Dynamic Range Imaging (HDRI)

In the last few years, picture resolution has increased as result of camera sensors becoming more sophisticated (passing from 4-6 Megapixels cameras to 18 Megapixel professional cameras). The information is also more precise and less noisy. The attention of the scientific community, and of the market, has recently focused on another aspect of photography, which is the issue of luminance. The dynamic range of a scene is the ratio between the lightest point of a scene and the darkest point of the same scene. Over the last few years the attention has also focused on the low dynamic range (LDR) images (typical 8-bit per channel image). This is due to the limits of displaying a high dynamic range on a traditional monitor, and by the format chosen to store the images themselves (JPEG). The most used format to store an image uses one byte for channel (8 bit Red, 8 bit Green, 8 bit Blue for a total of 24 bit RGB). This method results in a display of 16 million colours. However this large number is not enough to represent all the possible luminances of a real scene. The light coming from the sun at noon has an intensity of 10^7 cd/m^2 while the light reflected from a white paper under the moon light has an intensity between 10-1 and 10-5 cd/m2 (Oleari, 2008 p. 73).

[7] In 2004 Adobe launched the Digital Negative (DNG), an open raw image format based on the TIFF format, in order to create a standard (Adobe, 2004).

A classic example is the capture of a scene in a low lit room of somebody in front of a window. The dynamic range is too big to be reordered by the camera, and so usually only the subject inside (figure 29, a) or the environment outside (figure 29, b) can be correctly exposed.

(a) (b)

Fig. 29: A typical situation when trying to take a picture of a person in front of a window.

Moreover common images contain 3 or 4 order of magnitude of luminance, but traditional monitors can process only 2 orders of magnitude. This fact depends from the limit of excitability of the phosphors (in CRT monitors) or from the limit of the light intensity of the lamp (in LCD monitors). To create a HDR image some pictures of the same scene are captured using different exposure time (figure 30). An algorithm[8] will merge these images in a HDR image, which selects the correct value for each pixel using the information about the exposure value of each shoot. First of all the algorithm tries to approximate the camera response curve of the device in use. Then, this curve will be used to obtain a *radiance map* of the scene. This map contains information about the radiance values of the scene captured and, in the optimal case, it has to be proportional to the real values of the scene. However, can happen that these values are not proportional to real values because of the glare effect. Glare is an uncontrolled spread of light in the camera lens. Some experiments (McCann, et al., 2007 p. 2-7) have demonstrated that glare reduces the dynamic range that can be correctly measured by the camera.

[8] Here we are speaking in general case. This workflow is common to many software able to create HDR images.

Fig. 30: Multiple exposure to create an HDR image.

Often, this inaccuracy can be unheeded: if the aim of the image is to aesthetically pleasing or to lighten the underexposed zones of a picture, the precision is not important. However if the task is to simulate perceptive mechanisms accuracy is vital.
Many formats exist to memorize HDR image, for example (Reinhard et al, 2005, p. 89-98):

- .HDR: uses 32 bit in RGBE format, where 8 bit are used for a common exponent (E).
- .TIFF: IEEE RGB, uses 32 bit/channel.
- .EXR: uses 16 bit/channel.

It is not possible to display directly the radiance map on a typical monitor. A tone mapping operator is needed. This operation consist in re-mapping the values of the radiance map in a range that can be displayed by the LDR monitor in use preserving the main details.

5.2.1 The Tone Mapping

Tone mapping comprises re-mapping the pixels' values of the scene in a range that can be correctly displayed by the display device using the information contained in the radiance map. Working on the parameters of the tone mapping operator (TMO) different goals can be achieved. Some example: make an image aesthetically pleasing, lighten the underexposed zones of a picture or simulate some perceptive mechanism (figure 31). There are two main types of TMO:

- Global: they are the most simple and the fastest ones. They re-map each pixel according to its intensity and using some global information of the image itself. This family of tone mappers don't use spatial information of the pixel itself.
- Local: they map each pixel according to its intensity and to the intensity of adjacent pixels. This mechanism makes this family of tone mapper more slowly. They are often used when it's important to emphasize local contrast. These tone mappers are called *Tone Render*.

In the next chapters we will describes iCAM06 and the Retinex models. Aim of these models is to simulate some behaviours of the human visual system. They can work also on HDR images, remapping the pixels in a low dynamic range. In particular, both are implemented in order to keep in account the spatial characteristic of an image. For this reason they can be considered as local tone mapping operator.

Fig. 31: Tone mapping with Photomatix, iCAM06, RSR.

In figures 32 and 33 it is possible to see an another example of multiple exposure and HDR tone mapping with Photomatix, iCAM06, RSR.

Fig. 32: Another example of multi-exposure images.

Fig. 33: Again, Tone mapping with Photomatix, iCAM06, RSR.

5.2.2 The Existing Approaches

There are a range of techniques, developed over the past years, that attempt to capture the luminance of a natural scene. One technique consists of shots taken at different exposure times to create a radiance map of the scene. In 1995 Mann and Picard proposed to encode the pixel values of the acquired images as a floating point instead of using the traditional RGB encoding (Mann, et al., 1995). The goal of their work was to lighten the underexposed zones of the picture. In 1997 Debevec and Malik suggested a more accurate implementation of the algorithm that merged the shots (Debevec, et al.). They tried to obtain the camera response curve from the images and then recovered the radiance map of the scene. In their work, Mitsunaga and Nayar used a different approach (Mitsunaga, et al., 1999). They discover that many real devices shared a similar response curve. They decided to implement a radiometric model that, by changing some parameters, could represent many different devices. The major advantage of this method is that it did not need exposure values as an input as they are calculated by the model. In 2002 Adobe implement a procedure to create HDR images in Photoshop CS2 and in 2004 Grossberg and Nayar gathered a large number of response functions in order to create a model to obtain the camera response curve of any device changing a few number of parameters (Grossberg, et al., 2004). Another approach consisted in capturing the scene with multiple cameras at the same time. Each camera was set up to take the picture at different exposure times. The image of the scene was sent to all camera using a beam splitter. The advantage of this technique is the possibility to capture the pictures at the same time and removing all imprecision. The major disadvantage was the high cost of the equipment (Saito, 1996). In 2000 Nayar and Mitsunaga placed a greyscale grid in front of the sensor. In this way it was possible to take different measurements for each pixel. Data coming from an area covered with a lighter grey corresponded to an higher exposure time, data

coming from an area covered with a darker grey corresponded to an lower exposure time (Nayar, et al., 2000).

6. Experiments on Raw and HDR

6.1 Introduction

The goal of these experiments is to understand how the light in a real scene is captured by digital cameras. During the acquisition process, the signal is transformed more times. Moreover, these transformations have a non-linear process. There are many factors that can occur during the acquisition of the images, which can change the real values of luminance. For example, some in-built features of the camera itself (lens, sensors) or some post-processes (noise reduction, chromatic aberration) can change the digital values even before the image is stored. Moreover we wanted to understand if we can retrieve part of this "lost" luminance by applying an algorithm to create an high dynamic range (HDR) image of the scene. These algorithms require as input different shots of the same scene that are captured at different exposure times. In order to test if there are any changes in the values stored, a range of digital cameras are used to acquire the scene.

6.2 A test on the features of the cameras

During the acquisition of a real scene with a digital camera, many different transformations of the signals occur. These transformations depend on some in-built features of the camera (lens, sensor) and follow a non linear process. Moreover, in the camera's firmware, many post-processing elaborations take place, such as noise reduction or chromatic aberration. These elaborations will affect how the image is taken even before it is stored. In the JPEG this processing includes compression, that changes the values of the pixels. For this reason, professional cameras, have the opportunity to save the captured image as a RAW format; a file that saves the image without any kind of elaboration. The goal of this experiment is to understand how much the in-built features and the post-processing influences the image captured.

6.2.1 Experimental Setup

In this section the experimental setup is described, including the scene construction, the tools used and the procedures adopted. The scene has to be as close as possible to a real scene. For this reason the main characteristic of the scene are:

- A strong difference of luminance between the lightest point and the darkest point (this difference contributes to the creation of a high dynamic range in the scene);
- A strongly-oriented light source, that projects edges and gradients.

6.2.2 A description of the stage

In order to contain the objects in the scene, a small plywood cube is constructed. The cube is open at one side, and has a dimension of 40x40x40cm. The border of the cube is painted black, and the faces are covered with black cardboard to avoid unwanted reflections. Halfway into the cube is a dividing wall into which is cut a square hole, resulting in a space at the front of the box and a space in the back. The front portion is open on one side, enabling more light. The back portion one has only a small aperture, so it is darker; in this way we can achieve an higher dynamic range.

6.2.3 A description of the objects

The objects constructed for the scene are in plywood and painted with acrylic colours. For this experiment, we chose to work only with achromatics, as our interest was to focus only on the luminance. The faces of the objects were painted with 4 different colours: black, white and two different greys, labelled as: B, G1, G2, W (B is black, W is white and G1 is the light grey and G2 is the dark grey).

Ten samples of every face are measured with a colorimeter in order to check the chromatic uniformity. An object with the dimension of 25x10x10cm was placed in the bottom of the cube, behind the black wall, partially visible from the hole in the wall itself. Other two objects were placed in the frontal part of the cube and the two smaller were placed in front of the other objects (figure 34). This arrangement let all the objects of the scene clearly visible and, at the same time, the objects project edges and gradient on the other objects and on the cube. The light source was positioned at the same height as the objects, on one side of the scene. The spotlight was oriented to highlight the white face of the cube on the right and to let

in the dark the black face of the object behind the wall. This choice was made to maximize the difference between the lightest point and the darkest point. The tripod, on which all the cameras were mounted, was placed in front of the scene.

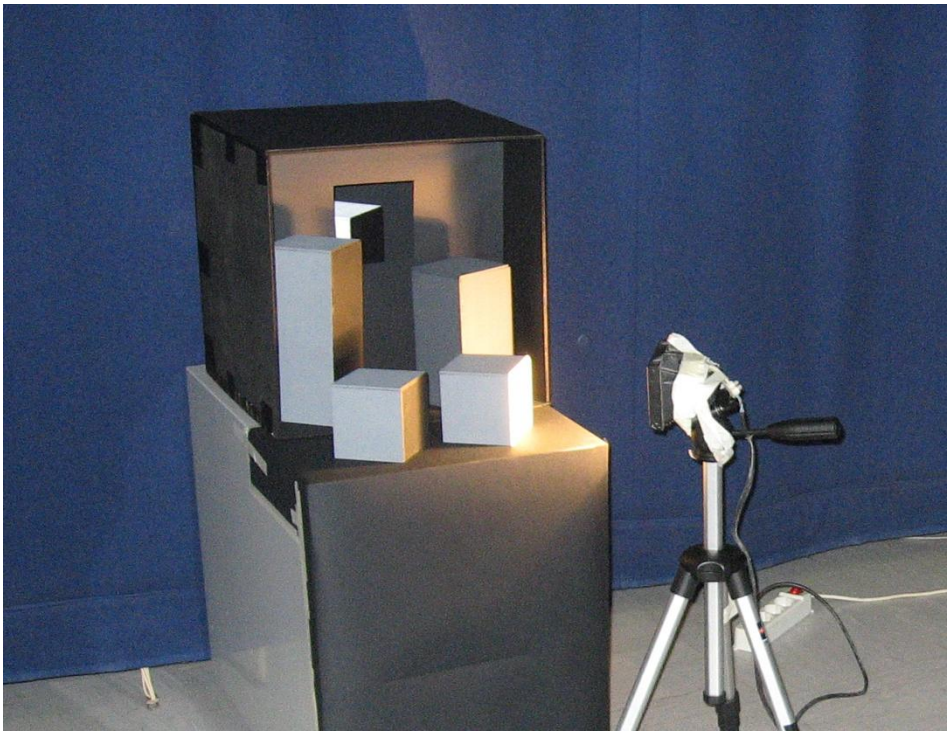


Fig. 34: The setup of the experiment.

6.2.4 The Tools

This section describes the tools used for the experiments. We can divide the tools in 4 categories:

- Measurement tools: colorimeter and spotmeter;
- Acquisition tools: the digital cameras used to capture the scene;
- Illumination tools: lamp used to enlighten the scene;
- Software tools: used to read the luminance values from the images.

An One-Eye Colorimeter is used during the construction of the scene to guarantee the colour uniformity of the object painted with the same colour. The Minolta spotmeter F is used to measure the luminance values of some samples in the scene, during the scene construction to assure that the scene has a significant dynamic range value.

A range of digital cameras are used to capture images of the scene. To cover all the typology of digital cameras we choose devices that belong to different classes: three “Reflex” digital

cameras (Canon EOS30D, Minolta Dynax 5D e Sigma SD9), one “Bridge” digital camera (Panasonic Lumix FZ38) and one “Compact” digital camera (Canon PowerShot A540). All the cameras enable complete manual setting.
The scene is highlighted using a spotlight of 150W, which is mounted on a tripod.

6.2.5 Procedures

This section describes the procedures used during the experiment. The tests were made on the same day, within one hour, in the same conditions. After the setup of the scene was complete the real dynamic range was measured with the Minolta Spotmeter. Several pictures of the scene were taken with all the cameras, applying the same settings: the aperture was fixed at F/5.0 and the exposure time was varied from 1/500 to 30 sec (eleven shots for each camera). The sensibility (ISO) was forced to 100 and the white was balanced manually. To avoid blur and micro-blur we used the self-timing shooting mode.

6.2.6 Results and Conclusions

In table 4 and table 5, the results are summarized. In particular in table 4 is shown the RGB values of the more overexposed images. The highest value is always clamped to 255. In the same way in table 5 the minimum value is always 0. These values are useless in order to get purposeful information. However the minimum values for the overexposed images and the maximum values for the underexposed images are not clamped and therefore are interesting for us. It is possible to see that these values are different for each camera. Looking at the first two rows of table 4, it is possible to observe that these values can vary quite a lot from each other. As the setting and the shooting conditions are the same, the only factors that may have changed the values in the pictures are the camera construction features and the post-processing.

EXP : +5	Rmin	Gmin	Bmin		Rmax	Gmax	Bmax
Panasonic Lumix FZ-38	13	17	26		255	255	255
Canon PowerShot A540	51	47	62		255	255	255
Minolta Dynax 5D	28	28	26		255	255	255
Canon EOS 30D	14	18	17		255	255	255
Sigma SD9	32	40	51		255	255	255

Tab. 4: Data regarding the more overexposed images

.

EXP : -5	Rmin	Gmin	Bmin		Rmax	Gmax	Bmax
Panasonic Lumix FZ-38	0	0	0		179	151	178
Canon PowerShot A540	0	0	0		177	179	176
Minolta Dynax 5D	0	0	0		178	177	159
Canon EOS 30D	0	0	0		219	205	208
Sigma SD9	3	3	3		181	166	147

Tab. 5: Data regarding the more underexposed images.

6.3 Experiments on raw and jpeg formats

6.3.1 The camera response curve

The camera response curve is a function that correlates the quantity of light incoming from a point of the scene and the digit value of that pixel stored in the camera. Two possible response curves are achievable: one related the raw file, one related to the 8-bit (usually jpeg) file. The following experiment aims to reconstruct the response curve for three different cameras: A Canon EOS 7D, a Canon EOS 400D and a Konica Minolta Dimage A200.

In order to achieve this purpose it is necessary a scene with a dynamic range of luminance of at least 10 Exposure Values (EV) [9]. The EVs can be measured with a Minolta F spotmeter. Thereafter, the digital output of the camera are related with the measured EVs to reconstruct the response curve.

6.3.2 Canon EOS 7D

To reconstruct the response curve of this camera, the image shown in figure 35 has been printed on paper and front lit using fluorescent[10] lights.

[9] The term EV can be referred to different concepts. In the field of photography a picture taken with EV 0 means that is correctly exposed. If EV is +1 the picture is overexposed of 1 stop, if it is -1 the picture is underexposed of 1 stop. The concept of correctly exposed means that to a particular area (in photography it can be usually the average of the whole scene or the central spot) the value of 128 will be assigned, in a range between 0 and 255. The same term can be used as a measurement of luminance. Knowing the ISO and the calibration constant k it is possible to derive the luminance in cd/m^2.

[10] The two lights used are: OSRAM L36W/765, Philips TLD 18W/33.

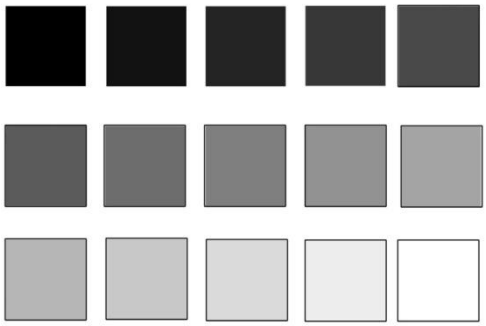


Fig. 35: image used to reconstruct the response curve of the Canon EOS 7D.

However, in this way, it was not possible to reach the 10 EV necessary, as increasing the light makes the black patch reflecting too much light, and decreasing the light makes the white patch not bright enough. Therefore, taking just on picture with the camera, it was not possible to a have an image with values ranging from 0 to 255. To cover all the dynamic range of the camera more pictures of the same scene have been taken. The aperture of the diaphragm and the exposure time were fixed, while the illumination varied. This is equal to take a single image with a high dynamic range (figure 36).

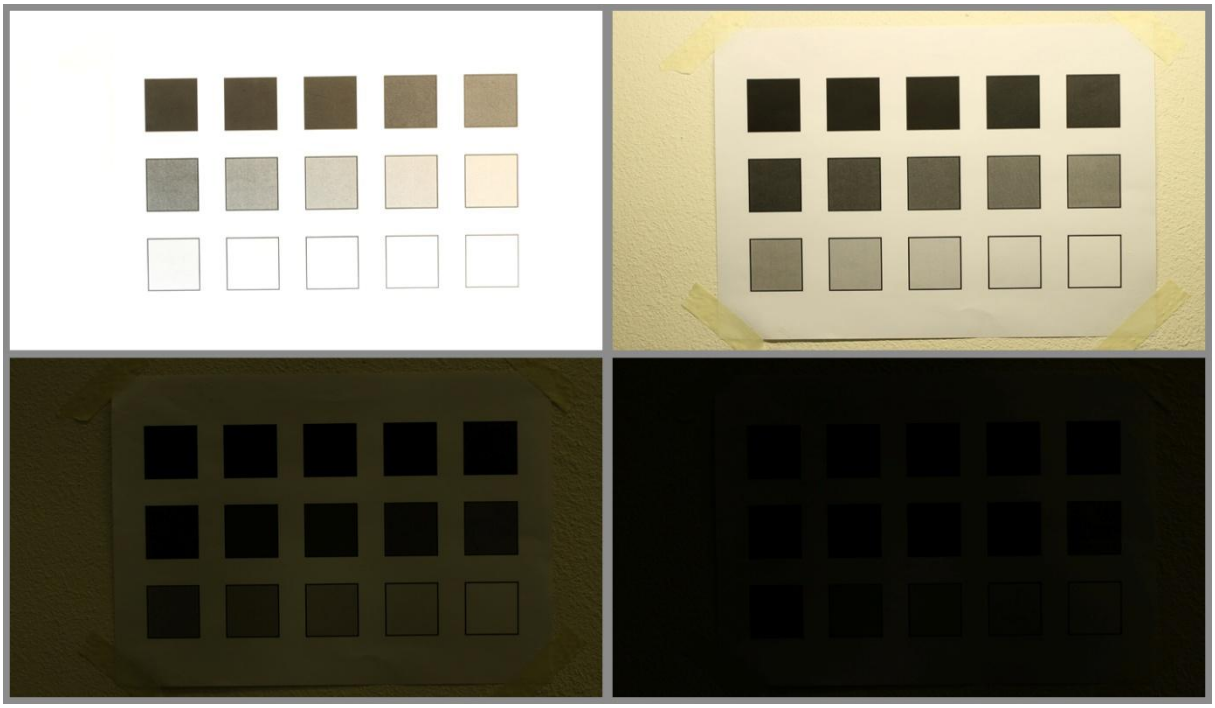


Fig. 36: The four images taken, varying the intensity of the light.

Two fluorescent lights have been used, able to be moved in order to change the light reaching the patches (figure 37).

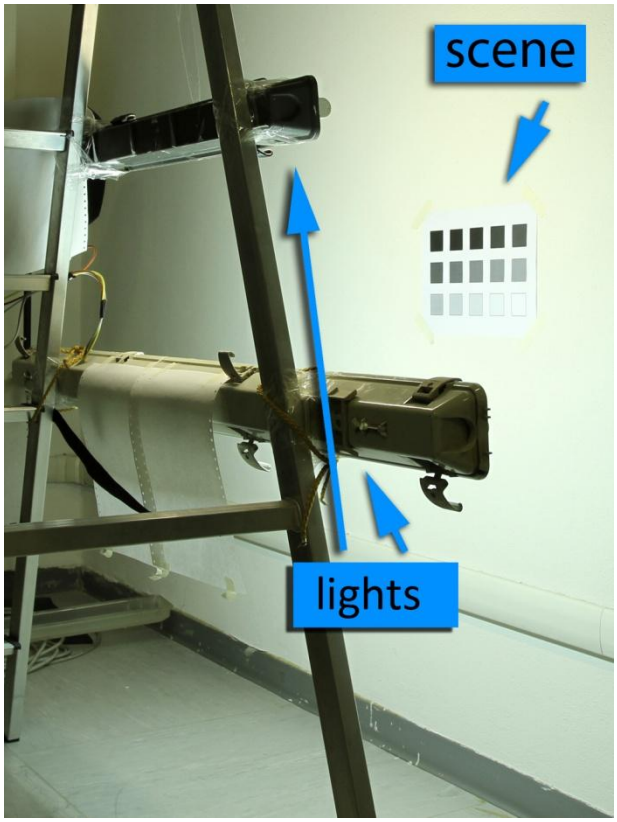


Fig. 37: The fluorescent lights can be moved, to vary the intensity of the light reaching the patches on the wall.

Four pictures have been taken, with the following camera setting: aperture: F/5.6, exposure time: 1/15 sec, ISO: 100. The output of the camera are four jpeg images and four raw images (Canon CR2 file format). The experiment operated as follows:

1 - The lights are placed near the patches, in order to achieve a high EV value for the white patch. The angle between the direction of the light and the normal to the printed patches was greater than 45°, in order to avoid reflection reaching the camera objective.

2 - The Exposure Values have been measured with the Minolta F spotmeter, from the point of view of the camera.

3 - A picture with the camera has been taken, with the camera setting fixed.

4 - The lights are moved back, in order to decrease the intensity of light on the patches, then the actions started again from step 2.

These steps are repeated until four images have been captured. It is now possible to extract the information from the files generated by the camera, both raw and jpeg. The native Canon CR2 raw files are read with Rawnalyze 2.10.4, able to displays the minimum, maximum and

mean values of a selected area. To extract the data from the jpeg file ImageJ 1.43 by Wayne Rasband has been used. This software is able to displays the minimum, maximum and mean values of a selected area too. In table 6 interesting data concerning the four images are indicated (we referred to them as first for the darker, and fourth for the lighter). The first column contains the EVs measured with the spotmeter F, the second column contains the conversion from EV to cd/m^2, the third macro-column contains the data extracted from the raw image for every patch, for the red, green and blue channels (average on an area), the fourth macro-column contains the data extracted from the jpeg file (average on an area).

	Physical Values		RAW			JPEG		
Image	EV	cd/m2	red	green	blue	red	green	blue
First	err	err	2050	2051	2048	0,14	0,16	0,15
	1,3	0.3444	2054	2057	2051	0,64	1,012	0,51
	4,4	2.9040	2098	2137	2085	12.884	12.884	5.880
Second	2	0.56	2055	2060	2053	1.194	1.351	0,415972
	6,4	11.880	2211	2344	2177	44.506	40.728	19.992
Third	6,6	13.68	2198	2327	2192	35.875	32.039	20.262
	10	140	3763	5293	3828	210.891	202.924	172.147
Fourth	8	36	2407	2719	2423	77.436	68.747	50.510
	11,9	542.30	7680	13277	9692	254.446	254.427	253.941
	12,3	701.10	9758	13582	12021	255	255	255
	12,8	991.80	13129	13582	13583	255	255	255
	13,1	1177	13582	13584	13584	255	255	255

Tab. 6: Data regarding the Canon EOS 7D

From the data shown in table 6 it is possible to obtain the following considerations:

- The first image has a range between 1 or less (with the tool used it is not possible to measure EV smaller than 1) and 4.4 EV.
- The second image has a range between 2 and 6.4 EV
- The third image has a range from 6.6 to 10 EV
- The fourth image has a range between 8 and 13.1 EV
- For the jpeg it is possible to notice that the saturation points occur between 11.9 and 12.3 EV.
- For the raw it is not possible to figure out the saturation point from this experiment, however it is clear that the raw file saturated for a greater luminance value.

- For the darkest image the jpeg values are near the zero, while for the raw they are between 2048 and 2051.

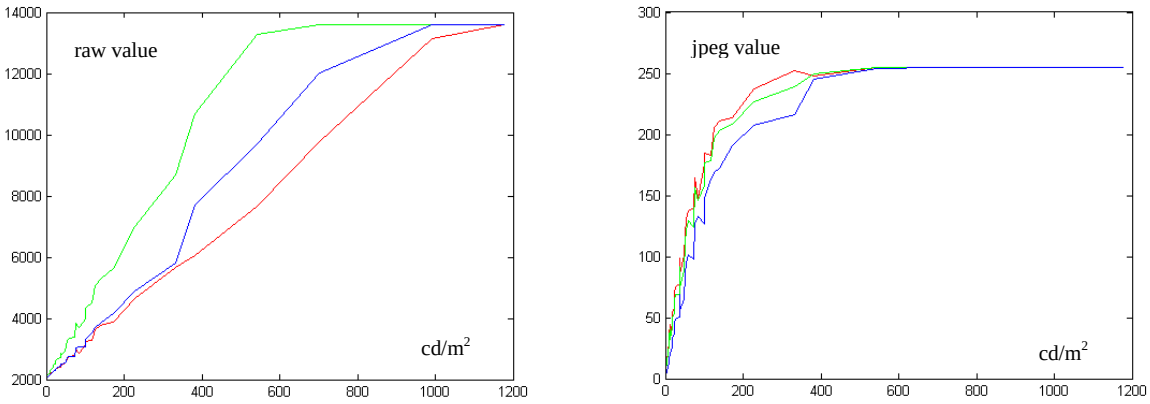
The last consideration can be misleading as one could think that decreasing the light the sensor is still able to record a lot of data. Therefore another test has been performed: a picture with the lens cover has been taken, to check the minimum possible value that the raw can assume. Using Rawnalize and measuring the whole image, the following values are extracted:

- Red channel: minimum: 2018, maximum 2087, mean 2048.
- Green channel: minimum: 2010, maximum 2083, mean 2048.
- Blue channel: minimum: 2019, maximum 2079, mean 2047.

Returning to the data in table 6, it is reasonable to suppose that the first row represents the minimum values recordable by the sensor, before falling into the noise.
To test the saturation level another test was performed. A picture is taken using a long exposure time. The jpeg image is completely white, the raw image has values:

- Red channel: minimum: 13583, maximum 13583, mean 13583.
- Green channel: minimum: 13582, maximum 13584, mean 13584.
- Blue channel: minimum: 13583, maximum 13584, mean 13584.

Therefore it is reasonable to assume that the last row in table 6 represents the maximum values recordable by the sensor, before it becomes saturated. The 13582 value for the red channel can be the last useful value before saturation. Interesting enough the raw channels saturate at slightly different points. In particular the green channel saturates before the other, around 12.3 EV, similarly to the jpeg channels. From these data it is possible to plot the response curves, both for the raw (figure 38, a) and for the jpeg (figure 38, b).
On the x-axis there are cd/m^2, on the y-axis there are raw values and jpeg values. The typical S-shape of the curve for the jpeg, as the image shown in figure 27 (p. 61) it is obtained using as input the EV values in place of the cd/m2 (figure 39). The curve is subject to noise, due to the small number of measurements with the spotmeter. To avoid noise more points have to be taken and then averaged.



(a) (b)

Fig. 38: Response curves for the three channels, for the raw (a) and the jpeg (b) files.

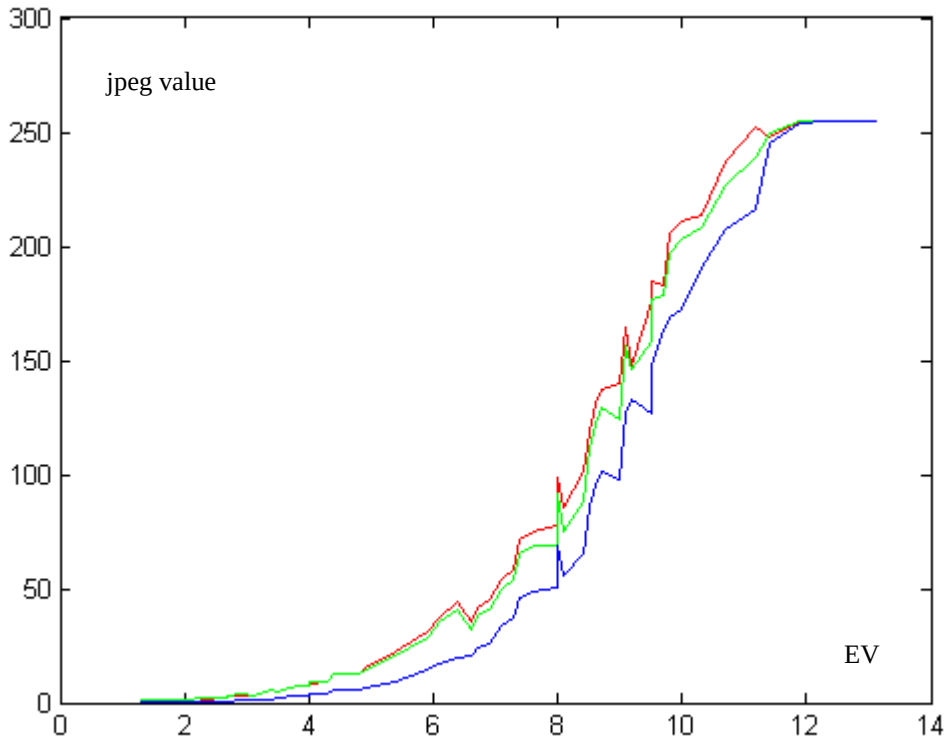


Fig. 39: Typical S-shape camera response curve.

From the images it is clear that, until the saturation points, the raw response for the Canon 7D is linear to real luminance values of the scene, while the jpeg has the characteristic gamma correction shape.

6.3.3 Canon EOS 400D

The same experiment performed in section 6.3.2 has been implemented with other two cameras: a Canon EOS 400D and a Konica Minolta Dimage A200. The patches used for the experiment have been re-printed, on matte photographic paper, in order to have a more uniform coloured area. Therefore, they have been printed bigger in size, to make them easily measurable. The fluorescent lights[11] are used again to illuminate the scenes (figure 40).

Fig. 40: the patches are visible on the wall. The fluorescent lights are placed vertically all around. The two cameras are used in the same time to take the pictures.

The setting for the two cameras are: aperture: F/5.6, exposure time 1/8 sec., ISO: 100. The white balancing was set on fluorescent light. As in the previous section, a portion of the data are shown below (table 7):

	Physical Values			RAW			JPEG		
Image	EV	cd/m2		red	green	blue	red	green	blue
First	err	err		256	256	256	0.46	1.15	0.61
	1.0	0.28		256	257	256	0.43	1.08	0.54
	1.2	0.31		257	258	256	0.48	1.33	0.68
	4.5	3.10		269	279	266	8.53	12.62	11.85
Second	2.3	0.68		260	263	259	1.84	4.41	3.65
	7.5	25.38		363	440	342	55.93	81.41	96.41
Third	5.6	6.84		308	346	302	27.83	40.35	54.69
	10.3	172.20		1049	1662	922	193.51	235.07	253.20

[11] In this case the lights used are: n°2 Philips TLD 58W/33, n°2 Mazdafluor Blanc Industrie 36W/33, n°1 Philips TLD 18W/33.

Fourth	7.0	18.00		437	540	396	94.30	110.66	138.17
	11.6	440.80		2585	3453	1559	254.93	254.94	254.94
	12.0	570.00		3001	3726	1826	255.00	255.00	255.00
	12.7	923.40		3726	3726	3378	255.00	255.00	255.00
	12.9	1065.90		3726	3726	3725	255.00	255.00	255.00
	13.0	1100		3726	3726	3726	255.00	255.00	255.00

Tab. 7: data concerning the EOS 400D.

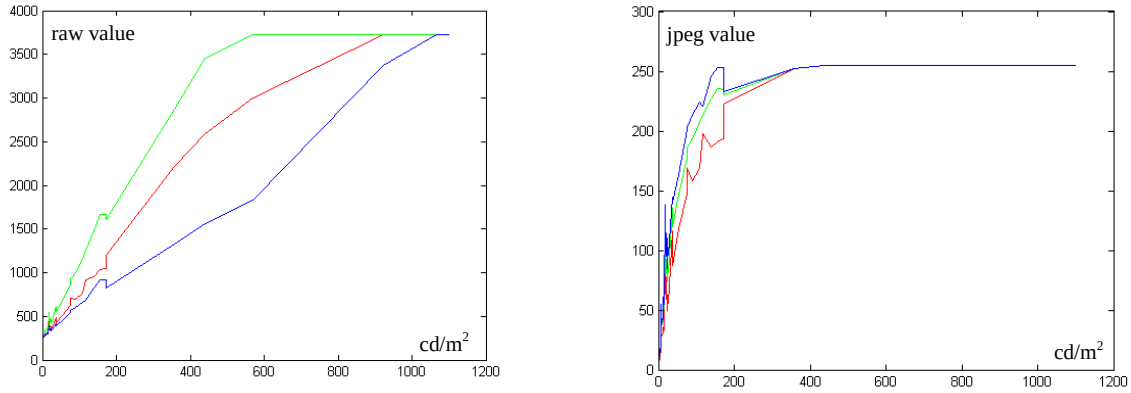
According to the data shown in table 7 it is possible to obtain the following considerations:

- The first image has a range between less than 1 and 4.5.
- The second image has a range between 2.3 and 7.5 EV
- The third image has a range from 5.6 to 10.3 EV
- The fourth image has a range between 7 and 13.0 EV
- For the jpeg it is possible to notice that the saturation points occur at 11.6 EV.
- The saturation points for the raw file are different for the three channels: between 12 and 12.7 EV for the red, between 11.6 and 12 EV for the green, between 12.7 and 12.9 for the blue .
- The raw file saturated for a greater luminance value.
- For the darkest image the jpeg values are near the zero, while for the raw they are 256.

Another test has been performed: a picture with the lens cover has been taken, to check the minimum possible value that the raw can assume. Using Rawnalize and measuring the whole image, the following values are extracted:

- Red channel: minimum: 249, maximum 264, mean 256.
- Green channel: minimum: 248, maximum 265, mean256.
- Blue channel: minimum: 249, maximum 265, mean 256.

Returning to the data in table 7, it is reasonable to suppose that the first and the seconds rows represents the minimum values recordable by the sensor, before falling into the noise.
The response curve for the raw (a) and the jpg (b) are shown in figure 41.



(a) (b)

Fig. 41: Response curves for the Canon EOS 400D, (a): Raw, (b): Jpeg.

As for the Canon EOS 7D until the saturation point the raw response for the Canon EOS 400D is linear to real luminance value of the scene, while the jpeg has the characteristic gamma correction. And, as the 7D, the raw channels saturate at slightly different points, with the green channel saturating before the others.

6.3.4 Konica Minolta Dimage A200

In this section data regarding the Konica Minolta Dimage A200 are shown. In table 8 the interesting data are shown.

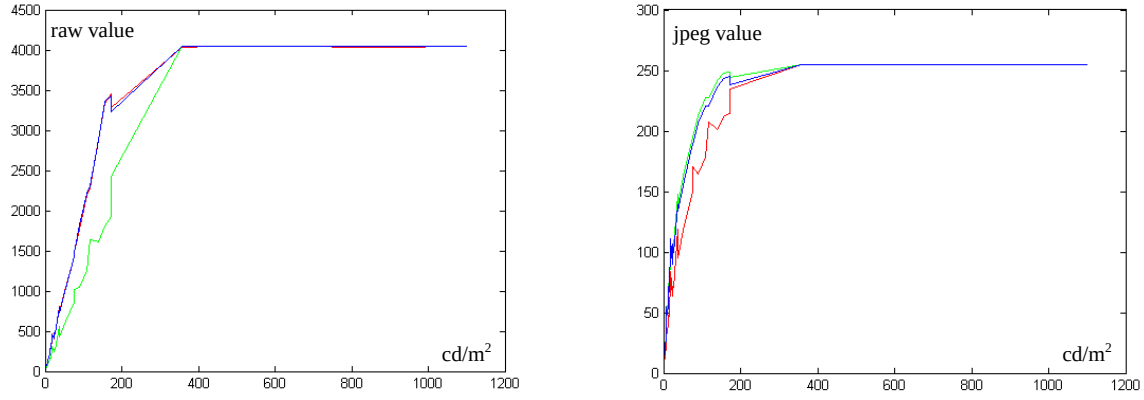
	Physical Values			RAW			JPEG		
Image	EV	cd/m2		red	green	blue	red	green	blue
First	err	err		12.2	3.4	11.0	1.92	2.14	2.05
	1.0	0.28		8.8	10.8	8.0	1.80	2.03	1.93
	4.5	3.10		64.0	47.2	76.4	11.94	17.70	16.29
Second	2.3	0.68		28.4	20.4	25.0	3.66	4.53	4.48
	7.5	25.38		467.8	275.6	473.4	70.45	105.14	99.96
Third	5.6	6.84		150.2	90.6	148.8	29.28	43.59	54.24
	10.3	172.20		3455.0	1934.0	3429.0	214.63	248.82	245.10
Fourth	7.0	18.00		418.2	298.2	435.0	80.67	99.05	109.17
	10.3	172.20		3280.0	2429.2	3227.0	234.63	244.21	238.46
	11.3	356.70		4039.4	4039.8	4039.8	255.00	255.00	255.00
	13.0	1100		4039.6	4040.8	4040.8	255.00	255.00	255.00

Tab. 8: Data regarding the Konica Minolta Dimage A200.

From the data shown in table 8 it is possible to obtain the following considerations:

- The first image has a range between less 1 than and 4.5.
- The second image has a range between 2.3 and 7.5 EV
- The third image has a range from 5.6 to 10.3 EV
- The fourth image has a range between 7 and 13.0 EV
- For the jpeg it is possible to notice that the saturation points occur between 10.3 and 11.3 EV.
- The saturation points for the raw file occur between 10.3 and 11.3 EV, similarly as the jpeg. Due to the big incremental step (10.3-11.3) it is impossible to comprehend if the raw file can record more data than the jpeg.
- For the darkest image the jpeg values are near zero, while for the raw they are between 3.4 and 11.

The test with the lens cover shows that for the raw file the minimum value is 0, the maximum 256, and the mean around 11, for the three channels.
Returning to the data in table 8, it is likely that the first and the seconds rows represents the minimum values recordable by the sensor, before falling into the noise.
The response curve are plotted in figure 42:



(a) (b)

Fig. 42: Response curve for the Minolta camera: (a): jpeg. (b): raw

As for the two Canon, the Raw, until the saturation point, is quite linear to real luminance of the scene, while the jpeg has a gamma correction. Interesting enough in this case, the raw and the jpeg file saturate around the same point.

6.3.5 Experiment conclusions

The experiments shown in this section aim to evaluate the raw and jpeg files from three different cameras: Canon EOS 400D, Canon EOS 7D, Konica Minolta Dimage A200. For all the cameras the raw sensor is able to acquire the data almost linearly proportional to the real light in the scene. The raw converter causes the jpeg to have a gamma corrected shape.
The declared number of bit for the raw file are now compared (table 9) with the real number of bit used to capture the data. The formula to calculate the real number of bit used is: $\log_2$(maximum value – minimum value + 1). For the Canon EOS 7D therefore is: $\log_2$(13584-2010+1) = 13.5 bit.

	Declared	Used
7D	14	13.5
400D	12	11.76
A200	12	11.98

Tab. 9: comparison between declared number of bit and actually used by the cameras.

It is possible to make an approximation regarding the number of EVs that the cameras can capture, before the sensor gets saturated [12] (table 10).

	Raw			Jpeg		
	red	green	blue	red	green	blue
7D	12.9	12.3	12.8	12	12	12
400D	12.7	12	12.9	11.6	11.6	11.6
A200	10.3-11.3	10.3-11.3	10.3-11.3	10.3-11.3	10.3-11.3	10.3-11.3

Tab. 10: number of EVs captured by the cameras.

According to these tests, it is possible to observe that both the Canon cameras can record more data in the raw mode, as they are able to record a more dynamic range (given by the number of EV captured). When converted in jpeg, this advantage gets lost. However the jpeg generated are able to keep a high dynamic range of luminance, 11.6 for the Canon EOS 400D, 12 for the Canon 7D. The Konica Minolta Dimage A200 is able to capture a smaller dynamic range. From the test performed it is not clear if having a raw format of 12 bit can be advantageous respect to the 8 bit jpeg format, as in both cases the saturation point is between 10.3 and 11.3. The big incremental step cannot be useful to appreciate when the real saturation point occurs.

[12] The minimum EV is always taken as 1, both for jpeg and for raw, despite the noise. The saturation point cannot be measure accurately for the three cameras, however the approximation should be realistic.

6.4 Experiments on the HDR format

6.4.1 Introduction

In the previous section we see that raw files use more bit than jpeg to record the data. In some cases this can leads to capture a greater dynamic range. However even if using raw, the high dynamic range that a real scene can achieve, cannot be captured by a single image. Therefore the HDR imaging has been developed to overcome this problem (chapter 5).

In the following section a range of experiments have been performed in order to understand how three HDR software work. In our experiments we tested :

- HDRShop v. 1: software from Paul Debevec (Debevec, et al., 1997).
- A Matlab implementation by me of the Debevec's algorithm (Debevec, et al., 1997).
- Luminance HDR 2.0.0, by Giuseppe Rota, that implements a range of HDR algorithms, among which the Debevec's algorithm, used in these experiments.

The declared intent of these software is to provide a luminance map, linearly proportional, up to a scale factor, to the real luminance values of the scene. The aim of the experiments is to prove if this purpose is verified or not.

6.4.2 Experiments on synthetic images

In the first experiment, the three software are tested on synthetic images, in order to avoid any complication due to the camera (effect of glare in the lens, vignetting, sensor noise). We simulated the behaviour of a scene captured at different exposure times. To achieve this intent, a scene composed of 100 stripes was drawn. This picture represent a real scene with stripes with a different luminance values. To obtain the photos of the scene, we have to change the RGB values of the rectangles in the scene by simulating the changing exposure (example of three different exposures in figure 43). Eleven shots are captured, in a range between -5 EV and +5 EV.

The values of the RGB are calculated using the sigmoid function, that has the typical camera response curve shape (figure 44). The formula of the sigmoid function is:

$$P(t) = \frac{1}{1 + e^{-t}} \tag{3.1}$$

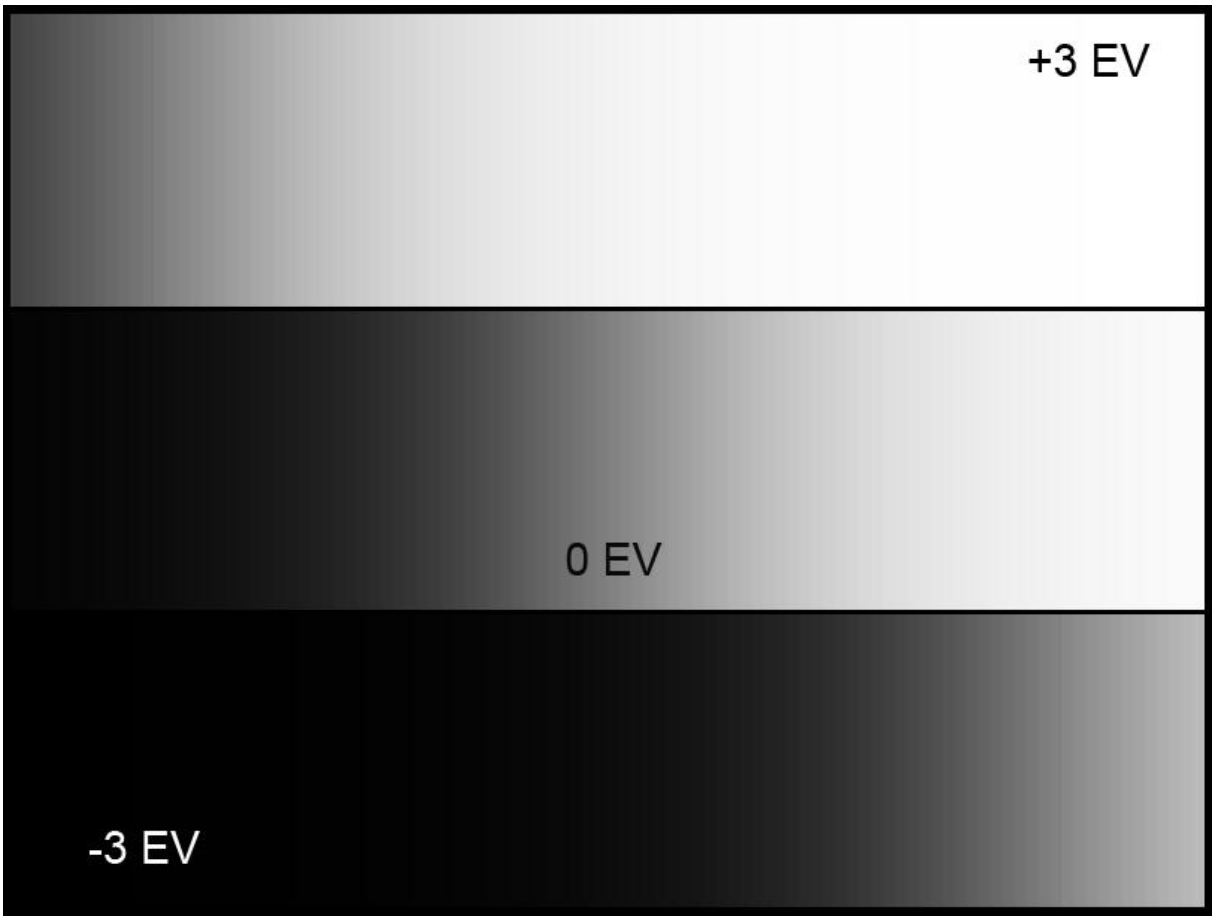


Fig. 43: Three different exposed images.

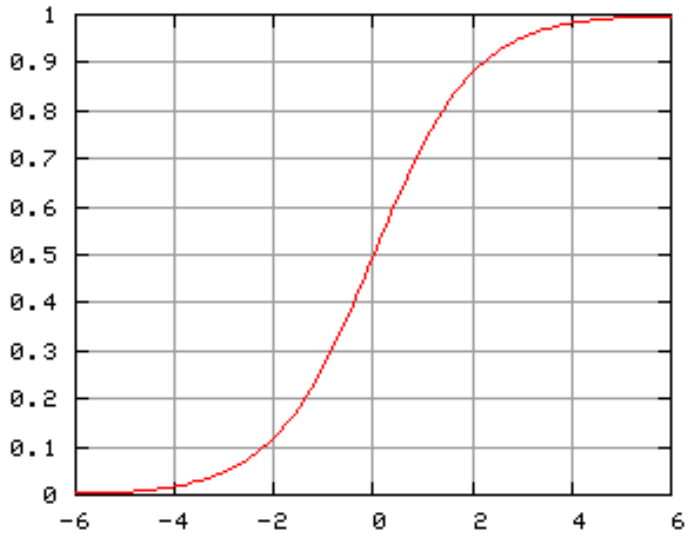


Fig. 44: graph of the sigmoid function.

The first step to test the models is to reconstruct the response curve of the camera (in our case a virtual camera with a sigmoid response curve). The graphs are shown in figure45.

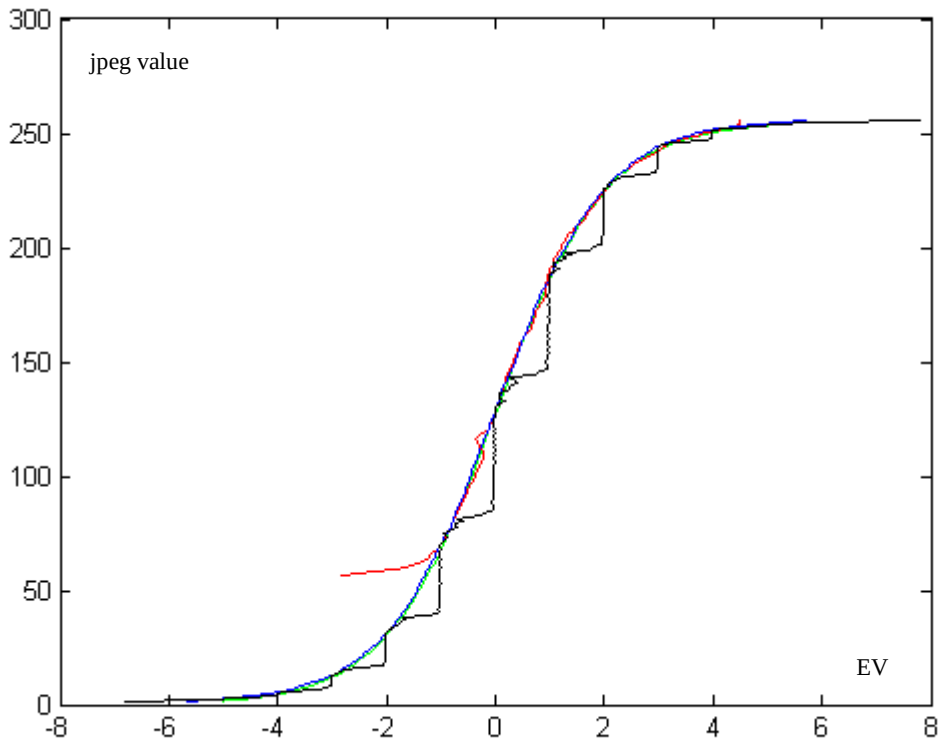


Fig. 45: Reconstruction of the response curve by the three software.

The green curve represent the sigmoid curve used to generate the image, and it is almost perfectly overlapped by the blue curve that represents my Matlab implementation of the Debevec's algorithm. HDRshop curve is the red one. The curve reconstructs in a good way the sigmoid only from -2 EV. Before it is totally wrong (the numerical data are not reported here). Finally the black curve represents the curve generated by Luminance HDR. The curve has a anomalous steps flow, however, it follows the sigmoid function shape.

From these curves the three software can generate three HDR image. Every image has its own unit, depending on the method used. To compare the data, the values are mapped in a range between 0 and 10000. For every of the 100 stripes, the values are extracted from the HDR image and plotted in the figure 46.

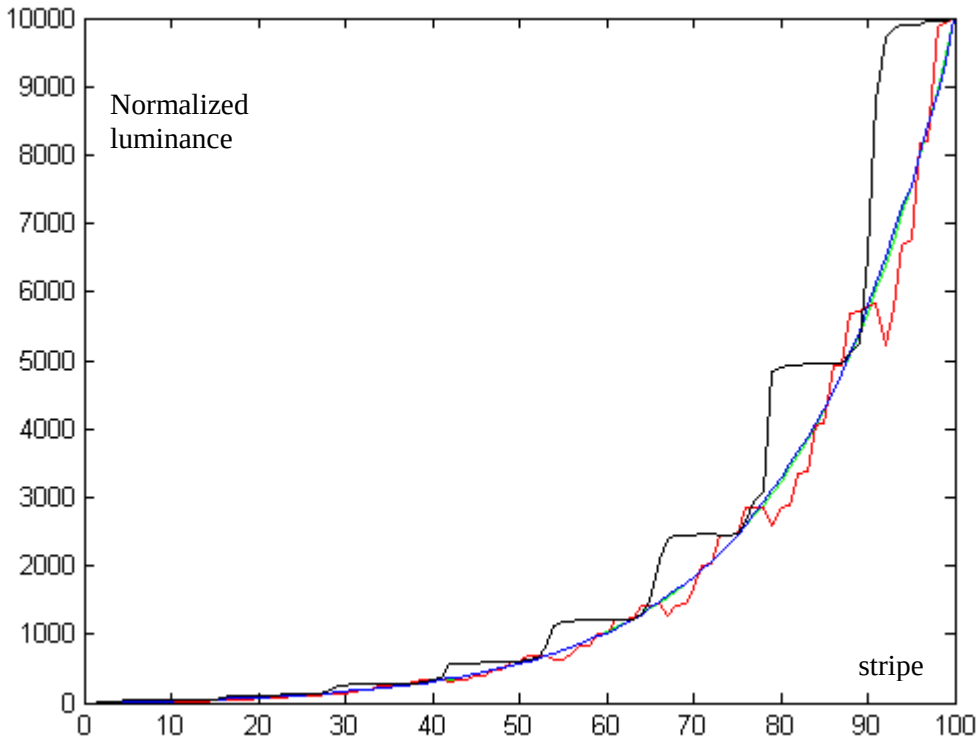


Fig. 46: Reconstruction of the luminance values of the stripes by the three software tested.

The blue curve is the Matlab implementation, and again is overlapped to the blue, that represents the real HDR values for the stripes. The red curve is the HDR Shop output, and the black curve is the Luminance HDR result. From these test it is clear that the three software are able to reconstruct quite well the HDR image. The noise of the HDR Shop curve, and the steps of Luminance HDR can be due to the sampling that the algorithms perform to reconstruct the response curve.

6.4.3 Reconstruction of an HDR image from a real scene

In this experiment a 3D scene is used. The camera is mounted on a tripod in front of the scene (figure 47). The exposure values of the scene has bee measured with the spotmeter, and the result are shown in figure 48.

Fig. 47: 3D scene used to test the HDR algorithms.

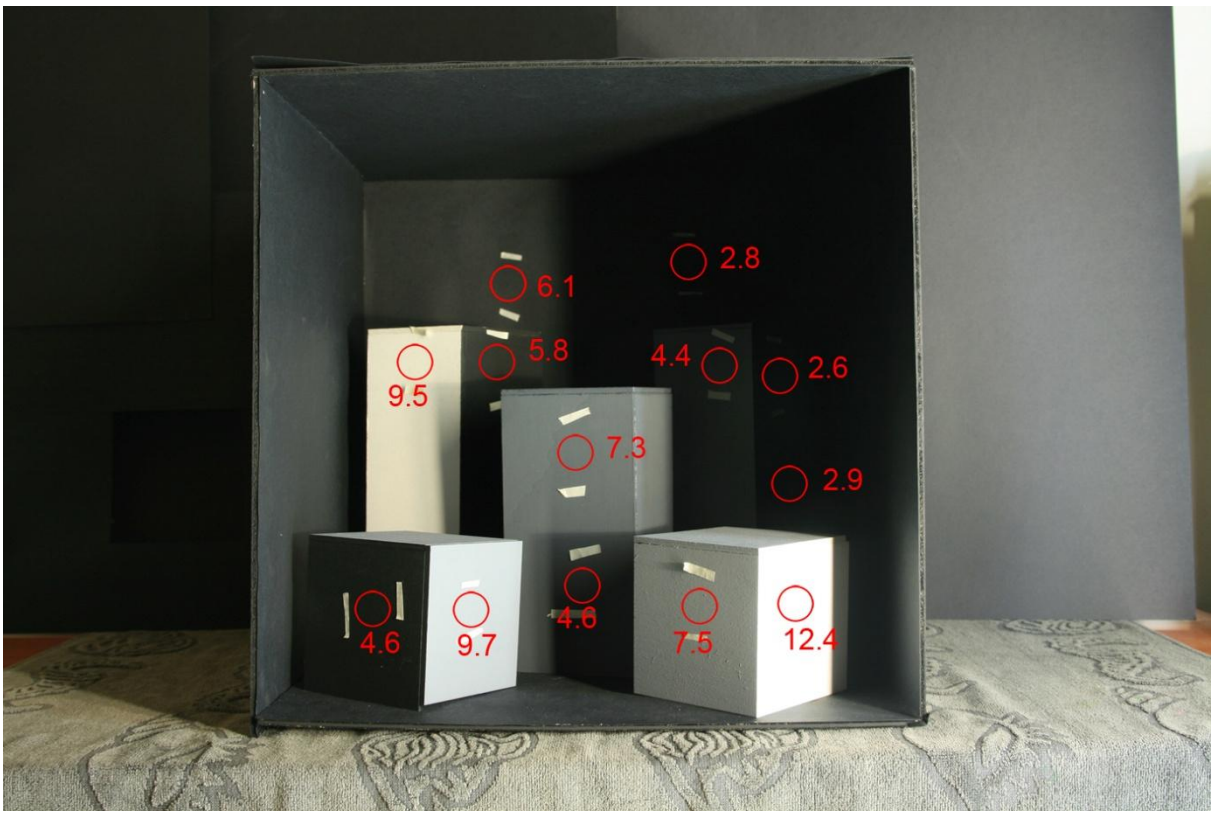


Fig. 48: EV measurements of a range of points.

With the Canon EOS 400D, 14 picture of the scene have been taken with aperture fixed at F/5.6, and ISO 100, varying the exposure time from 15 seconds to 1/500 second (5 shots are shown in figure 49).

Fig. 49: 5 of the 14 shots captured by the Canon EOS 400D

The curves generate by the three software are plotted in figure 50 only for the red channel (the blue and green channel gives similar results):

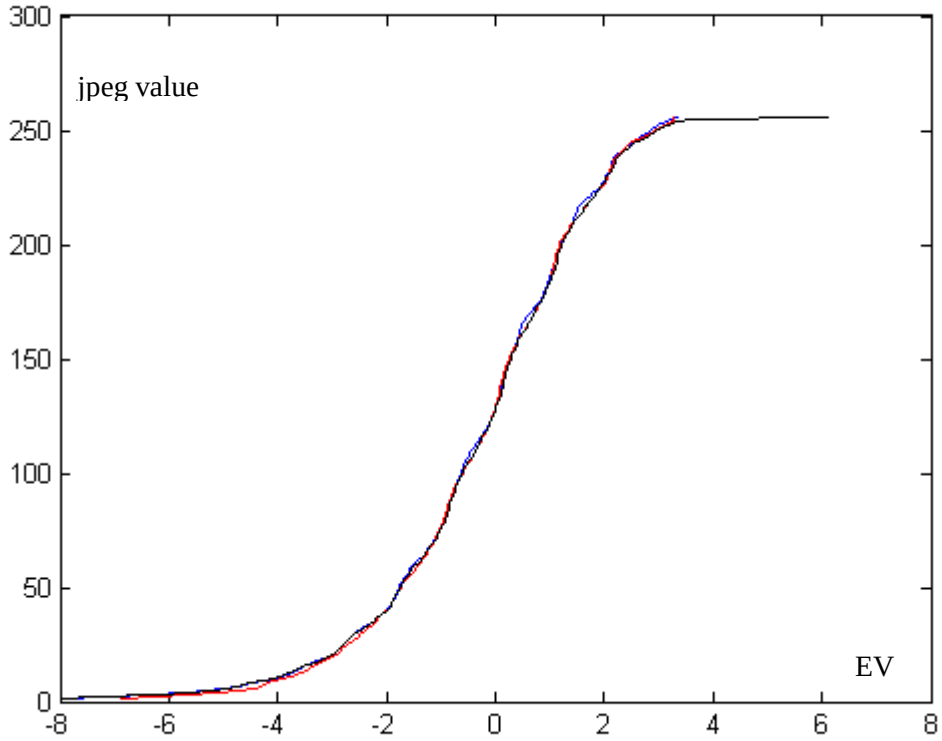


Fig. 50: response curves for the three software.

The blue curve is the Matlab implementation, the red curve is generated by HDR Shop, and the black curve is the Luminance HDR result. It is possible to see that the three curves overlap. From the experiments performed in section 6.3.3, concerning the front illuminated patches, it is possible to calculate the curve of the Canon (in figure 41b it is shown the curve respect to the cd/m^2. Here we have to show the curve respect to the EVs). From figure 51 it is clear that the result shows good consistency. The green curve (originated from the experiment in section 6.3.3) is more subject to noise, but the flow is similar. The curves calculated by the software are the result of an average of more points taken in the image, therefore they have a smooth shape.

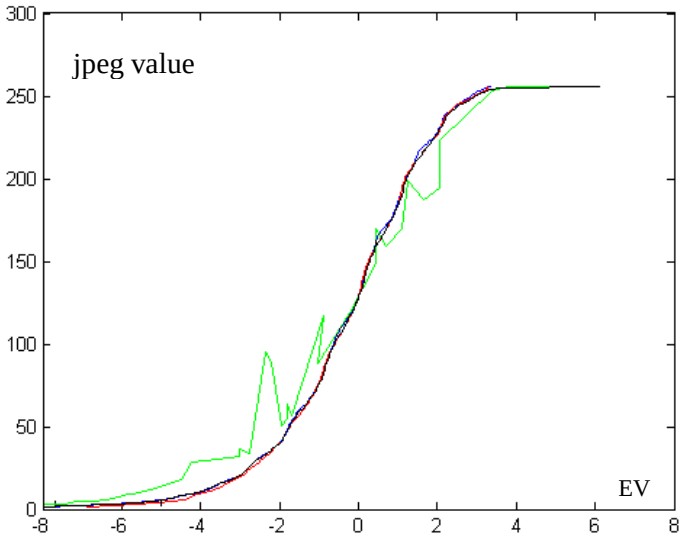


Fig. 51: overlapping of the curve calculated from the experiments with the front lit patches.

From the curves generated by the three software, three HDR images are created. The values for the points shown in figure 48 are extracted and mapped in a range between 10000 and 0 (table 11), the results can be plotted (figure 52):

cd/m2	Matlab	HDR Shop	Lum. HDR
10000	10000	10000	10000
1540	1476	1395	1481
1339	1471	1356	1442
326	333	275	309
283	326	300	297
116	85	69	77
92	93	73	82
33	35	29	30
33	27	22	25
27	27	23	25
2	12	12	11
1	0	0	0
0	11	11	10

Tab. 11: data regarding the values extracted from the HDR images created with the three software.

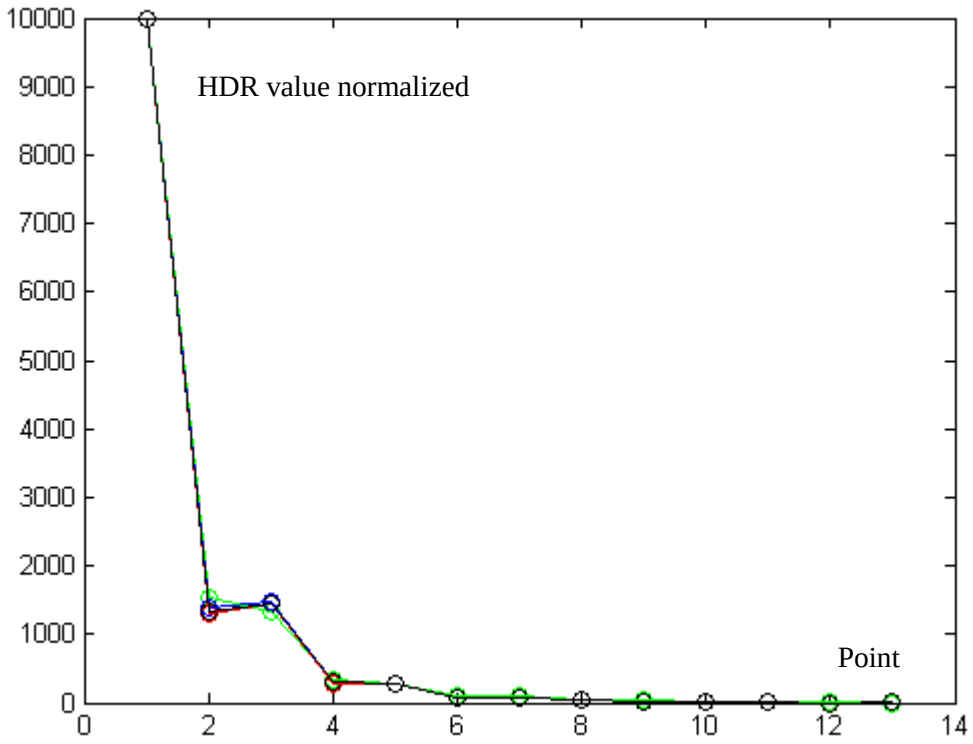


Fig. 52: the HDR data are plotted.

The green curve represents the real EV, the red curve is the HDR Shop result, the blue curve is the Matlab implementation and the black curve is the Luminance HDR reconstruction. The three software are able to reconstruct the luminance of the scene quite well.

6.4.4 Testing the glare in HDR images

In this experiment we are going to evaluate how the glare influences the reconstruction of the HDR image. In (McCann, et al., 2007) it is shown that glare affects the HDR image luminance values.

In our test we placed different semi-transparent patches on a light diffuser back lit by fluorescent lights[13]. The camera was placed in front of them, on a tripod (figure 53).

[13] The lights used are: n°2 Philips TLD 58W/33, n°2 Mazdafluor Blanc Industrie 36W/33, n°1 Philips TLD 18W/33.

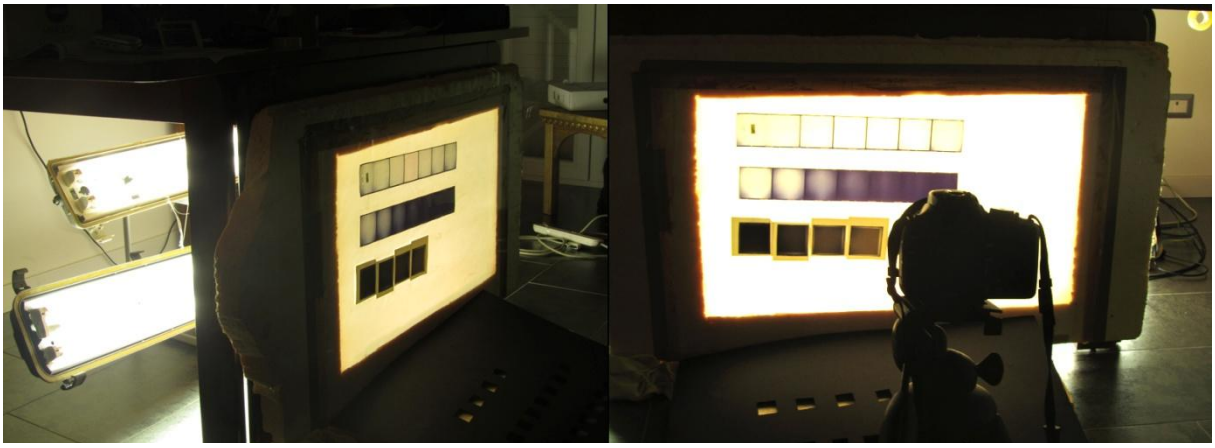
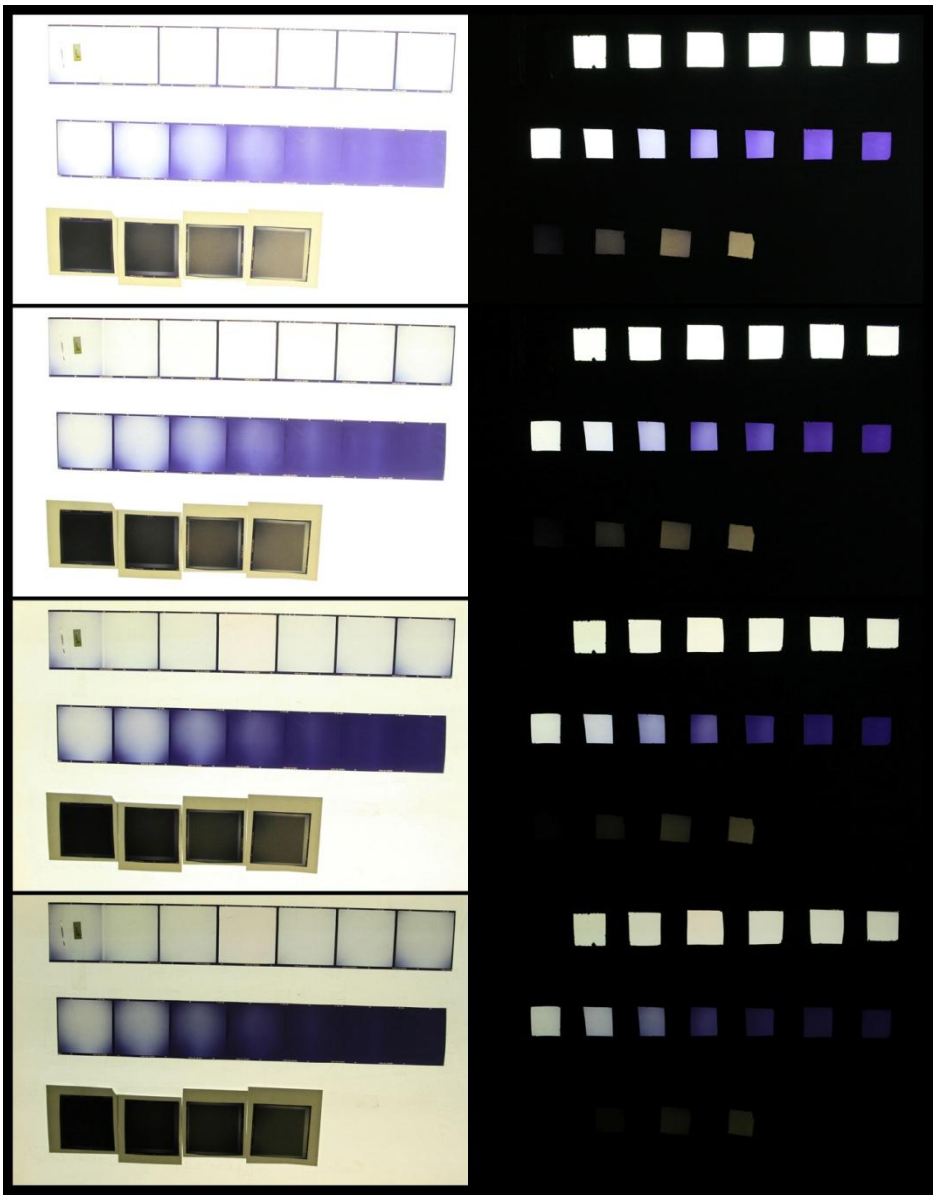


Fig. 53: the setup of the experiment.

Four pairs of images have been captured, with exposure times ranging from 1/15 to 1/2 second; the first set of images are taken with a mask, that have apertures in correspondence of the patches, the second set of images are captured without the mask (figure 54). The patches are the results of pictures previously taken on film. It is possible to observer the vignetting in the corners. Anyway, special attention have been paid in order to measure the centre of the patches to obtain the data in table 12.

The patches themselves, of course, have the same physical values, as the mask does not influence the light coming from the back.



(a) (b)

Fig. 54: pictures captured without the mask (a) and with the mask (b)

Five points have been taken into consideration, P1 is the lightest, and P5 is the darkest (figure 55).

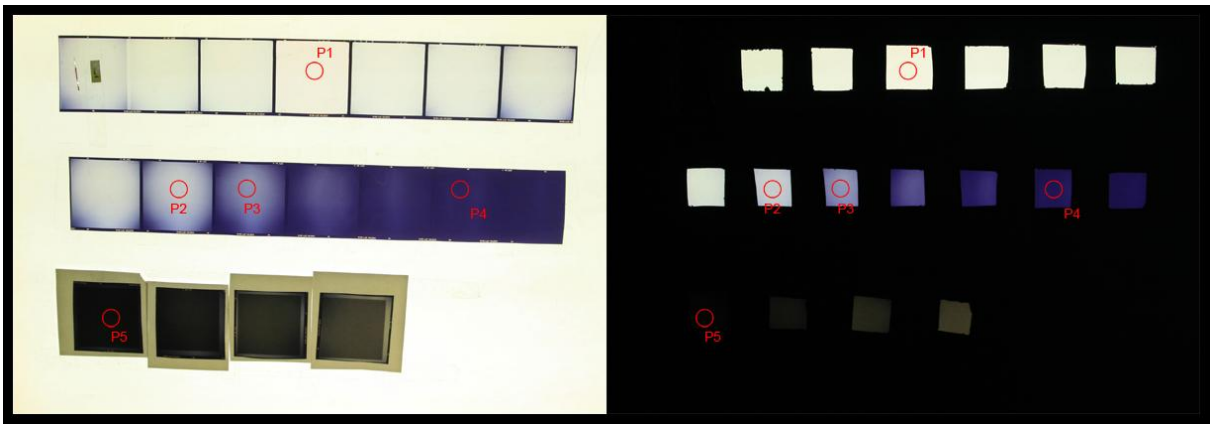


Fig. 55: The same five points taken in consideration, shown on the image without mask and with mask.

These points are measured in Photoshop (an averaged area around the points) for the five shots (table 12).

		Mask			No mask		
		red	green	blue	red	green	blue
P1	1/15	212	208	183	205	200	173
	1/8	244	244	220	243	237	214
	1/4	255	255	255	255	255	255
	1/2	255	255	255	255	255	255
P2	1/15	144	154	150	137	141	140
	1/8	199	206	204	196	199	198
	1/4	235	240	239	233	237	236
	1/2	255	255	255	255	255	255
P3	1/15	87	90	114	82	85	105
	1/8	143	152	184	141	145	175
	1/4	195	204	232	196	201	226
	1/2	234	241	255	237	240	255
P4	1/15	19	16	38	26	23	39
	1/8	33	21	82	47	35	86
	1/4	73	46	145	96	78	155
	1/2	135	99	231	162	141	236
P5	1/15	2	2	2	6	7	6
	1/8	5	6	5	14	14	14
	1/4	11	11	11	26	28	24
	1/2	21	22	22	51	53	51

Tab. 12: data concerning the five points, for every shot.

The data of the table are plotted in figure 56.

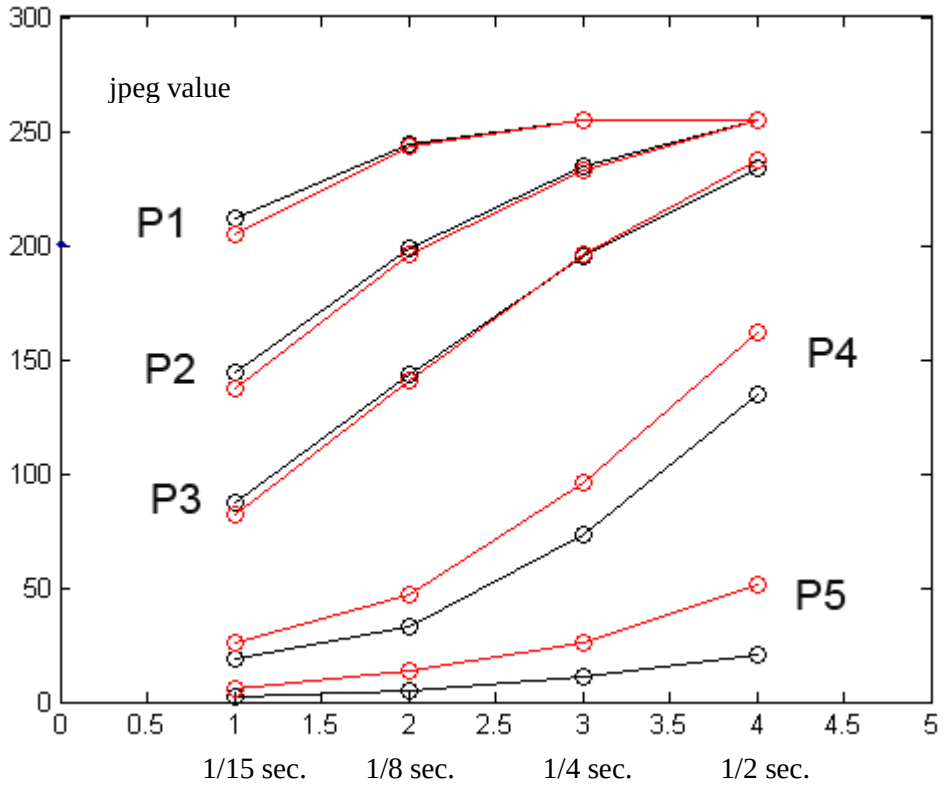
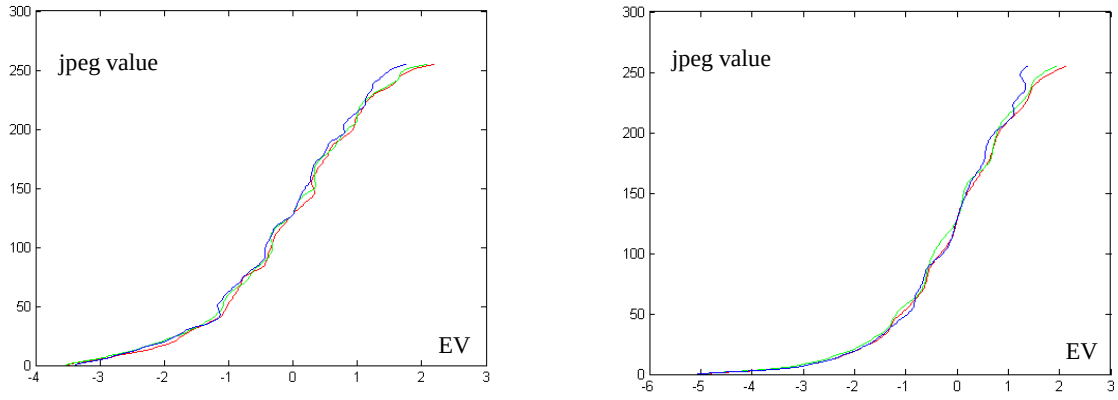


Fig. 56: the data of table 11 are plotted.

On the x-axis 1-2-3-4 represent the shot number (1: 1/15 sec., 2: 1/8 sec., 3: 1/4 sec., 4: 1/2 sec.), on the y-axis the digital value for the red channel (same graphs are obtained for the green and blue channels). Black lines are referred to the measurements of the points with the mask, the red lines represents the measurements of the points without the mask. From the figure it is possible to see that the lightest points has similar values (P1 – P2 – P3), while the darkest point (P4 – P5) are lighter in the images withot the mask. This effect is due to the glare in lens of the camera, that increase the luminance of the points.

The response curves generated by the Matlab implementation of the Debevec's algorithms are plotted in figure 57: (a): without the mask, (b) with the mask.



(a) (b)

Fig. 57: response curves. (a): points without the mask, (b): points with the mask.

Plotting in the same graph the curves for the red channel, for both the masked and no-masked images the result is (figure 58):

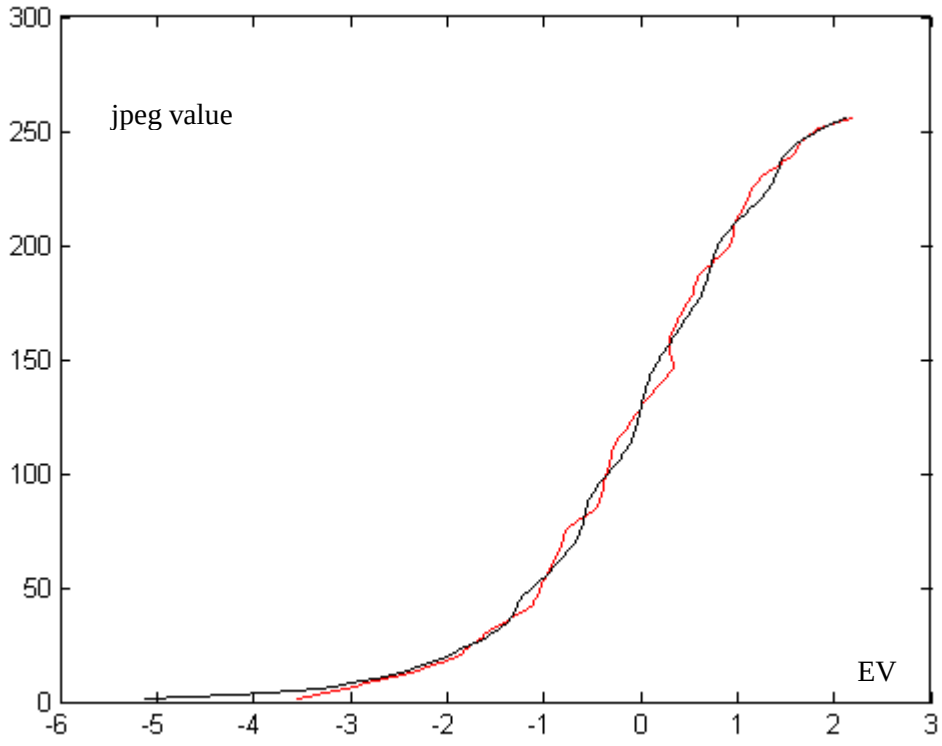


Fig. 58: black curve: response curve of the point without the mask, red: response curve of the point with the mask

The black line represents the curve reconstructed starting from the masked images, the red represents the curve reconstructed starting from the not masked images. As it is possible to

see, they overlap quite well, even if the reconstruction done using the pictures taken with the mask (black curve) generates values also between -5 and -4 EV.

Finally, the HDR values for the five points are given in table 13:

	Mask			No Mask		
	Red	Green	Blue	Red	Green	Blue
P1	41.8500	35.1000	24.8000	38.8500	34.5500	23.2500
P2	17.2000	17.6000	17.6250	16.9750	16.6500	17.9750
P3	8.6875	8.8750	13.6000	8.9875	8.8875	12.3375
P4	2.0438	1.2063	4.5375	2.8500	2.0625	5.7313
P5	0.2844	0.2570	0.2863	0.7828	0.7414	0.7672

Tab. 13: HDR values for the two situations.

These data can be plotted: the graph of the five points for the red channel is shown in figure 59.

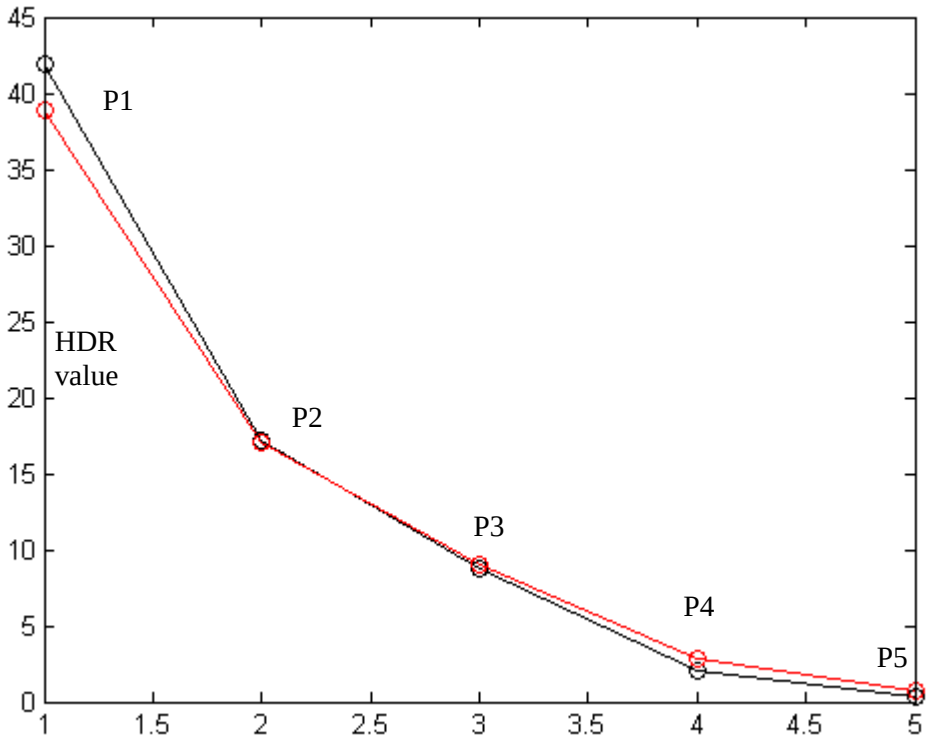


Fig. 59: HDR values for the two situations.

The black lines is referred to the measurements of the points with the mask, the red line represents the measurements of the points without the mask. From the graph the

reconstruction seems good, however calculating the percent relative difference between corresponding points in the two cases (with/without mask) as:

$$Diff. = \frac{abs(P_m - P_w)}{\max\left(abs(P_m) - abs(P_w)\right)} * 100$$

Where P_m is the point (from P1 to P5) covered with the mask, and P_w is the point (P1-P5) without the mask, the result is (table 14):

Point	Diff.
P1	7.1685
P2	1.3081
P3	3.3380
P4	28.2877
P5	63.6689

Tab. 14: percent relative difference of corresponding points

It is possible to see that for the darker points the percent relative difference becomes great. This is an effect due to the glare, that affects more the darker points without the mask, increasing the luminance captured by the camera.
From this test it is clear that, as shown by McCann and Rizzi (2007), the glare can influence how the darker points are captured. The algorithm proves to be quite robust, in absolute percentage, but in a relative evaluation the error on the dark points becomes considerable. Probably, increasing the exposure time, the glare will affect even more the data. This test will be performed in future.

6.5 Conclusions

Many tests have been performed to understand how the luminance of the scene is captured by digital cameras and to learn what aspects occur during the creation of a HDR version of the scene. The results show that the images are:

- Camera-dependent: the dynamic range captured depends by some in-built features and by some operations of post-processing of the different cameras.
- Glare-dependent: glare changes the luminance values of the pixels in the scene, changing the calculation of the response curve.

The raw format is able to record the luminance of the scene linearly proportional to the real world scene. We tested three cameras, and in two cases (Canon cameras) the raw file extends the dynamic range of captured luminance respect to the jpeg. In these cases an user can take advantage in shooting in raw mode and processing the file with ad hoc software. It is possible in fact to control more the intensity of the light.

Finally a range of experiments on high dynamic range have been performed. The dynamic range of a scene is the ratio between the most highlightened point and the less highlightened point of the scene. It depends on:

- The intensity and position of the source of light;
- The geometry of the scene that create gradient and edges and contribute to define the most and the less higlightened point;
- The glare, a perceptual spread of light that modify the values of luminance in the scene.

A photography is not a perfect copy of the real world. The luminance values of a real scene exceeds by many order of magnitude the values that can be represented in a single photographic image. In our test the objects photographed has a maximum intensity around 13 EV, corresponding approximately to 1100 cd/m^2 while, in a typical sunny day, the luminance of a natural scene can be greater (also 10^4-10^5 cd/m^2). In our test with back lit patches, the glare influence the capturing of the luminance of the darkest points, therefore it is likely to suppose that in a natural scene, the glare can affect even more the dark areas of an image. It is possible by using the HDR technique to extend partially the dynamic range of an image, however the efficiency of these algorithms are limited by glare, which damages both acquisition and the process of merging (by changing the pixel values of the scene) and therefore it cannot be considered an accurate representation of the luminance of the scene.

7. Retinex and iCAM06: experiments

In the last decades the appearance topic has been widely considered. Appearance describes how a colour is perceived by the human observer. As we have seen in chapter 2 the same colour stimulus can be perceived differently depending on the context where it's observed.
Classic colorimetry is based on the extra-context stimulus, therefore more complex models are needed, to be able to study the colour in different viewing conditions. Various theories have been formalized, giving rise to chromatic appearance models (CAM), the aim of which is to simulate the behaviour of the human visual system (HVS).
As follows some experiments are described in order to obtain a quantitative evaluation of the appearance (based on the McCann et al. experiments (2009)).
A selection of observers were asked to select and apply colours , using different rendering techniques, to a graphic reproduction of the scene in front of them. The aim is to analyse how the eye – brain system perceives the scene. Several shots of the scene have been taken with different exposure times. From these images a high dynamic range image (HDR) has been assembled and used as input to some colour appearance models.
To estimate qualitatively and quantitatively the models, the results of the computations will be compared with the colours selected by the observers. In particular the considered models are that previously described: iCAM06 and a HDR version of RSR. The visualization of the idea is shown in figure 60: the HDR image is a representation of the real scene, the observers selected colours to reproduce the appearance of the reality, the algorithms reproduce the appearance staring from the HDR image.

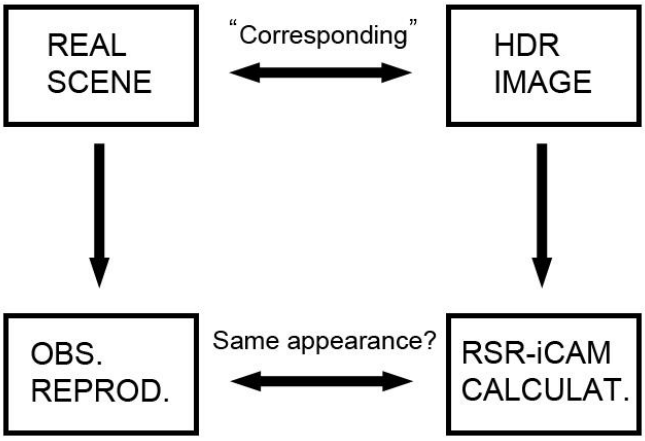


Fig. 60: visualization of the idea of the experiments.

7.1 Description of the experiments.

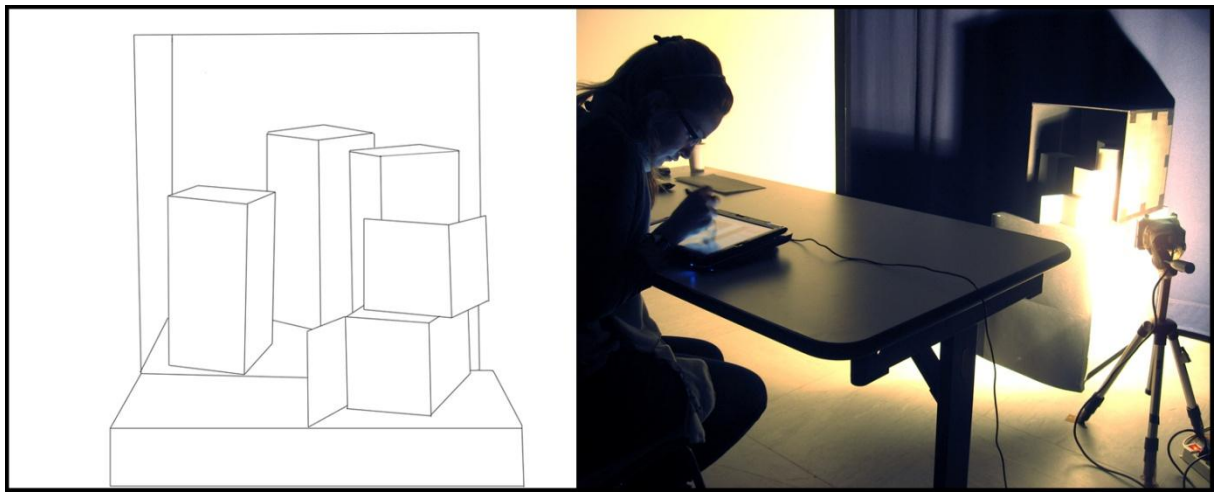
To study different properties of the HVS, three different experiments have been performed. Two real scene have been constructed with a collection of plywood blocks, painted with acrylic colour and placed inside a cube, which is constructed of plywood, covered with black cardboard, and has a side opening. In the first experiment the blocks has been coloured in shaded grey. In order to produce a high dynamic range scene, a halogen spotlight with a colour temperature of 2910 K was used to highlight the objects. The aim of this experiment is to investigate if the simultaneous contrast phenomenon subsists for the human observers in a three dimensional scene, and how it is computed by the CAMs (section 7.2). In the second experiment, coloured objects have been used in place of those in shaded grey. The spot light is replaced with a lamp that has a strong green colour cast. The aim is to study the chromatic constancy phenomenon in the observers, and the simulation computed by the CAMs (section 7.3). The objects used in the second experiment are used also in the third. However, in this experiment it is used a tungsten lamp. The aim is still to study the chromatic constancy phenomenon both in the observers and CAMs (section 7.4).

Once the scene was set up, a graphical reproduction is created, which outlines trace out the edges of shapes (figure 61, a). In order to replicate their sensation of the scene the observers were asked to colour the faces of the reproduced scene. Subjects were divided in two groups. The first group had attended arts school and were asked to use pigments to paint the scene. The second group who had no specific ability persons, were assigned to select colours and fill each face in Photoshop.

7.2 First experiment

7.2.1 Performing the experiment

The painted shaded grey blocks were placed inside the cube so that the light illuminates the front faces of the objects isolating the faces a the bottom in shadow.
A Minolta (Spotmeter F) light meter was used to measure the exposure values (EVs) of many points on the faces of the blocks figure 62, from the point of view of the observer.



(a) (b)

Fig. 61: a: the contoured image. b: a phase of the experiment.

The scene is built so that the face painted with white is placed inside the cube (figure 62: face A, red contoured), and therefore in the shadow, reflected approximately the same amount of light of a grey coloured face of another cube placed in the frontal part of the scene (figure 62: face B, blue contoured). Furthermore, the orientation of the scene in respect of the light is studied in order to simulate a kind of "3D simultaneous contrast". In fact the surround of the face A is darker than the face itself and the surround of the face B is lighter than the face itself.
Figure 61a illustrates the representation of the scene, which includes only the edges of the blocks. Each observer was supplied with this image either printed or as a digital file.

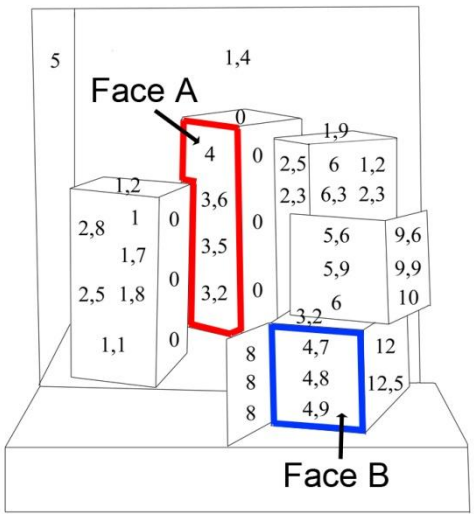


Fig. 62: EV Map of the real scene measured with light meter.

The observers were requested to colour the final image as close as possible to a photo or a realistic canvas. Group 1 used pigment colours on paper, group 2 filled each face with Photoshop, adding eventually the shadows.

In figure 61b it is visible the situation where the human observer operate. The source of light that illuminates the scene is different from the source of light under which the observers operate. The aim of the experiment is to investigate which colours are chosen by the observers, as they can be seen under natural light, while they are looking at a scene illuminated with a different light. In order that colour constancy occurs in this case, it is necessary that the phenomenon is almost instantaneous, and not the result of an adaptation that takes some time. In his paper, Land (1977) describing the experiment (see also chapter 3), in which two different patches but reflecting the same amount of light are perceived as different, hypothesizes that is not matter of adaptation. He write: "*In order to demonstrate that the colour sensations in these experiments do not involve extensive chromatic adaptation of retinal pigments the projectors are equipped with synchronized shutters so that the Mondrians can be viewed in a brief flash, a tenth of a second or less in duration. Regardless of the brevity of observation the results of the demonstrations are not altered. Thus one can say that neither chromatic adaptation nor eye motion is involved in producing the observed colours."* (Land, 1977 p. 9)

Therefore it is necessary to distinguish the dark and light adaptation, that is a physiological process, from the chromatic adaptation, that is an hypothesis to explain colour constancy (von Kries, 1902), (Breneman, 1987)(Fairchild, 2005) but is still an open topic to discuss (McCann, 2004).

The question affects the final result. Our task is to point out the colour sensation, that is the chromatic appearance without any interference of the recognition mechanism and/or memory. As this distinction is not easy to explain, we try to ask a question related to common experience. Ask to reproduce the scene in a realistic way it is similar to ask to reproduce its appearance. To this end it is interesting the McCann's work about sensation and perception (McCann, et al., 1983), (see also chapter 1)

Figure 63 demonstrates a range of results. The observers have used different colouring techniques.

Until now ten experiments have been performed. The results are summarized in table 15. The table contains the RGB values for the two faces A and B. Subjects (1), (2), (7) belong to the group with artistic ability, and used pigments. The others used Photoshop tools, including brushes, paint bucket and gradient tool. To obtained the RGB values of the surfaces coloured by the group with artistic ability, it has been used a spectrophotometer to meausure the points on the surface, and then the spectra have been converted in the sRGB linear colour space.

The first macro column refers to face A (painted with white). The second macro column refers to face B (painted with grey, but reflects more light). It is also reported the grey value for every point obtained as a weighted average: grey=0.3*R + 0.59*G + 0.11*B [14]. The last column of the table contains the ratio between A and B face values. If the ratio is equal to one the two faces are perceived as equal, if the ratio is greater than one A face is perceived as lighter than B. Note that eight times out of ten A face is perceived lighter than B, even if the latter reflects more light.

[14] The formula derive from that used in the colour television system to find the Y component, that is the relative luminance luminance (Oleari, 2008 p. 346), (Hunt, 1995 p. 394-395).

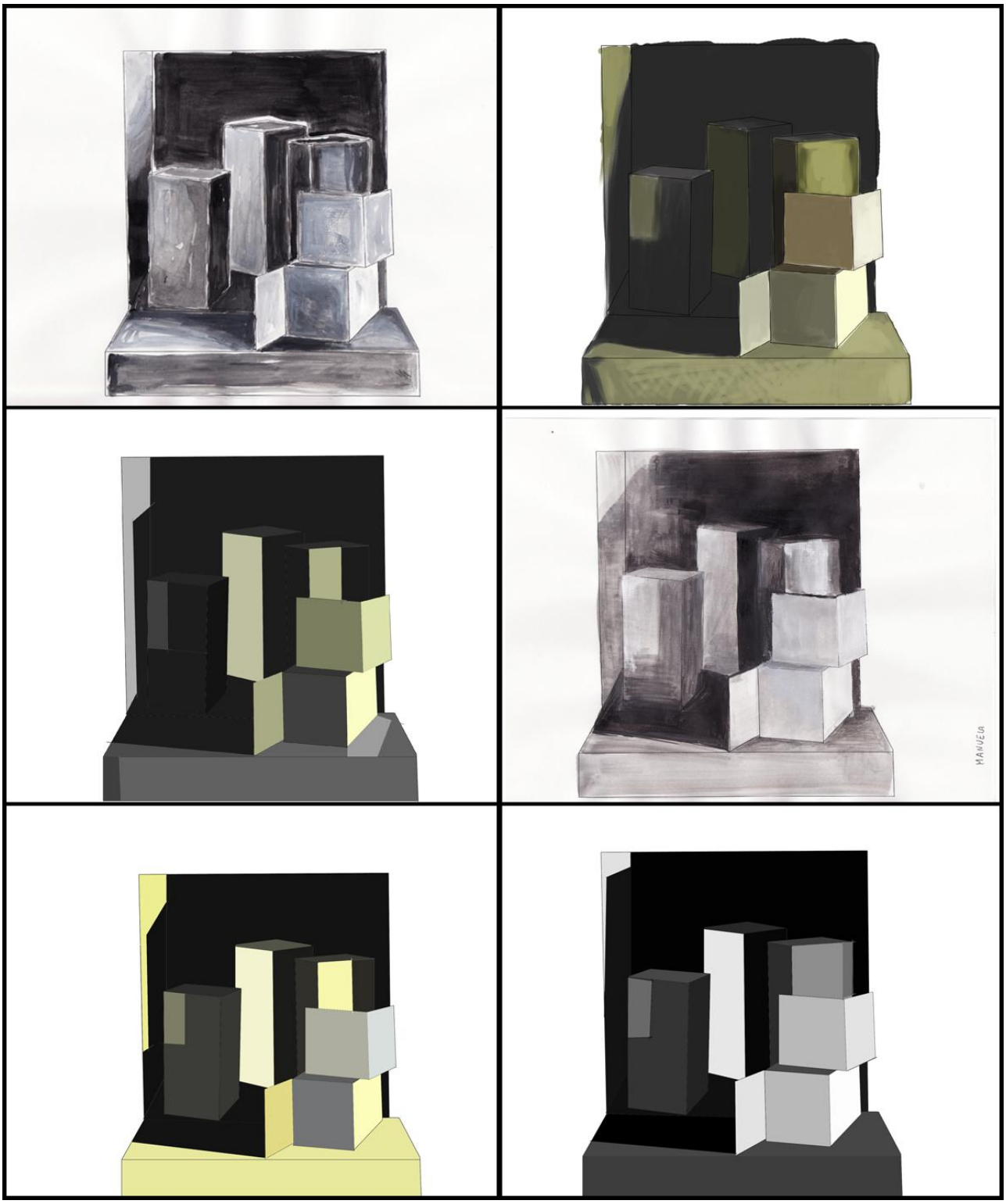


Fig. 63: some reproductions of the observers.

7.2.2 Filtering with Retinex and iCAM06

The image in input to the model has been created with HDR Shop, starting from eleven shoots, keeping an aperture fixed to F/3.5 and varying the exposure time from 15 sec to 1/60 sec (figure 64).

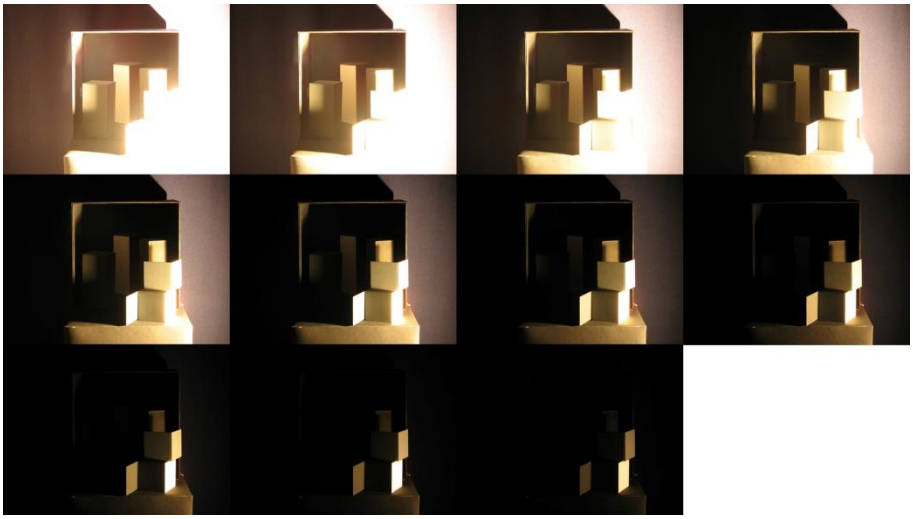


Fig. 64: the eleven shoots used to creation of the HDR image.

The results of the RSR and iCAM06 filtering is shown in figures 65-67. In the case of iCAM06 we filtered the HDR image with D=1 (total chromatic adaptation), and D=0.5 (partial chromatic adaptation).

It is possible to ascertain the removal of the yellow cast of the halogen light, reordered by the photos (the white balance was set on *natural light*). More considerations about this behaviour will be discussed later, after the performance of some experiments related to the colour constancy phenomenon. Furthermore the local contrast has been enhanced, compressing the global contrast in order to map the HDR image in a LDR one. This behaviour is also noticed in the images produced by the human observers.

Finally, note how the algorithm is able to simulate the simultaneous contrast phenomenon in a 3D space. This fact is noticed both visually (figures 65-67) and quantitatively, measuring the filtered image as previously depicted. The measurements related to the considered faces are indicated at the bottom of table 15. RSR aims to make face A lighter, and face B darker, according to the experimental data. The iCAM06 algorithm is also able to simulate the phenomenon, as clearly visible in table 15. The difference between the faces A and B using the Retinex filtering demonstrates a more exaggerated effect. Face B is much darker than A, resulting in a greater ratio than the iCAM06 result. Therefore the mean is closer to the

iCAM06 results rather than RSR. This result can be better visualized in the graph which represents the data (figure 68)

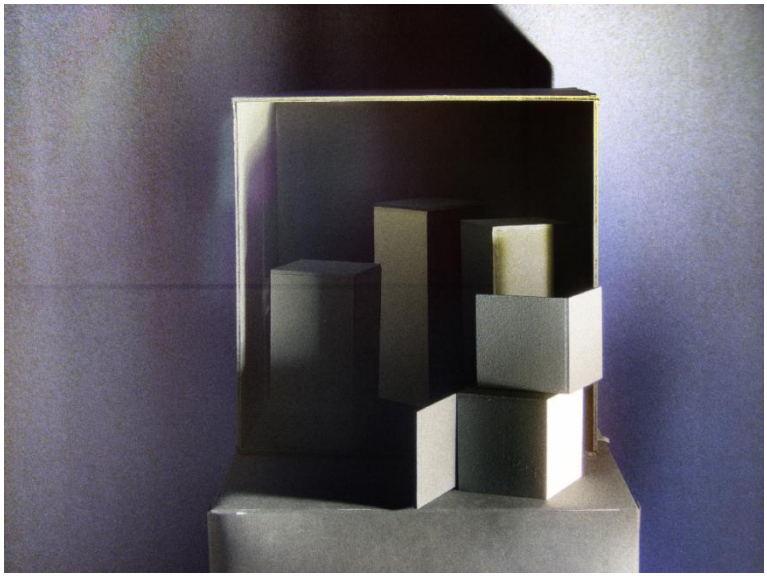


Fig. 65: RSR filtering.

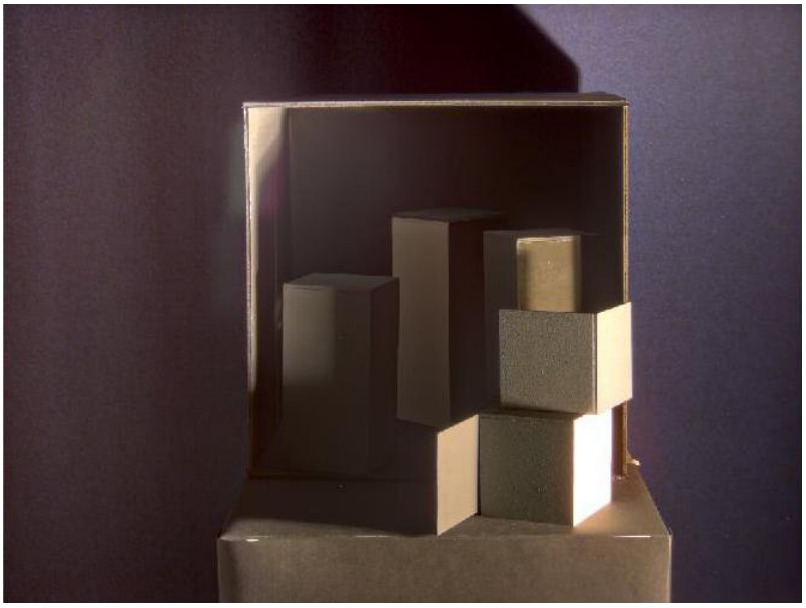


Fig. 66: iCAM06 filtering with D=0.5.

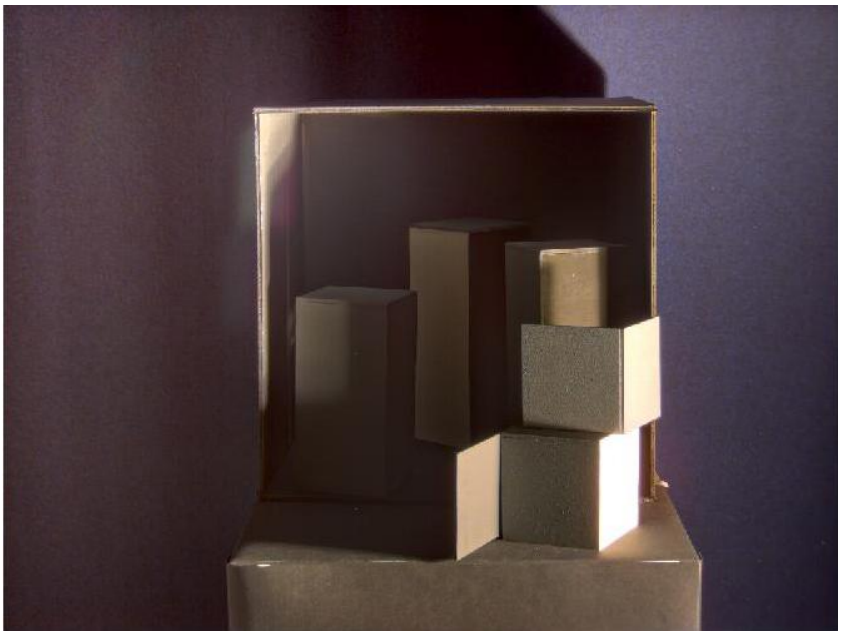


Fig. 67: iCAM06 filtering with D=1

	A					B				A/B
	R	G	B	Grey		R	G	B	Grey	Grey
Obs. 1	245	239	139	230		40	39	39	39	6.13
Obs. 2	132	134	90	129		12	12	12	12	10.60
Obs. 3	222	222	222	222		134	134	134	134	1.65
Obs. 4	228	228	159	221		44	44	46	44	5.13
Obs. 5	203	203	203	203		124	124	124	123	1.64
Obs. 6	255	255	255	255		30	30	30	30	8.54
Obs. 7	160	155	160	158		49	54	71	54	2.87
Obs. 8	149	142	149	145		142	141	155	143	1.01
Obs. 9	9	10	5	9		31	30	18	29	0.33
Obs. 10	109	129	140	124		144	186	196	174	0.69
Mean	155	157	137	154		65	67	68	66	2.35
Retinex	33	29	17	29		3	4	3	3	7.31
iCAM 1	20	13	7	14		15	12	7	12	1.09
iCAM 05	23	15	9	17		15	11	6	11	1.40

Tab. 15: data regarding face A and face B. The last column reports the ratio between the two faces.

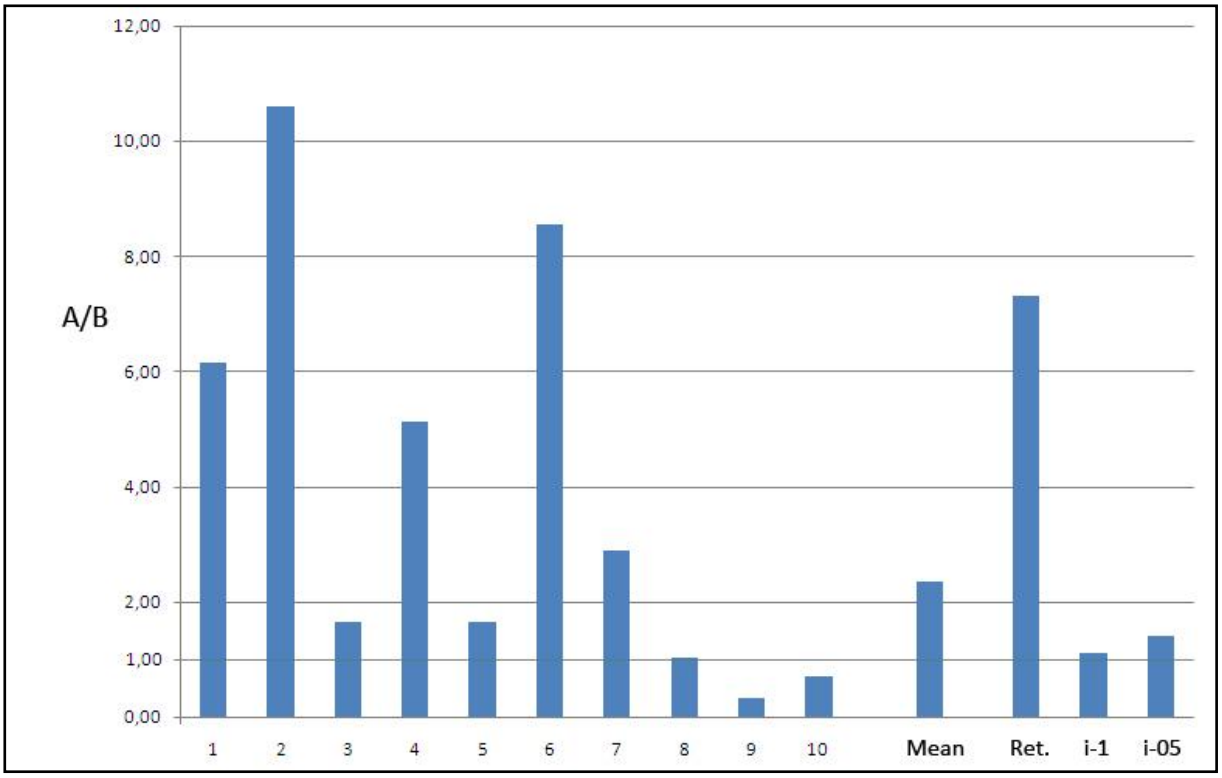


Fig. 68: Graph representing the data of the Table 15.

7.3 Second experiment

In the second experiment five hues have been used to paint the faces of the solids: magenta, green, cyan, yellow, white. A light bulb with a strong green cast has been used to highlight the scene which is placed sideways.

As in the previous experiment eleven picture have been taken to reconstruct the HDR image (figure 69) We decide to keep untouched the white balance (*natural light*). To the observers the same request as in experiment one has been made: to paint the scene in order to reproduce the appearance. The aim is evaluate if the colour constancy phenomenon occurs even under these extreme conditions.

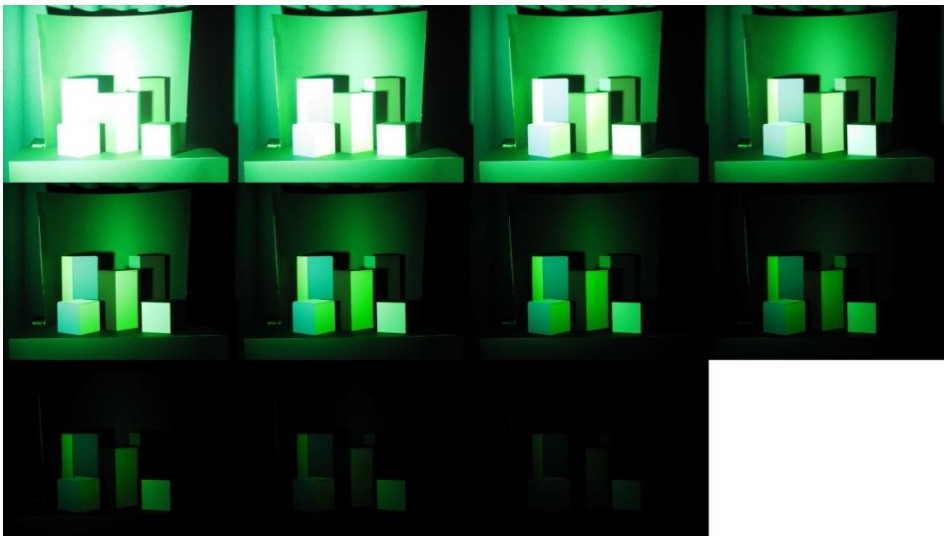


Fig. 69: Eleven shots to reconstruct the HDR image.

7.3.1 Performing the experiments and results

The coloured results of some observers can be seen in figure 70. Note how the colour cast is strongly perceived by the observer. The front face of the cube on the right is in fact painted with white, but almost all the observers have coloured it green. Observing the images it is possible to guess how far we are from the colour constancy. However some observers assert to be able to distinguish the “real” colours of the solids.

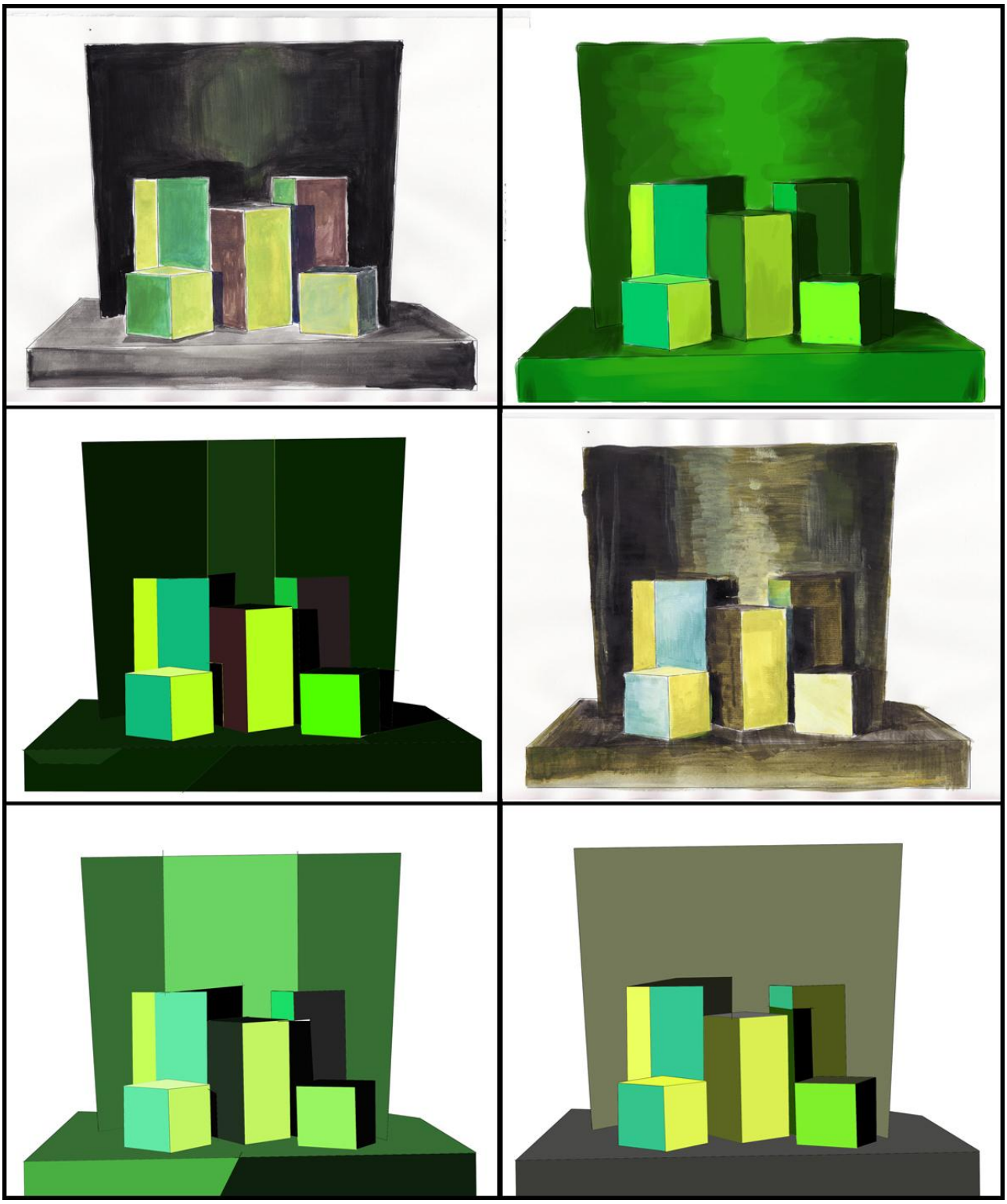


Fig. 70: Some reconstructions of the observers.

As in the previous experiment, the HDR image have been filtering with iCAM06 and RSR.
It is necessary to define the parameters used for the iCAM06 filtering, in order to understand how the chromatic adaptation works.

The actual input for the algorithm are the XYZ values of the image. As in an HDR image the data are stored as RGB, it is necessary to pass from RGB to XYZ. As in a HDR image the RGB values are linear to absolute luminance, this step can be accomplished with a 3x3 matrix, without gamma correction. Since the specific camera characteristics are usually unknown, in equation 4.38 (chapter 4) the matrix chosen was the matrix that allows to pass from the sRGB in XYZ space. It is important to note that this colour space has a D65 white point. As in our experiments the white point is not D65, it would be necessary to use another matrix to pass from RGB to XYZ. This step refers to a global white balance, done before running the iCAM06 model. However, as aim of the experiment is evaluate the model even if the source of light is unknown (or can be not measured), it is possible to apply only the local chromatic adaptation (equations 4.40, 4.41, 4.42) In most cases they have a very similar behaviour. Furthermore advantages can derived from this choice when different sources of light are present in the scene.

In order to better analyze the properties of this model, the HDR image is filtered in two ways with iCAM06: one setting D=0.5, that means partial adaptation (figure 72) one setting D=1, that means total chromatic adaptation (figure 73). It is possible to note how the images filtered by the two algorithms, both iCAM06 (figures 72-73) and RSR (figure 71) present an almost perfect colour constancy, visible especially in the white face of the frontal cube at left. In fact, due to the strong colour cast of the light, this particular face is calculated by the algorithms as white but the observers perceived it as greenish hue (figure 70).

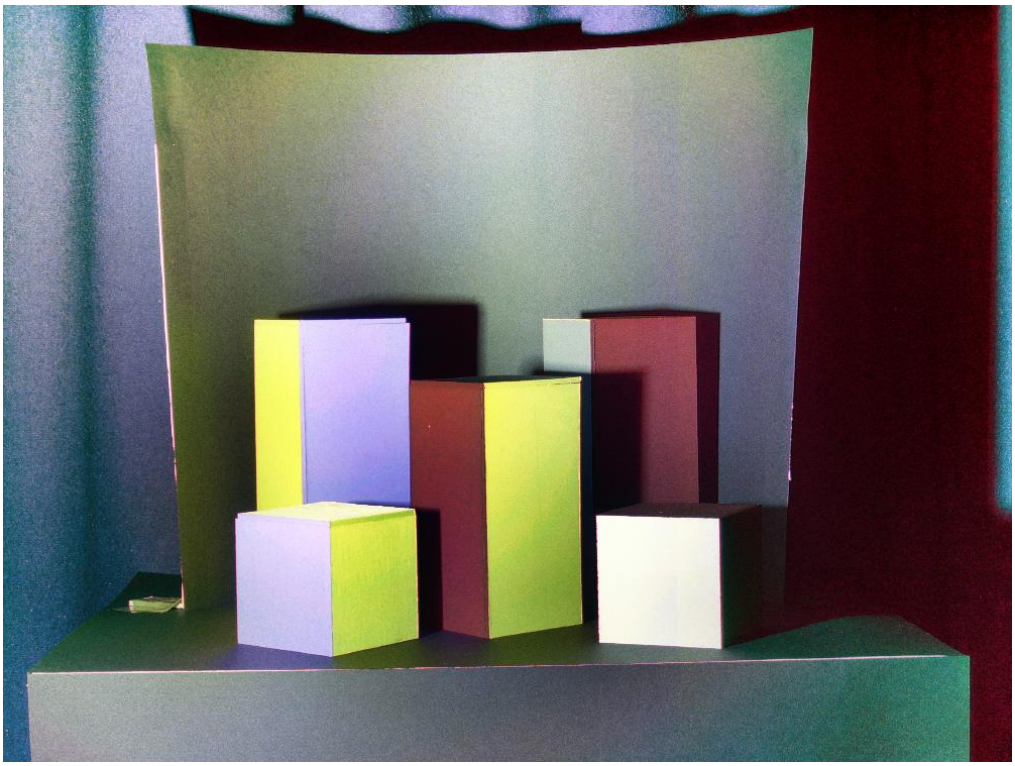


Fig. 71: RSR filtering.

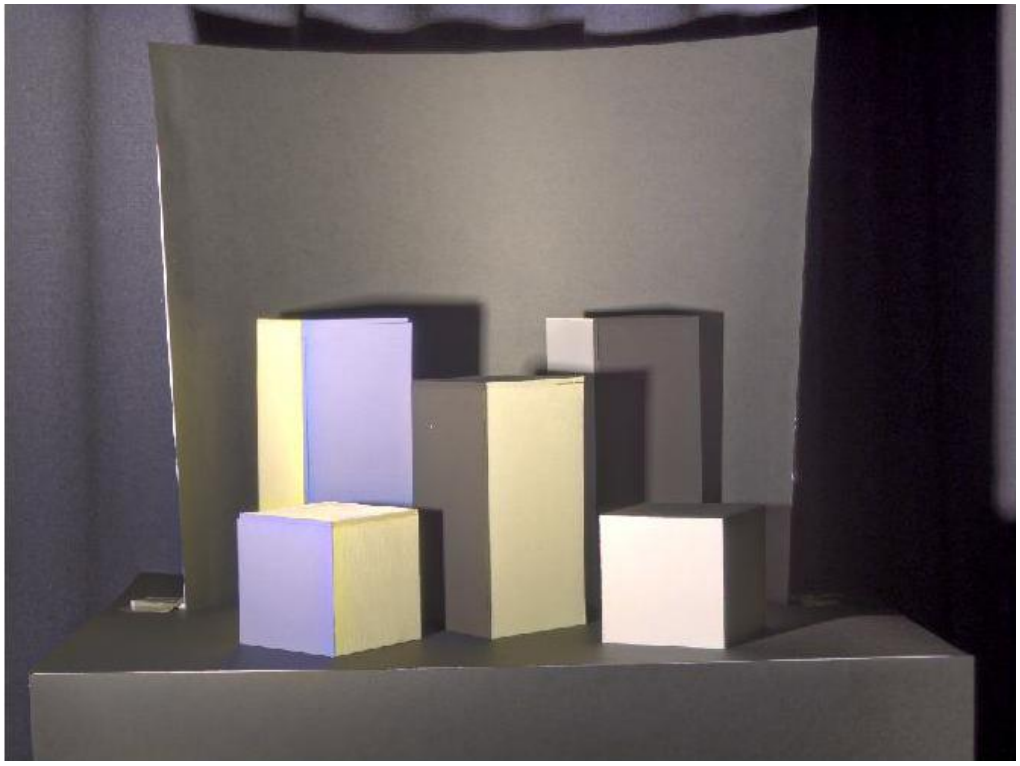


Fig. 72: iCAM06 filtering, with D=0.5.

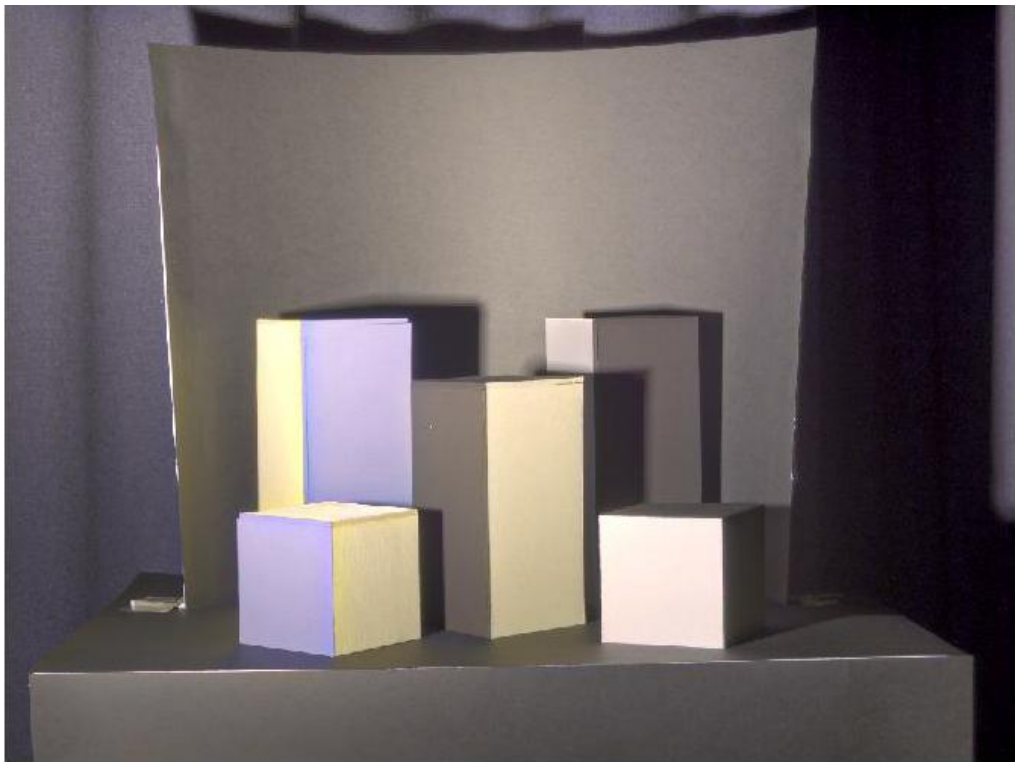


Fig. 73: iCAM06 filtering with D=1.

In order to obtain quantitative evaluations of the results, we decided to pass from the RGB space in the CIELAB (white point: D50) uniform colour space (table 16), considering ten points on the image (the points taken in account are shown in figure 74).

With regards to the luminance value of the HDR images (that is considered as an approximation of the luminance of the real scene), they have been normalized in respect to the global maximum of the image.

Fig. 74: Points taken in account for the comparison

Once the images are in the CIELAB colour space, the $\Delta E^*_{ab} = \sqrt{(\Delta L^*)^2 + (\Delta a^*)^2 + (\Delta b^*)^2}$ is used to compare the colours. The aim is ascertain how what has been painted by the observers moves away from the real scene, in other words, how the appearance moves away from the reality. For a certain number of points, the ΔE^*_{ab} between the colour chosen by the observer and the corresponding colour of the HDR has been calculated table 17, left. As we are particularly interested in the chromaticity in table 16, right, only the chromatic difference has been reported (called here ΔC).

	Obs. Mean			Retinex			iCAM1			iCAM05			HDR		
	L	a	b	L	a	b	L	a	b	L	a	b	L	a	b
1:Red	28	-4,9	19	31	32	14	52	4	7	49	3	7	26	-28	27
2:Yellow	88,1	-33	70,5	93	-14	64	90	2	22	90	1	23	71	-63	63
3:White	87,3	-50	64,3	95	-2	6	94	6	7	95	5	7	73	-63	53
4:Yellow	15,6	-2	6,2	6	-2	3	22	1	3	26	2	3	3	-6	5
5:Yellow	91,5	-29	73,6	97	-15	69	99	-7	25	99	-7	26	88	-77	78
6:Cyan	69,3	-40	14,4	81	15	-27	82	9	-16	81	9	-19	58	-45	18
7:Cyan	74	-43	18,3	87	10	-18	78	10	-18	77	10	-20	56	-43	16
8:Yellow	89,3	-28	69,7	89	-19	70	83	1	26	82	0	28	63	-58	58
9:Green	67,4	-47	36,7	83	-6	4	84	5	6	82	5	6	52	-48	38
10:Red	23,9	-3,1	15,2	42	37	12	45	5	2	43	4	3	17	-19	17

Tab. 16: CIELAB coordinates of the points considered.

	ΔE^*_{ab} 1	ΔE^*_{ab} 2	ΔE^*_{ab} 3	ΔE^*_{ab} 4
1:Red	37,36	28,27	25,44	24,53
2:Yellow	20,49	59,72	58,33	35,51
3:White	75,78	80,26	79,66	22,5
4:Yellow	10,12	7,76	11,59	13,27
5:Yellow	15,29	53,67	52,77	48,82
6:Cyan	70,14	59,38	60,77	12,72
7:Cyan	65,14	63,95	65,05	18,15
8:Yellow	9,41	53,04	50,98	41,29
9:Green	54,34	62,21	61,71	15,53
10:Red	44,11	26,17	23,75	17,43

ΔC 1	ΔC2	ΔC3	ΔC4
37,24	14,94	14,37	24,45
19,89	59,69	58,30	31,12
75,39	79,98	79,29	17,38
3,20	4,39	5,12	4,18
14,26	53,14	52,23	48,70
69,16	58,00	59,63	5,84
63,83	63,83	64,99	2,35
9,40	52,67	50,45	31,83
52,05	59,96	59,96	1,98
40,23	15,49	14,12	16

Tab. 17: ΔE^*_{ab} and ΔC between the colours.

In table 17 the following ΔE^*_{ab} has been calculated:

- ΔE^*_{ab} 1: between the observers mean and RSR.
- ΔE^*_{ab} 2: between the observers mean and iCAM06 with D=1.
- ΔE^*_{ab} 3: between the observers mean and iCAM06 with D=0.5.
- ΔE^*_{ab} 4: between the observers mean and the HDR image.

These results can be better illustrated in figure 75, in which the graph of the ΔE^*_{ab} of six points has been reported. The six points are: red, yellow, white, yellow in the shadow, cyan, green respectively. It is possible to note how the ΔE^*_{ab} 4 (deltaE4 in the label), for the most part of the cases is less than the other, meaning the appearance perceived by the observer is closer to the HDR image than to the appearance calculated by the algorithms.

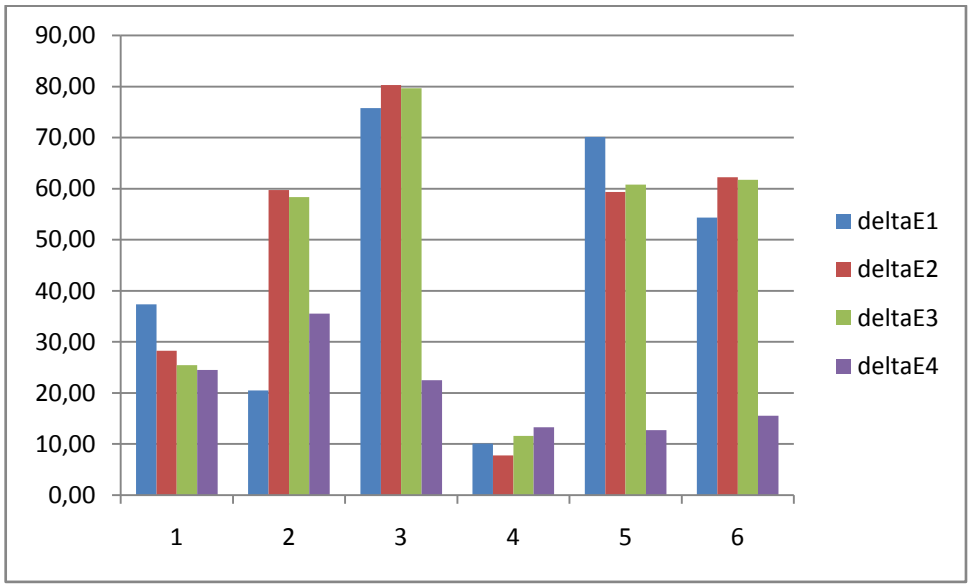


Fig. 75: graph of the ΔE^*_{ab} calculated for six points (red, yellow, white, yellow in the shadow, cyan, green).

In figure 76 a range of points are plotted in the CIELAB space. In the figure, L* is coming out the page, the circle has a ray equals to 100 and it is used as reference for the a* and b* values. The broken coloured line represents the gamut of the sRGB space. The white face (point 3) is of great importance, as it gives the colour cast. The face painted with white is perceived by the observers as greenish, and it is quite close to the real light reflected (the HDR image). RSR and iCAM06 compute the face as white, failing the prediction of the appearance (the observers result).

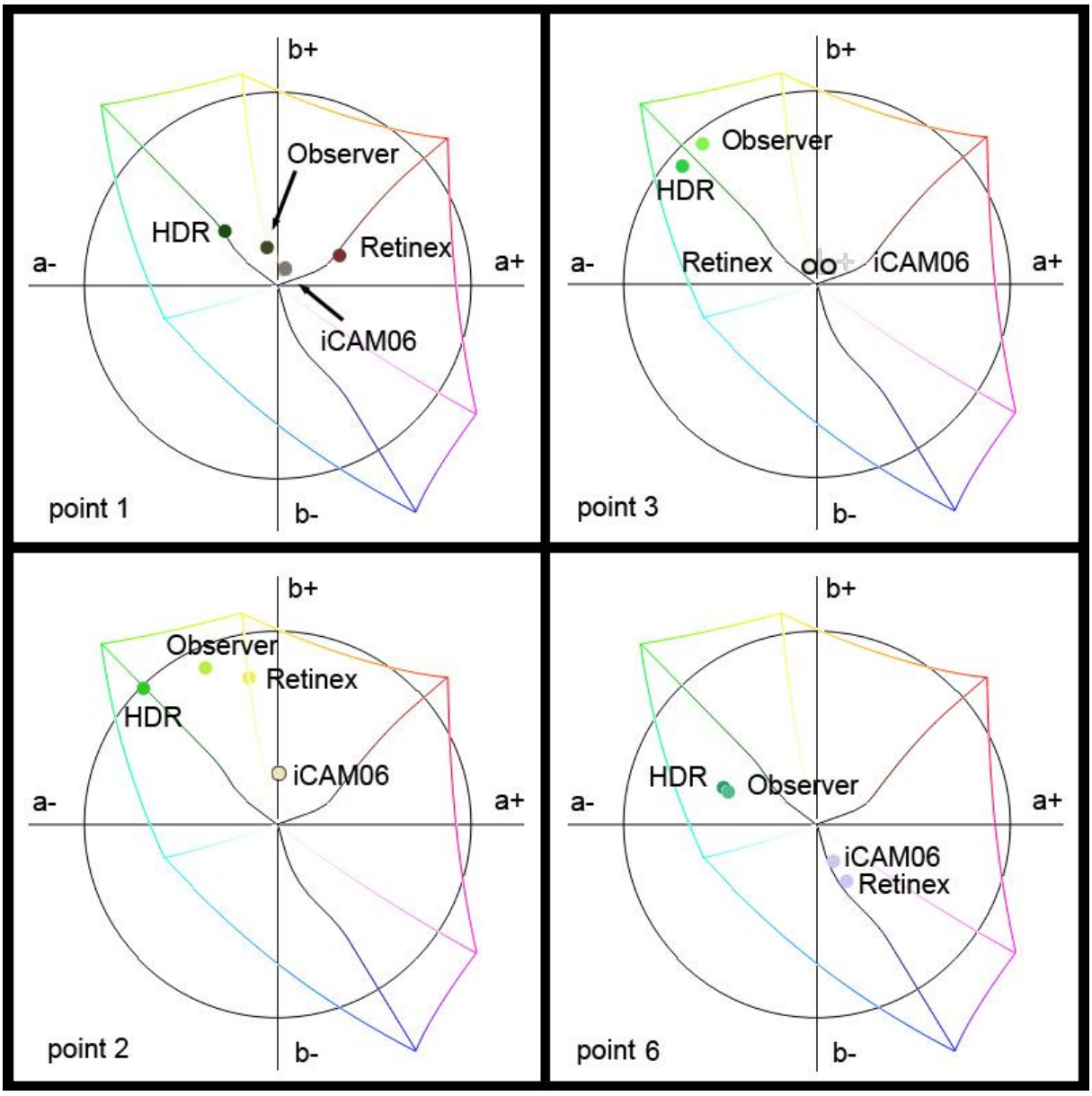


Fig. 76: Some points in the CIELAB space (point 1: red, point 2: yellow, point 3: white, point 6: cyan).

In this experiment it is shown that in extreme conditions, in which a strong colour cast is present, the human colour constancy is far from being perfect. The two algorithms considered instead calculate the appearance as almost perfect colour constancy, calculating, incorrectly, the white face as white, while it is perceived as greenish.

7.4 Third experiment

The scene set up for the previous experiment has been used for the third, replacing the green light with a typical tungsten light bulb. This kind of light introduces a reddish colour cast, visible in the pictures taken to create the HDR image (figure 77).

Fig. 77: The eleven pictures taken to create the HDR image.

Observing the frontal face of the cube on the right side, it is possible to see, especially in the less exposed images, how the light reflected has a reddish component.

The same modality and conditions of the previous experiment have been used to undertake the third experiment. In figure 78 a range of reproductions by the observers are illustrated. In contrast to the previous experiment the observers seem not to perceive the dominant colour introduced by the light. This fact can be ascertain observing that almost the persons paint the face of the cube on the right as white.

In figures 79-81 the filtering with RSR, iCAM06 (D=1) and iCAM06 (D=0.5) are shown. The same procedure of transforming the data in the CIELAB (white point: D50) space and compare the data using the ΔE^*_{ab}, has been adopted (tables 18-19).

The same consideration of the previous experiment about the iCAM06 filtering are valid in this one.

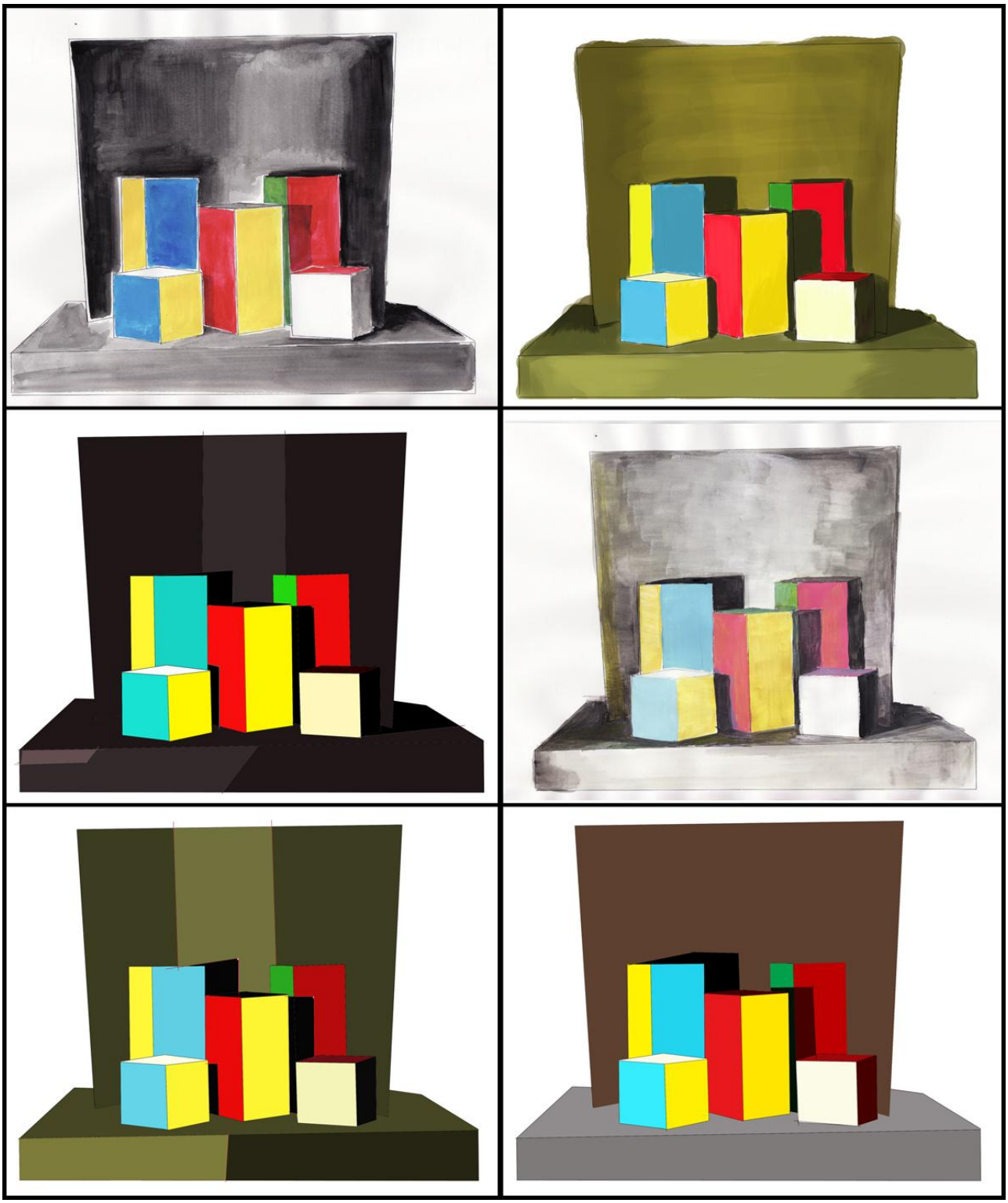


Fig. 78: A range of reproduction by the observers.

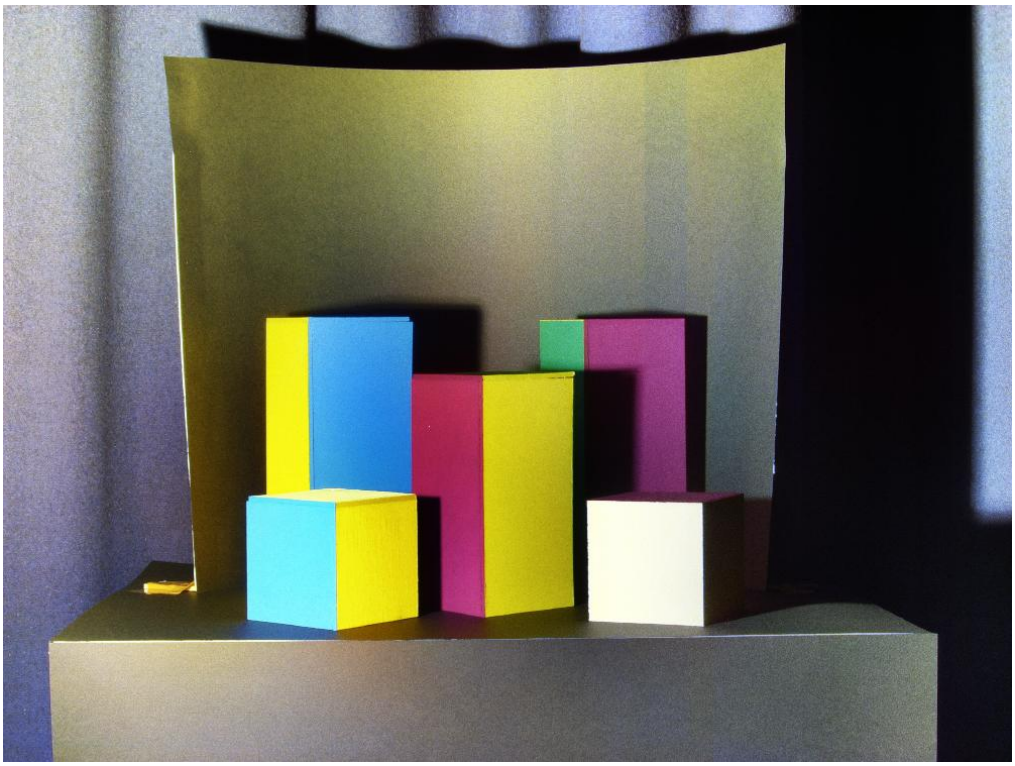


Fig. 79: RSR filtering.

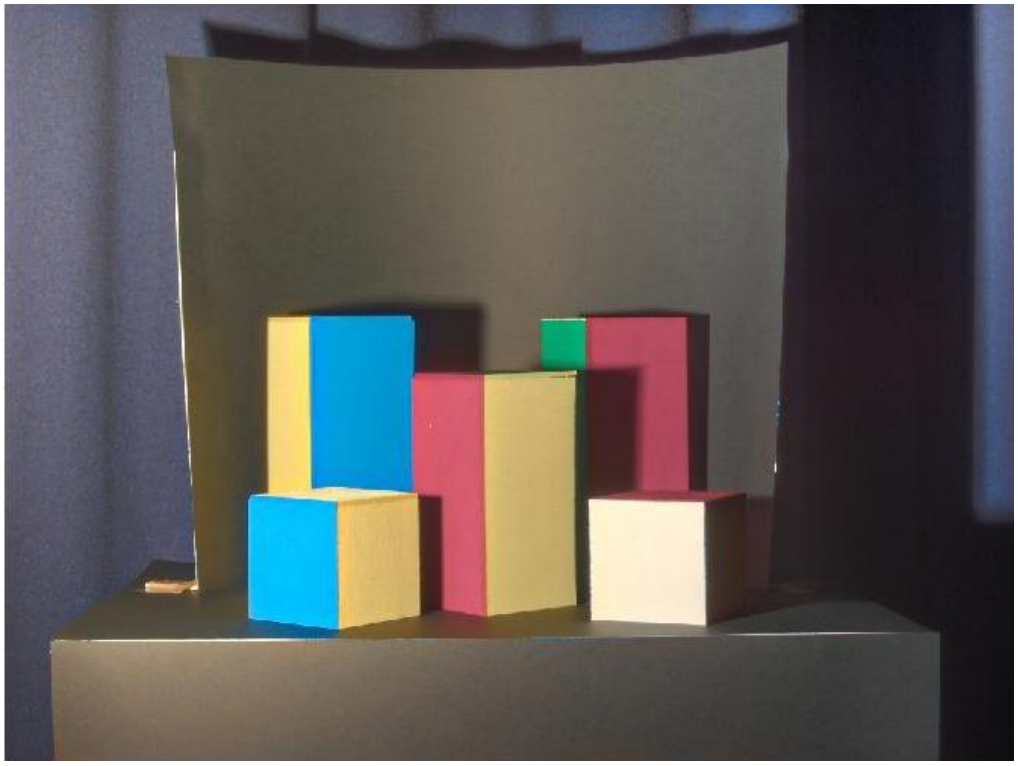


Fig. 80: iCAM06 filtering, with D=1.

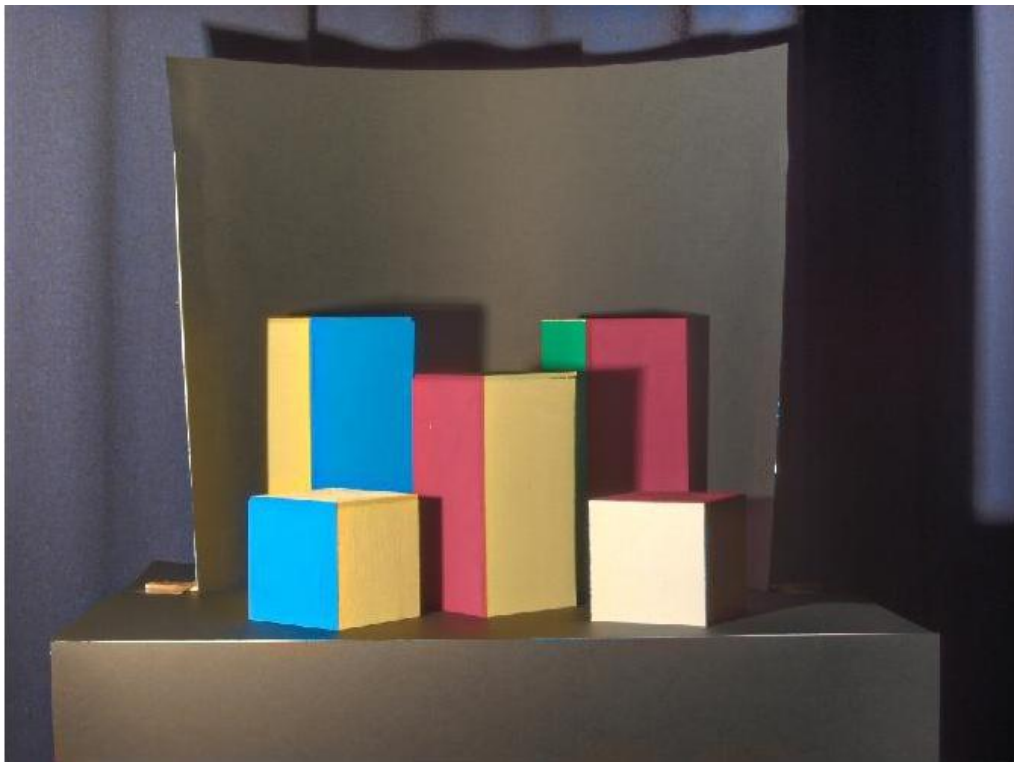


Fig. 81: iCAM06 filtering, with D=0.5.

In figure 82 are shown the points taken in account to perform the comparison.

Fig. 82: points considered in the scene.

	Obs. mean			Retinex			iCAM1			iCAM05			HDR		
	L	a	b	L	a	b	L	a	b	L	a	b	L	a	b
1:Red	52,5	68,7	47,2	42	56	11	56	52	14	54	52	16	31	45	33
2:Yellow	90,2	-8,5	78,8	82	-10	78	77	4	39	76	6	41	49	24	57
3:White	98,1	-2,6	11,9	96	-2	6	95	3	14	95	2	16	75	31	63
4:Yellow	21,8	7,8	5,6	8	4	5	21	5	5	22	5	3	4	7	6
5:Cyan	70,9	-22	-28	88	-35	-17	73	-26	-41	71	-23	-44	49	-11	-10
6:Yellow	92,4	-9,8	78,7	91	-12	87	80	4	45	79	7	48	59	29	65
7:Yellow	94,1	-11	78,6	94	-14	88	90	2	49	88	4	54	58	28	65
8:Cyan	64,5	-18	-31	64	-12	-46	60	-19	-42	59	-18	-41	30	-9	-6
9:Green	61,4	-48	47,4	66	-41	18	60	-48	17	58	-47	18	25	-16	27
10:Red	47,7	65,8	45,5	44	47	-16	47	-19	-42	46	47	16	18	34	22

Tab. 18: CIELAB data for 10 points, scene with tungsten light.

	ΔE^*_{ab} 1	ΔE^*_{ab} 2	ΔE^*_{ab} 3	ΔE^*_{ab} 4	ΔC 1	ΔC 2	ΔC3	ΔC4
1:Red	39,77	37,33	35,42	35,01	38,36	37,16	35,39	27,63
2:Yellow	8,37	43,76	42,9	56,82	1,70	41,72	40,49	39,13
3:White	6,29	6,74	6,9	65,37	5,93	5,98	6,16	61,16
4:Yellow	14,33	2,97	3,83	17,82	3,85	2,86	3,82	0,89
5:Yellow	24,22	13,57	15,83	30,53	17,16	13,41	15,83	21,27
6:Cyan	8,7	38,47	37,47	53	8,59	36,42	35,00	41,15
7:Cyan	9,93	32,51	29,35	54,71	9,93	32,25	28,71	41,11
8:Yellow	16,27	12,26	11,69	43,27	16,26	11,40	10,31	26,12
9:Green	30,59	30,43	29,62	52,65	30,25	30,40	29,42	38,03
10:Red	64,42	121,85	35,02	49,45	64,31	121,85	34,98	39,54

Tab. 19: ΔE^*_{ab} and ΔC between 10 points.

- ΔE^*_{ab} 1: between the observers mean and RSR.
- ΔE^*_{ab} 2: between the observers mean and iCAM06 with D=1.
- ΔE^*_{ab} 3: between the observers mean and iCAM06 with D=0.5.
- ΔE^*_{ab} 4: between the observers mean and the HDR image.

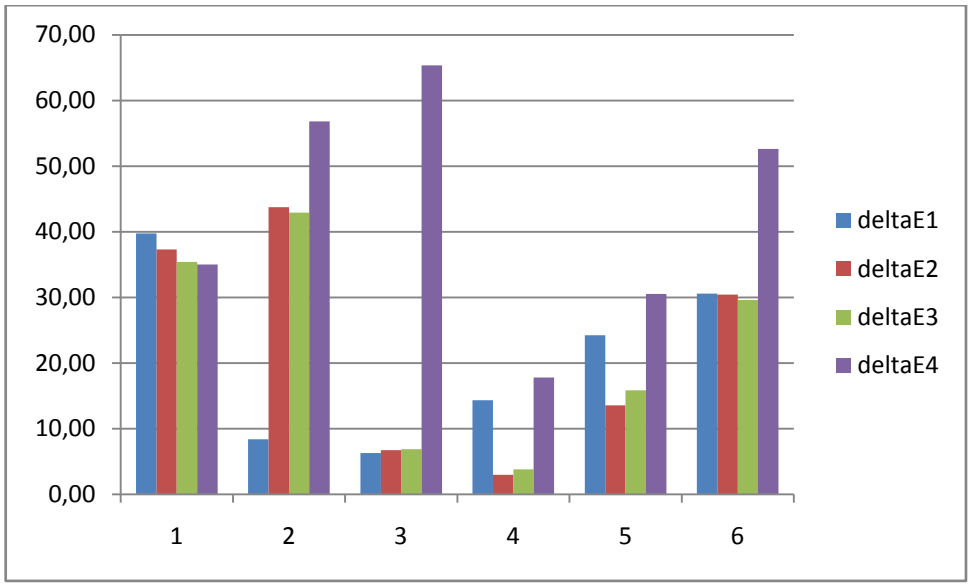


Fig. 83: graph reporting the ΔE^*_{ab} for six points (red, yellow, white, yellow in shadow, cyan, green).

These results can be better visualized in figure 83 in which the graph of the ΔE^*_{ab} of six points has been reported. The six points are: red (1), yellow (2), white (3), yellow in the shadow (4), cyan (5), green (6) respectively. We want to remark again the importance of the white face. It is possible to see from the graph (figure 83) and from figure 84, how the face of the right cube (point 3) is painted with white by the observers, and calculated as white by RSR and iCAM. In this experiment therefore colour constancy take place in the human observer and the predictions of the two algorithms agree with their results. It is clearly visible in fact from the HDR image, taken as description of the reality, that the light reflected from the white surface as a reddish component that is not perceived so strongly by the human visual system. RSR and iCAM are both able to simulate this behaviour.

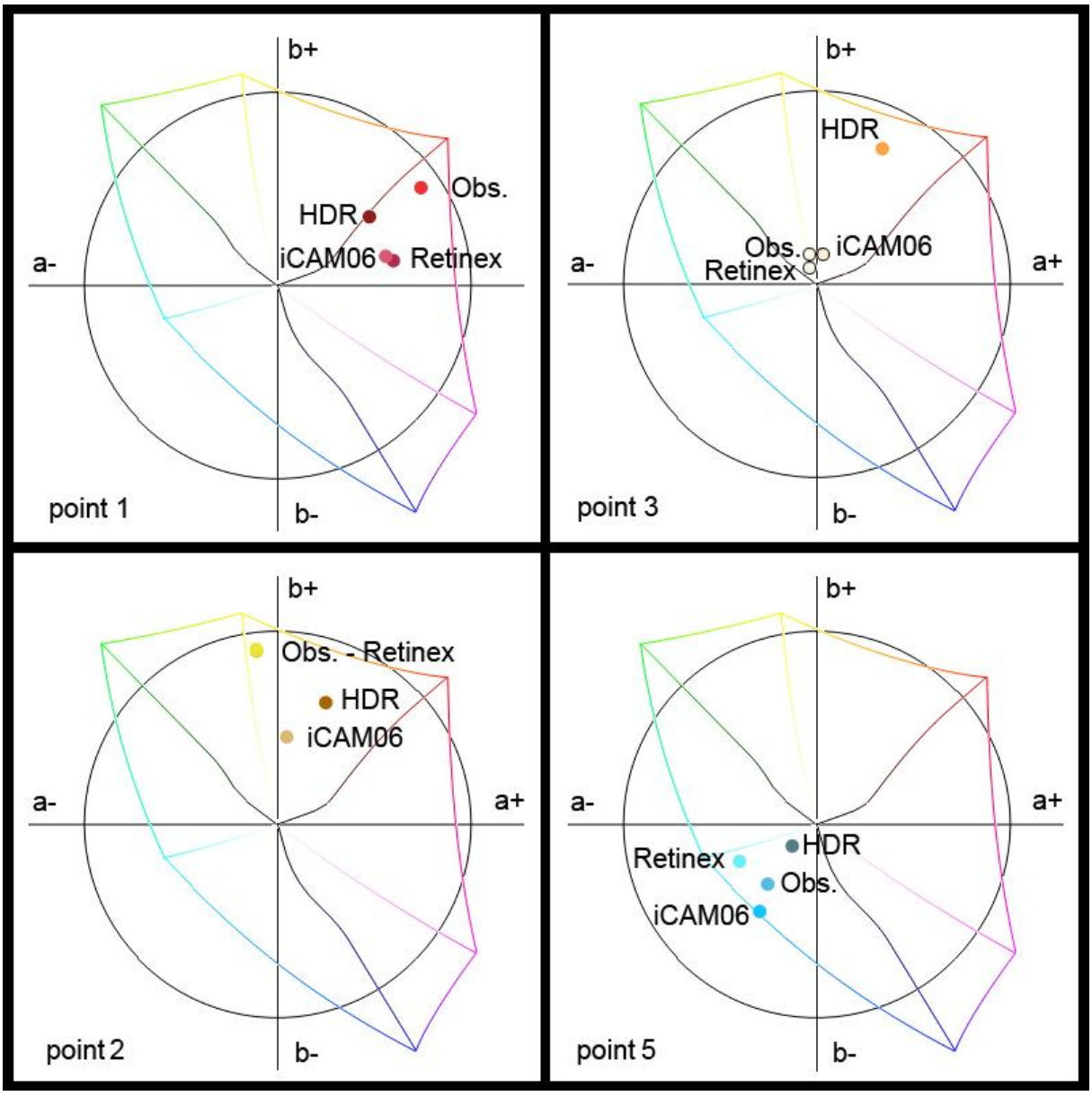


Fig. 84: Some points representing the appearance calculated by the observers, and the algorithms, and the HDR value.

7.5 Conclusions

The experiments presented in this chapter show that the colour appearance in related not only to the wavelength reflected by the objects, but also to factors that define the context. Through the colouring process (painted or digital) we attempt to understand how the eye-brain system perceived the scene.

In experiment one we simulated a situation that can be defined as 3D colour contrast, placing a white painted face surrounded by darker area, and a grey painted face surrounded by lighter area. Most of the people perceived the white face as lighter than the grey face, even if the latter reflected more light. From a sequence of shots with different exposure time an HDR

image as been created, used as input for a Retinex model and iCAM06 model. In both the cases the computation render the white face as lighter than the grey face, according to the experimental data, validating the sensation of the subjects.

In experiments two and three the chromatic constancy phenomenon has been considered. We demonstrated that in extreme conditions, in presence of a strong coloured light, this ability of the visual system is far from being perfect. The algorithms considered instead prove to be effective in simulating the phenomenon even in condition in which the human visual system is not able to do. In more typical condition, simulated in the experiment with a tungsten light, the colour constancy is verified both for the observers and for the algorithms.

8. A new printing pipeline

Walking through a city, our gaze is caught by many advertising billboards that cover walls, buildings, buses shelters, undergrounds (figure 85) These billboards are printed starting from a single file that the graphic designer send to the printer, without considering the different environments in which the billboard is going to be set and viewed. Moreover, the advertisings are presented on different supports and, characterized by different visual properties. Thus, a billboard seen under the sunlight in the summer, is different from a back-illuminated billboard seen during the night, or a billboard placed on the underground platform, with specific neon lights and different backgrounds.

Fig. 85: advertising billboards in the city.

A project has been developed for an Italian advertising agency in order to study the colour and contrast reproduction of the billboard in indoor and outdoor environments, with different lighting conditions, settings and supports.

8.1 The proposed method

As already mentioned human vision has an interesting property, it is able to perceive a scene in an almost steady way regardless of the change in illumination. Thus the colours of a printed image viewed inside an office or outside, in a bright summer afternoon, appear similar, despite their colorimetric difference. In our work we try to simulate how an image viewed under different kind of illumination, in a certain context is perceived by our visual system.
The procedure that is described consists of two phases. First it is necessary to learn the characteristics of the inks and paper. For this aim some patches of particular combinations of

colours are printed and measured with a spectrophotometer. In this way it is possible to establish the relationship between the digital values of the file and the actual inks printed on the paper. We need to perform this step only once, for a specific couple ink/paper. For this project three kinds of support with the respective inks used are considered: sticky paper, PVC canvas, blueback paper (figure 86).

Fig. 86: Some of the patches used in the learning phase. From left to right: blueback paper, sticky paper, PVC canvas

In the second phase an estimation of the final appearance of the advertising is computed. The starting point is the evaluation of the spectral composition of each point of the image. From a CMYK image, knowing the characteristics of the inks and of the paper that will be used to print the picture, Neugebauer (Neugebauer, 1937) and Yule-Nielsen (Yule, et al., 1951) models are used to estimate the spectral image. The next step consists adding an illuminant to the image, in order to simulate different illumination conditions, both artificial or natural. Starting from this spectral composition the illuminant component is superimposed. No geometrical lighting differences are considered so far; this point is one of the future direction of research. After that the spectral image is converted in RGB. The last operation simulates and display the visual appearance of the image in the context, that is how the colour appearance of the image will change depending on the type of illumination and visual context in which the image is placed. This step is performed using image enhancement models based on how the human visual system works. Figure 87 represents graphically the pipeline.

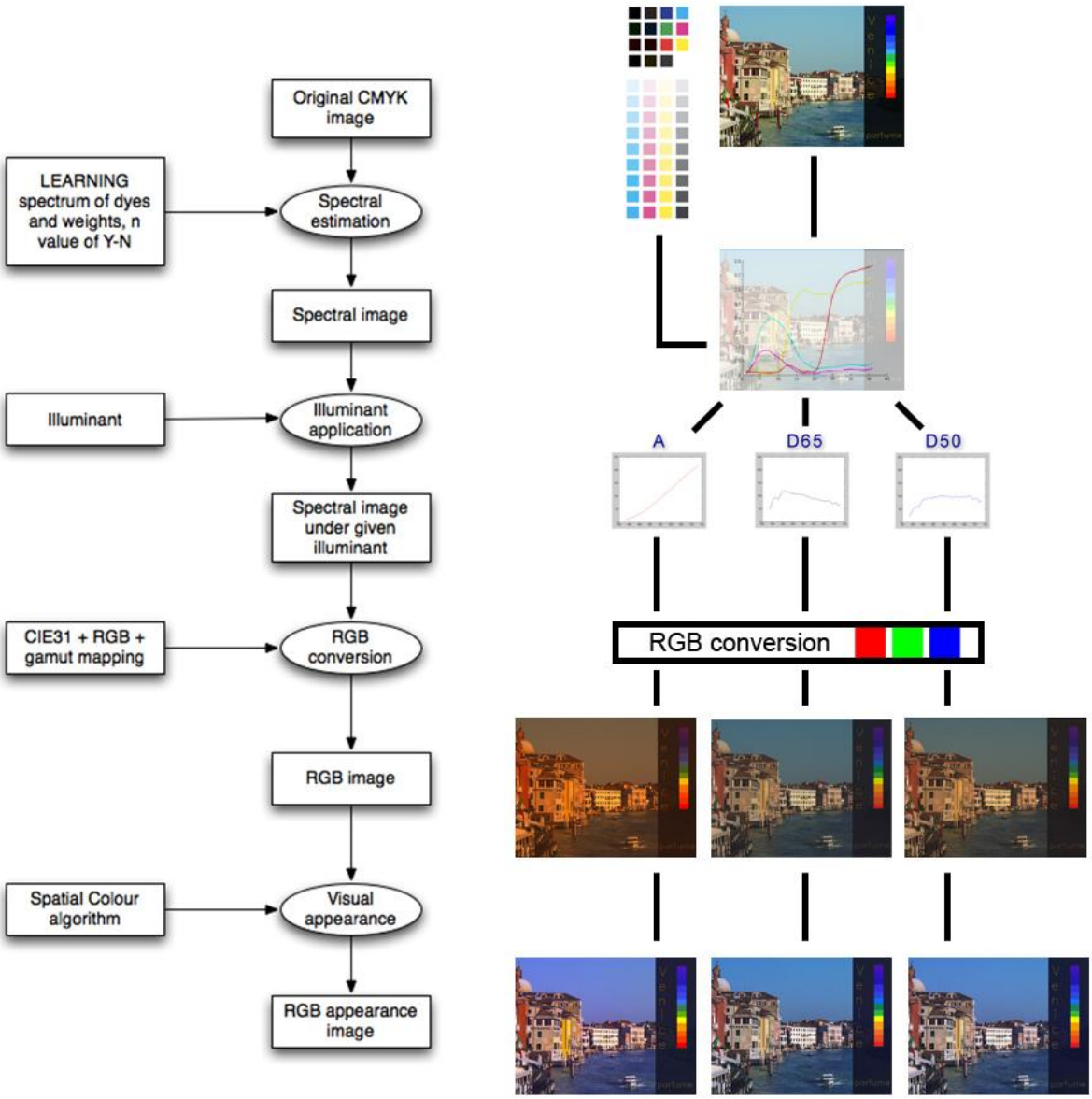


Fig. 87: graphic pipeline

8.2 The learning phase

In order to convert a CMYK image in a spectral image, it is necessary to estimate the relationship between the digital values and the ink percentage placed on the paper by the hardcopy device. This estimation can be done through the Neugebauer (Neugebauer, 1937) model that allows to predict the spectrum of every colour, having measured the spectra of some combinations of basic colours. In the case of a CMYK colour printer, the samples to measure consist in the combinations of the four inks: cyan (C), magenta (M), yellow (Y), black (K) and of the paper (W). The samples are generated combining these colorants in

amount of 0% or 100% (figure 88 ,on the left). The reflectance spectra of these patches are measured with a spectrophotometer for wavelength λ in the range 380 - 730 nm, with a step of 10 nm. In order to predict the spectral reflectance of the colour is used the following formula, a linear combination of the 16 colours, for which the spectra has been previously measured.

$$R(\lambda)_{cmyk} = \sum_{i=1}^{16} w_i * R_i(\lambda) \qquad (8.1)$$

$R(\lambda)_{cmyk}$ indicates the value of a certain λ of the colour spectrum to calculate, $R_i(\lambda)$ is the value of the spectrum, at the same wavelength, of the i-th measured sample.
The sixteen basic colours for the linear combination are:

$\{R_W(\lambda), R_C(\lambda), R_M(\lambda), R_Y(\lambda), R_K(\lambda), R_{CM}(\lambda), R_{CY}(\lambda), R_{CK}(\lambda), R_{MY}(\lambda), R_{MK}(\lambda), R_{YK}(\lambda), R_{CMY}(\lambda), R_{CMK}(\lambda), R_{CYK}(\lambda), R_{MYK}(\lambda), R_{CMYK}(\lambda)\}$

where $R_1(\lambda)$ indicates $R_W(\lambda)$, that is the white sample, $R_2(\lambda)$ indicates $R_C(\lambda)$, the sample with 100% of cyan ink and so on up to $R_{16}(\lambda)$ that indicates $R_{CMYK}(\lambda)$ the sample given by the superimposition of all the inks.
The w_i indicate the weights that link the digital values C, M, Y, K with the dot areas c, m, y, k printed. In fact, for constructive reasons, C, M, Y, K and c, m, y, k usually are not linearly related.
w_i belong at the following set: $\{w_1 = (1-c)*(1-m)*(1-y)*(1-k), w_2 = c*(1-m)*(1-y)*(1-k), ..., w_{16} = c*m*y*k\}$
Where i.e. $w_2=c*(1-m)*(1-y)*(1-k)$ indicates that it will be printed only cyan ink, but not magenta, yellow and black.

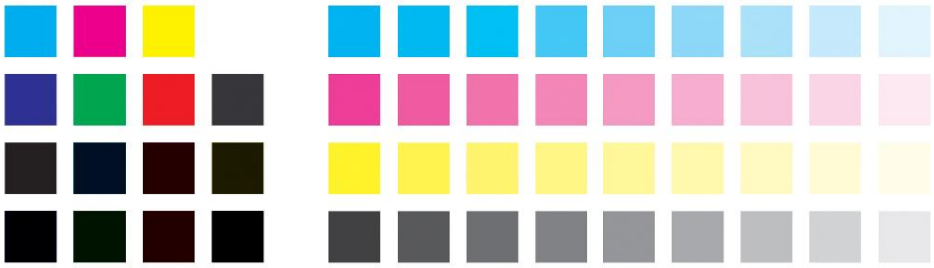


Fig. 88: Patches corresponding to the basic colours (on the left) and patches necessary to calculate the weights, representing the percentage of the basic inks from 10% to 90%.

The dot areas values can be empirically estimated printing and measuring some patches of the primary colours, at different percentages, for example C=0, C=10, ...,C=90, C=100 (figure 88, on the right). The Neugebauer model is then used to predict the colours minimizing the ΔE^*_{ab} between the predicted and the measured colour.

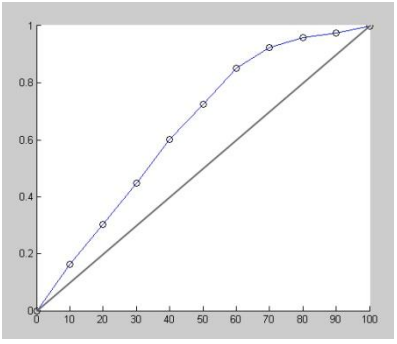


Fig. 89: Relationship between digital values and dot areas of the blue ink, for the Sticky support.

In this way, it is possible to obtain for each of the four colorants a relationship like that shown in figure 89. The model just described is quite limited, and an important improvement is given by the introduction of the Yule-Nielsen n factor (Yule, et al., 1951), that modify the equation in the following way:

$$R(\lambda)_{cmyk}^{1/n} = \sum_{i=1}^{16} w_i * R_i(\lambda)^{1/n} \qquad (8.2)$$

The aim of this modification is to keep in account the scattering of the light in the paper substrate. Introducing this non-linear factor clearly enhances the prediction of the spectrum.
To calculate this value, n is fixed, and the optimal dot areas are calculated, as previously described. After that the root mean square (RMS) between the predicted and the measured

spectra is calculated and a mean is estimated. This procedure is repeated for a range of values of n, for example from 1 to 30, and the n that minimize the mean RMS is chosen (an example of how the RMS varies for different n, for the sticky paper support is shown in figure 90).

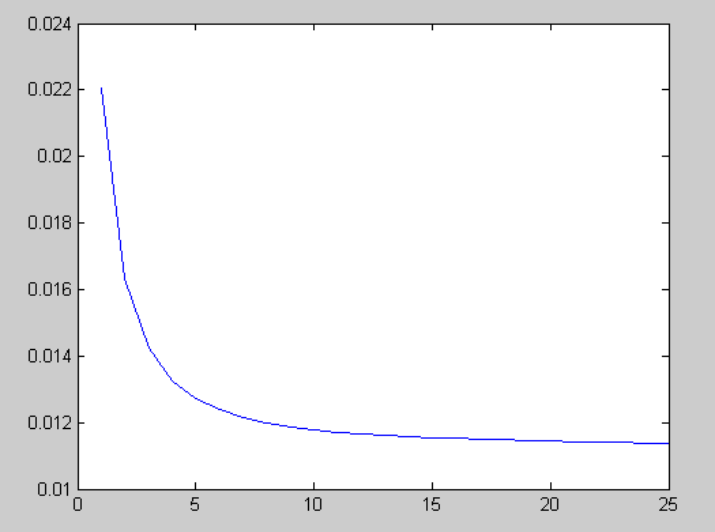


Fig. 90: RMS at varying n.

Once implemented, it is possible to validate the model, printing and measuring with a spectrophotometer a range of test patches. The spectra of these patches is then predicted with the model, and finally the ΔE^*_{ab} between measured and predicted colours is calculated.

Testing the model with around 300 patches (figure 91) for the three supports leads to an average ΔE^*_{ab} between 2.787 and 3.045 and maximum ΔE^*_{ab} between 6.1951 and 7.849 (table 20).

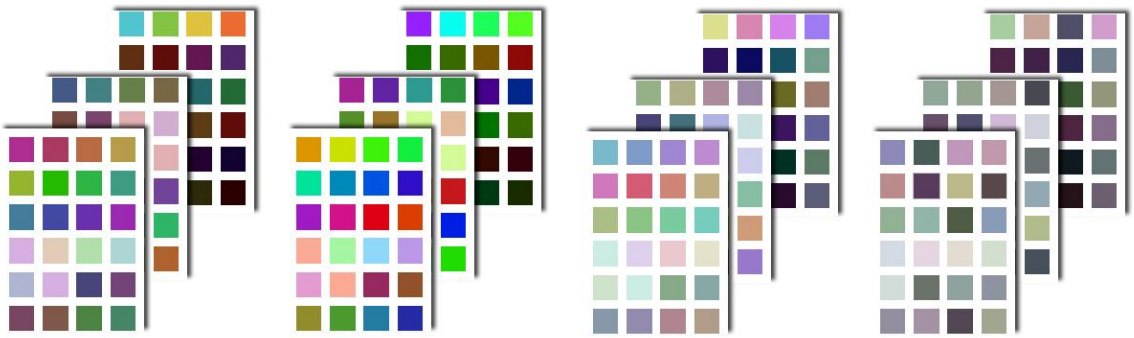


Fig. 91: test patches used to validate the model.

Support	average ΔE^*_{ab}	Max ΔE^*_{ab}
Sticky paper	2.7870	6.1951
PVC Canvas	3.0449	7.8495
Blueback paper	2.9835	7.4050

Tab. 20: average and max DE for three different supports .

8.3 Spectral image

In this phase, the algorithm receives as input a CMYK image and, using the data from the learning phase, transforms it in a spectral image, keeping in account the type of hardcopy device (in figure 92 are shown the spectra of four colours present in the image, calculated with this model). For every pixel of the image is estimated a spectral version, applying the Yule-Nielsen modified Neugebauer model. In this way the four C, M, Y, K channels of the pixel became a vector of 36 components, as our spectrophotometer can achieve measures in the range 380 - 730 nm, with a step of 10 nm (figure 93).

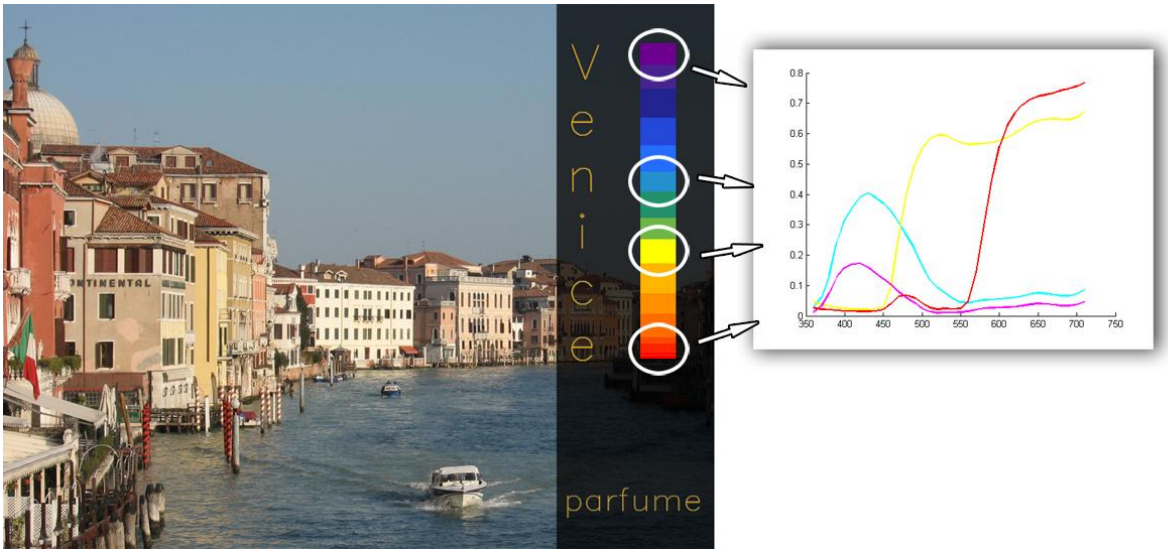


Fig. 92: Four spectra calculated from the CMYK image

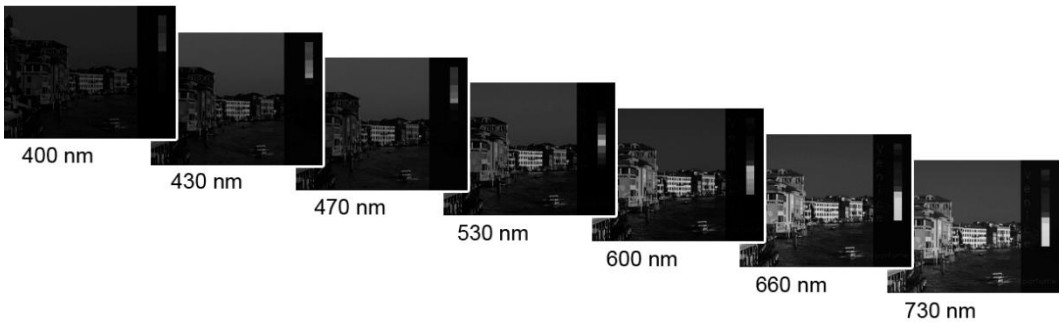


Fig. 93: A range of spectral channels calculated from the CMYK image.

8.4 Adding the illuminant

The spectral image, as calculated in the previous step, is multiplied for a certain illuminant. This operation is performed in the next step. Every pixel of the image, that now consists in a 36-elements vector, is multiplied, component per component, with the spectra of the chosen illuminant. Our program has pre-programmed well-known illuminants, as D65, D50, A, F2, and eventually allows to load a custom illuminant from file.

The result of this process is a spectral image as uniformly illuminated by a certain source of light. No geometrical lighting non-uniformities are considered so far. Some preliminary test with the geometry of the backlit boxes has been done, but this point will be investigated in the future.

8.5 RGB transform

To display the images, it is necessary to transform the spectral image in a RGB image. To this aim, we need to pass through the XYZ tristimulus values. In order to perform this operation we use the CIE 1931 Colour Matching Functions.

The following formulae are used for the conversion:

$$N=\sum_i cmf(y)_i * I_i \tag{8.3}$$

$$X=(1/N) * \sum_i cmf(x)_i * S_i * I_i \tag{8.4}$$

$$Y=(1/N) * \sum_i cmf(y)_i * S_i * I_i \tag{8.5}$$

$$Z=(1/N) * \sum_i cmf(z)_i * S_i * I_i \tag{8.6}$$

Where N is a normalization factor, cmf indicate the CIE 1931 Colour Matching Functions.

I is the spectral power distribution of the illuminant and S is the spectral reflectance of the sample. Repeating this process for every pixel of the image it is possible to obtain a device independent XYZ image.

At this point, according to the device involved, the correct RGB space has to be chosen. In our preliminary test, in order to be as general as possible, we used sRGB. Every RGB space has a own white point, for example D65 for the sRGB, thus it should be necessary to perform a chromatic adaptation transform from the reference white of the input XYZ to the final XYZ, before the RGB conversion. However this step is not accomplished, because the

aim is to observe the image in an “objective” way. It will be the last step that simulate the visual appearance and thus the white adaptation.

8.6 Appearance model

The last step has the aim to estimate how the human visual system will perceive the visual information of the image considering its position, illumination and context. In order to make this possible a spatial colour algorithm is applied. In this kind of algorithms every pixel of the image is recomputed according to the spatial distribution of the other pixel.

Starting from the Land's Retinex theory in the last years some models has been implemented that modify colour information in the visual contest, for example Retinex, ACE, RSR, Race. They try to emulate some mechanisms of the human visual system in order to estimate the final visual appearance. In figure 94 a screenshot of the application developed is shown.

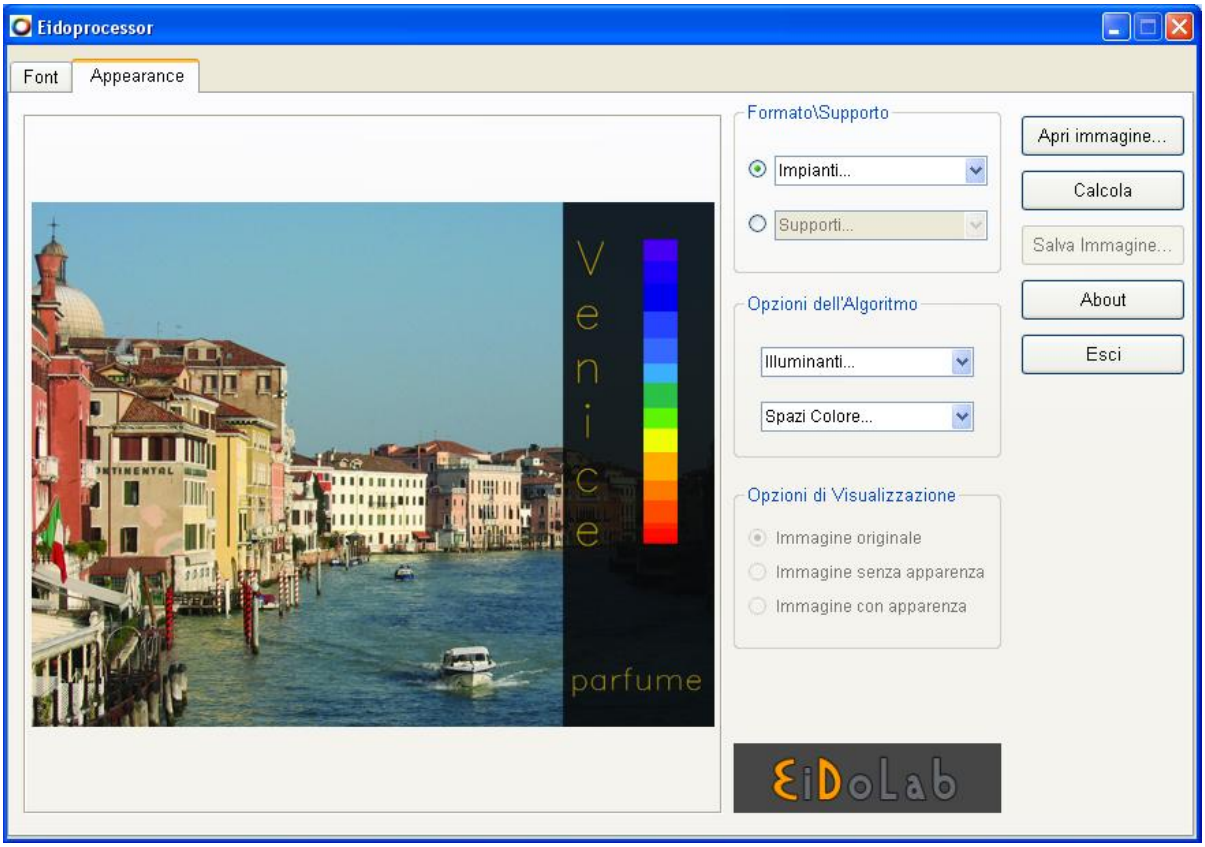


Fig. 94: screenshot of the application.

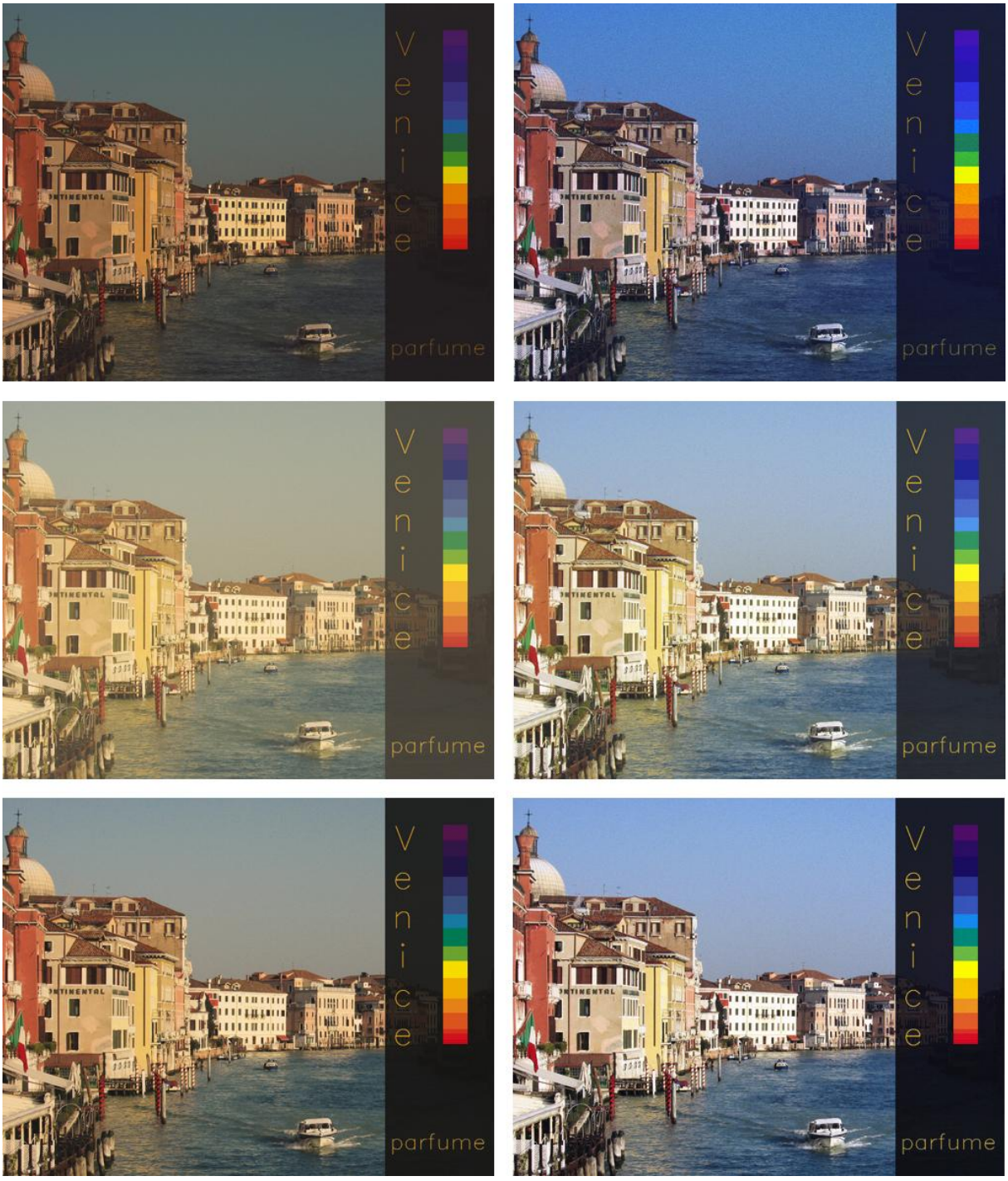


Fig. 95: In the left column the same image printed on different support (sticky paper, blueback paper, PVC canvas) and seen under the same illuminant (D75). At the right side the corresponding image after the application of the appearance step.

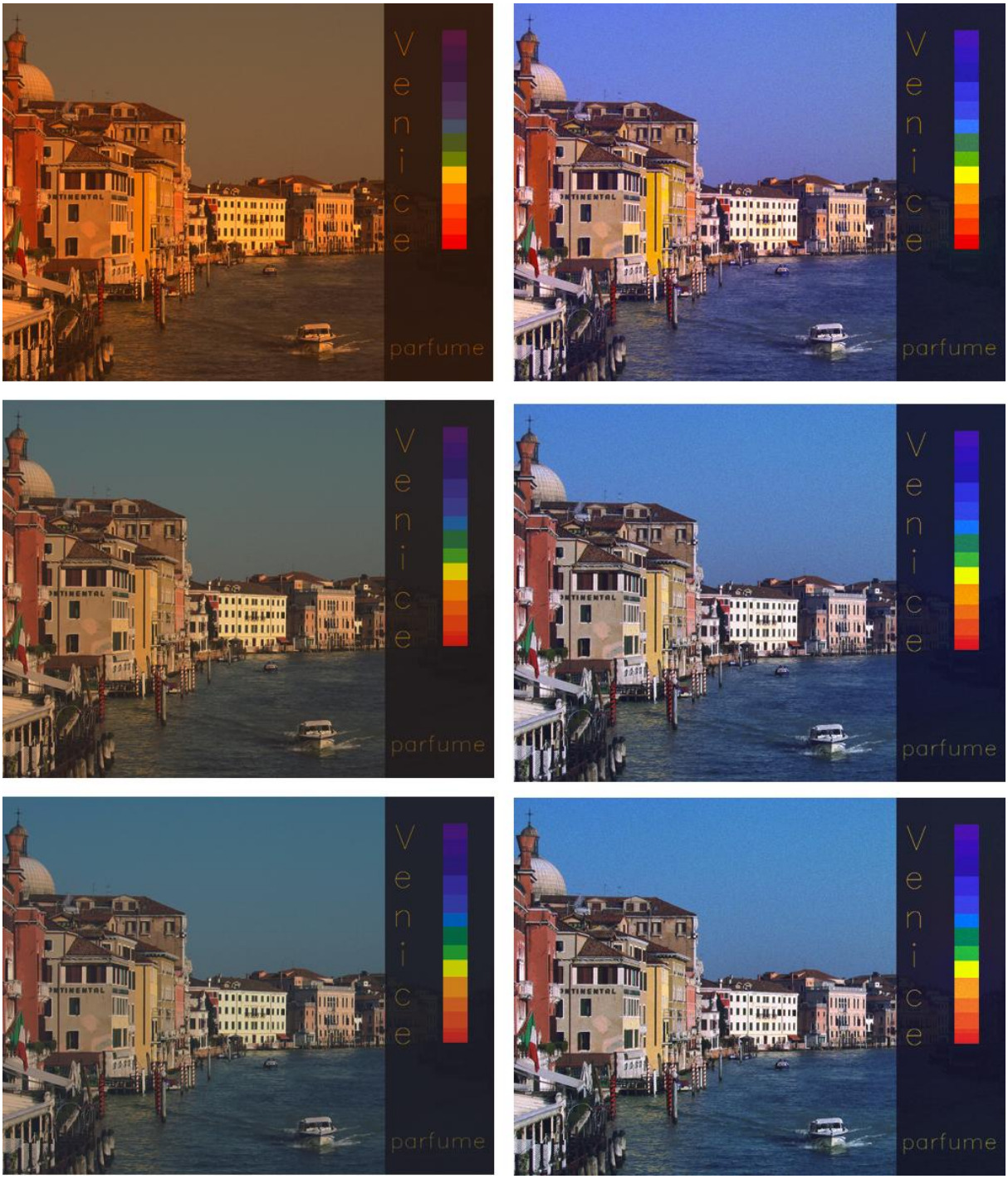


Fig. 96: Left: the image printed on the same support (sticky paper) but with different illuminant (A, D50, D75). At right: the corresponding image after the appearance step.

In figures 95-96 it is possible to see the result of the application of the appearance model in different situations. When simulating different supports or different lights, the images are colorimetrically different. When applying the algorithm that simulates the appearance, these difference are considerably reduced (due to the colour constancy, see also chapter 7).

8.7 Conclusions and future possibility

We presented a method to process a CMYK image in order to simulate how it appears to a human observer in a certain context, under a given illuminant. The methods produce a useful estimate that can be used to check for problems of visualization, according to the various type of advertising. Billboards in fact can be back-illuminated or with frontal illumination, both natural or artificial. All these combinations can be taken into account.

Further accuracy can be add to the method. The Neugebauer model, that transforms the CMYK image in a spectral image, can be improved by using a cellular model, reducing the space over which the interpolation is done adding more primaries (Rolleston, et al., 1993). Furthermore Zuffi and her colleagues improved the model keeping in account how the inks interact when superimposed on the paper (Zuffi, et al., 2006).

The second improvement could be in the simulation of the illumination. At the moment the billboard is uniformly illuminated, but since the billboard is seen in an outside environment, a more complex geometry of illumination is needed.

9. Conclusion and future work

The colour is an attribute of our visual perception. The basic colorimetry is no more able to explain certain visual phenomena such as colour constancy or simultaneous contrast. Some colour appearance models are needed to perform this task. In this thesis we compare different models with human evaluations of the appearance.

In chapter 2 is given an overview of the human visual system and it describes a range of colour appearance phenomena. Chapter 3 describes the spatial colour models, based on the Land's Retinex theory, that try to simulate the behaviour of the human visual system. In chapter 4 are presented the models proposed by the CIE, based on equations that attempt to quantify some perceptual attribute such as Lightness, Brightness, Chroma, Colourfulness, and iCAM06, a model derived from the CIECAM02. In chapter 5 we present background on HDR and raw processing, while in chapter 6 a series of experiment on High Dynamic Range (HDR) images are perfomed, in order to study the procedure that leads to the creation of an HDR image. HDR image can increase the dynamic range of a single image, and with the tone mapping operator better results can be obtained, resulting in images able to show more details in the dark or light regions. Anyway we showed that HDR image cannot be considered as a perfect copy of a real scene as it is not able to record correctly the luminance of the scene. Different software produce different results, and moreover the glare can affect in the lens can affect the luminance that is captured. In Chapter 7 we performed a range of specific and targeted experiments in order to study the behaviours of the humans and of the algorithms in presence of specific colour appearance phenomena. Both iCAM06 and Random Spray Retinex are able to simulate the simultaneous contrast starting from an HDR image of a three-dimensional scene. The same phenomenon takes place in the human observers. Performing a second experiment, it has been shown that, in presence of strong coloured dominant, colour constancy seems not to take place in humans. The algorithms however simulate, incorrectly, the phenomenon. On the other end, in typical situations in which a slight colour cast is introduced the human visual system is able to perceive the colour as natural, and both the algorithms work according to this fact. The models act in the right directions, but other factors have to be taken in account to simulate any complex aspect of the human visual system. Finally a new printing pipeline has been shown in chapter 8. A graphics designer can take

advantage in knowing in advance the appearance of an image, printed on a specific support and with specific kind of inks. As a billboard is placed in a certain environment, this pipeline could be extended in order to keeping in account the context in which the image is seen. This pipeline can be used in a real production workflow as the image processing is performed with the new CUDA Nvdia parallel technology, reducing noticeably, compared with the CPU, the computational time of the spatial algorithm (Appendix B). Considering the calculation power that arise from the CUDA technology new algorithms can be developed.

A big challenge can be the fulfilment of a physically-based render engine. Nowadays render engines work with RGB value, an improvement can derive from using a spectral approach. This could led to a more photorealistic image, avoiding problems like metamerism. In this case the billboard could be placed in a 3D environment for a more photorealistic visualization. Before displaying the image on the screen a colour appearance model can be used in order to enhancement the general effect.

Appendix A: Colour-appearance terminology

Since the aim of the this kind of CAM is to describe the colour appearance in a mathematical way, some specific terms are necessary. It is important to note that these definitions concern perceptual terms, and are not colorimetric quantities. In the following are quoted some definitions taken from the International (CIE, 1987).

Colour. Attribute of visual perception consisting of any combination of chromatic and achromatic content. This attribute can be described by chromatic colour names such as yellow, orange, brown, red, pink, green, blue, purple, etc., or by achromatic colour names such as white, gray, black, etc., and qualified by bright, dim, light, dark, etc., or by combinations of such names.

Surprisingly enough the definition of colour contains the world colour, therefore the authors added a note:

Perceived colour depends on the spectral distribution of the colour stimulus, on the size, shape, structure, and surround of the stimulus area, on the state of adaptation of the observer's visual system, and on the observer's experience of the prevailing and similar situations of observations.

Before listing the colour attributes it is important to remark that some of them are typical of only related colours, that are colours viewed in relation to other colour stimuli. The definition is provided here:

Unrelated Colour. Colour perceived to belong to an area or object seen in isolation from other colours.

Related Colour. Colour perceived to belong to an area or object seen in relation to other colours.

Related colours are viewed in relation to other colour stimuli, unrelated colours are viewed completely in isolation.

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

Achromatic Colour. Perceived colour devoid of hue.

Chromatic Colour. Perceived colour possessing a hue.

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

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

Note.
Only related colours exhibit lightness.

Lightness express the brightness of a colour in a relative scale. Therefore the brightness refers to absolute level of the perception, and lightness is a sort of relative brightness, normalized typically from 0 (black) to 100 (white). A white paper seen in two different environments, for example inside an office or outside in sunny day, could have very different brightness but approximately the same lightness (probably it is the lightest stimulus, and therefore it appears white).

Colourfulness
Attribute of a visual sensation according to which the perceived colour of an area appears to be more or less chromatic.

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

Practically colourfulness indicates the amount of hue of a colour. In general colourfulness increases if the luminance level increase, since it is an absolute quantity.
Chroma is the colourfulness of a colour judged as a proportion of the brightness of the white.
It is used to indicate how much a colour move away from a grey with the same lightness.

Saturation. Colourfulness of an area judged in proportion to its brightness.

Both saturation and chroma can be thought as relative colourfulness. The difference is that, while saturation is the colourfulness of a stimulus relative to is own brightness, chroma is colourfulness relative to a white area brightness.
As these definitions could be confusing, it can be useful to think in terms of equation, not in a mathematical sense but as a description of the relationships between the colour attributes.

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

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

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

From these equations it is possible to find another definition for the saturation attribute.

$$Saturation = \frac{Colorfulness}{Brightness} = \frac{Chroma * Brightness(white)}{Lightness * Brightness(white)} = \frac{Chroma}{Lightness}$$

Note that from these equation it turns out that saturation is a redundant attribute, knowing the other five.

Appendix B: Working with CUDA™

An important aspect concerning this studies was the utilization of the new NVidia™ graphic board. By the use of the NVidia™ CUDA™ technology it is possible to solve high cost computation algorithms in a fraction of the time, thanks to the GPU parallel computing architecture.

Image processing algorithms often can take a lot of advantage from parallelization since the same operations are performed on every pixel of the image. Having many microprocessor it is possible to elaborate more pixels at the same time saving a huge amount of time.

To show the potentiality of this new technology an example is shown.

Two versions of Random Spray Retinex algorithm has been implemented, one based the simple CPU, and the second based on CUDA™ technology.

More versions of a square image, scaled from 128 pixels to 1024, has been used to test the two implementations. The parameters of the algorithms were 200 points per spray, and 20 repetitions. The output images are of course the same for the two versions, but the computational time is really different. The following graphics (figure 97) show the results, on the x-axis is represented the dimension, n pixel, of the image an on the y-axis the computation time in seconds.

Chosen the 1024x1024 image the time to compute the algorithm in CPU was about 15 minutes, in GPU about 12 seconds.

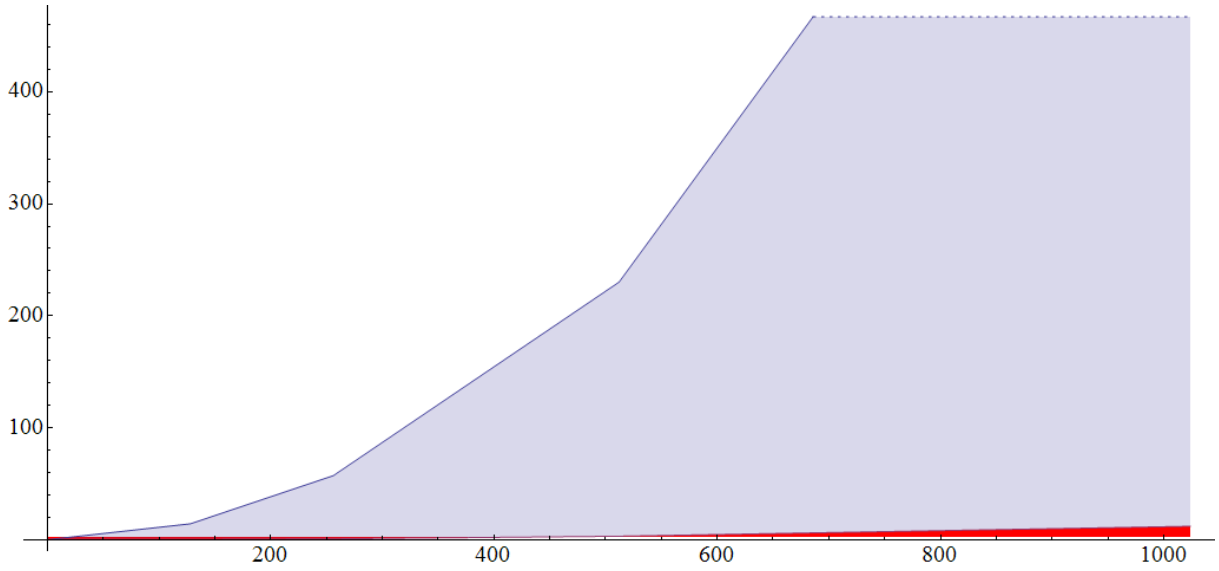


Fig. 97: in grey is represented the time request to run RSR in CPU. In red the same calculation performed on the Nvidia Graphic board.

Bibliography

Adelson, E. H., 1995. [Online 05 11 2010.] http://web.mit.edu/persci/people/adelson/checkershadow_illusion.html.

Adobe, *Digital Negative (DNG) specification*, v 1.0.0.0. 2004.

Albers, J., *Interaction of Color*, Yale University Press, 1963.

Bonanomi, C.; Villa, D.; Rizzi, A.; Marini, D., *Towards the simulation of visual appearance for indoor and outdoor advertising*, 11th Congress of the International Colour Association (AIC), Sidney, 2009.

Boynton, R. M., *Human Color Vision*, Holt, Rinehart and Winston, 1979.

Breneman, E.J., *Corresponding chromaticities for different states of adaptation to complex visual fields*, J. Opt. Soc. Am. A 4, 1987.

Brown, P. K. and Wald, G., *Visual pigments in single rods and cones of human retina*, Science 144, 1964.

Chambah, M., *Reference-free Image Quality Evaluation for Digital Film Restoration*, Society of Dyers and Colourists, SDC Journal : Colour Design & Creativity, 2008.

Chevreul, M. E., *The principles of harmony and contrast of colours, and their applications to the arts*, Henry G. Bohn, 1854.

CIE.,*International Lighting Vocabulary*, Vienna, CIE Publ. No. 17.4, 1987.

CIE., *The CIE 1997 Interim Colour Appearance Model* (Simple Version), *CIECAM97s.*,CIE Pub. 131, 1998.

CIE TC1-34, *Testing Colour Appearance Models*, Progress Report. 1997.

Damasio, A.; Yamada, T.; Damasio, H.; Corbett, J.; McKee, J., *Central achromatopsia: behavioral, anatomic and physiologic aspects*, Neurology 30, 1980.

Daw, N. W., *Why After-Images are Not Seen in Normal Circumstances*, Nature 196:1143–1145, 1962.

Debevec, P. and Malik, J., *Recovering High Dynamic Range Radiance Maps from Photographs*, 1997.

Durand, F. and Dorsey, J., *Fast bilateral filtering for the display of high-dynamic-range images*, Proceedings of SIGGRAPH 2002, 2002.

Ebner, F. and Fairchild, M. D., *Development and testing of a color space (IPT) with improved hue uniformity*, IS&T/SID 6th Color Imaging Conference, 1998.

Fairchild, M. D., *Refinement of the RLAB color space*, Color Res. Appl. 21, 1996.

Fairchild, M. D. and Johnson, G.M., *Meet iCAM: a next-generation color appearance model*, IS&T/SID 10th Color Imaging Conference, 2002.

Fairchild, M. D. and Johnson, G.M., *The iCAM framework for image appearance, image differences, and image quality*, J. of Electronic Imaging 13, 2004.

Fairchild, M. D., *Formulation and testing of an incomplete-chromatic-adaptation model*, Color Res. Appl. 16, 1991.

Fairchild, M. D., *Color Appearance Models*, 2nd Ed., Wiley-IS&T, 2005.

Funt, B., Ciurea, F. and McCann, J. ,*Retinex in Matlab*, Proc. 8th Color Imaging Conference, 2000.

Gadia, D.; Villa, D.; Bonanomi, C.; Rizzi, A.; Marini, D., *Local Color Correction of Stereo Pairs*, San Jose, California, IS&T/SPIE Electronic Imaging 2010, 2010.

Gao, X.; Hong, K.; Passmore, P.; Podladchikova, L.; Shaposhnikov, D., *Colour Vision Model-Based Approach for Segmentation of Traffic Signs*, EURASIP Journal on Image and Video Processing, 2008.

Gao, X.; Anishenko, S.; Shaposhnikov, D.; Podlachikova, L.; Batty, S.; Clark, J., *High-precision Detection of Facial Landmarks to Estimate Head Motions Based on Vision Models*, Journal of Computer Science 3 (7), 2007.

Gatta, C., Rizzi, A. and Marini, D. , *ACE: an Automatic Color Equalization algorithm*, CGIV02 the First European Conference on Color in Graphics Image and Vision, Poitiers (France), 2002.

Gatta, C., Rizzi, A. and Marini, D. *Local Linear LUT Method for Spatial Color Correction Algorithm Speed-up*, IEE ProceedingsIEE Proc. Vision, Image & Signal Processing, Volume 153, 2006.

Geisler, W. S., *Adaptation, Afterimages and Cone Saturation*, Vision Research 18, 1978.

Goethe, Johann Wolfgang von., *Theory of Colours (1810)*, translated from the German by C. L. Eastlake, John Murray, 1840.

Gregory, R., *Perceptual illusions and brain models*, Proc. R. Soc, London Ser. B, 1968.

Grossberg, M. and Nayar, S., *Modeling the Space of Radiometric Response Functions*, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol.26, No.10, pp.1272-1282, 2004. Vol. 26.

Grossberg, M. and Nayar, S., *What can be Known about the Radiometric Response Function from Images?*, European Conference on Computer Vision (ECCV), Vol.IV, pp.189-205, 2002. p. 189–205.

Helson, H., Judd, D.B. and Warren, M.H., *Object color changes from daylight to incandescent filament illumination*, Illum. Eng. 47, 1952.

Hering, E., *Outlines of a Theory of the Light Sense*, Harvard University Press, 1964.

Hunt, R. W. G., *A model of colour vision for predicting colour appearance*, Color Res. Appl. 7, 1982.

Hunt, R. W. G. and Luo, M. R., *Evaluation of a model of colour vision by magnitude scalings: Discussion of collected results*, Color Res. Appl. 19, 1994.

Hunt, R. W. G. and Winter, L. M., *Colour adaptation in picture-viewing situations*, J.Phot. Sci. 23, 1975.

Hunt, R. W. G., *Light and dark adaptation and the perception of color*, J. Opt. Soc. Am.42, 1952.

Hunt, R. W. G., *The reproduction of colour*, 5th ed., Fountain Press, 1995.

Hunt, R. W. G., *The specification of colour appearance. I. Concepts and terms*, Color Res. Appl. 2, 1977.

Hurvich, L. M. and Jameson, D., *An Opponent-Process Theory of Color Vision*, Psychological Review 64, 1957.

Kuang, J. and Fairchild, M. D., *iCAM06, HDR and Image Appearance*, IS&T/SID 15th Color Imaging Conference, 2007.

Fairchild, M. D. and Johnson, G. M., *Image appearance modeling*, Proc. SPIE/IS&T Electronic Imaging Conference, 2003.

Kitaoka, A., *A brief Classification of Colour Illusion*, Colour: Design & Creativity (5), 2010.

Kuang, J., Johnson, G. M. and Fairchild, M. D.,*iCAM06: A refined image appearance model for HDR image rendering*, Journal of Visual Communication, 2007.

Lam, K.M., *Metamerism and colour constancy [DPhil thesis]*,University of Bradford, 1985.

Land, E. H., *Color vision and the natural image: part I*, Proceedings of the National Academy of Sciences, Vol. 45, No. 1, pages 115-129, 1959.

Land, E. H., *The Retinex Theory of color vision*, Scientific American, 237/6, published by W. H. Freeman and Company, 1977. p. 108-128. Vol. Scientific American 237.

Land, E. H. and McCann, J. J., *Lightness and Retinex theory*, Journal of Optical Society of America 61, 1971.

Li, C.; Luo, M.R.; Rigg, B.; Hunt, R.W.G, CMC 2000 Chromatic Adaptation Transform: CMCCAT2000, Color Res Appl 27, 2002.

Luo, M.R. and Hunt, R.W.G., A chromatic adaptation transform and a colour inconstancy index, Color Res. Appl. 23, 1998.

Luo, M.R.; Clarke, A.A.; Rhodes, P.A.; Schappo, A.; Scrivner, S.A.R.; Tait, C.J., Quantifying colour appearance. Part I. LUTCHI colour appearance data, Color Res. Appl. 16, 1991a.

Luo, M.R.; Clarke, A.A.; Rhodes, P.A.; Schappo, A.; Scrivner, S.A.R.; Tait, C.J., Quantifying colour appearance. Part II. Testing colour models performance using LUTCHI color appearance data, Color Res. Appl. 16, 1991b.

Luo, M.R., Lo, M.C. and Kuo, W.G.,*The LLAB colour model*, Color Res. Appl. 21, 1996.

MacAdam, D.L., *A nonlinear hypothesis for chromatic adaptation*, Vis. Res. 1, 1961.

Mann, S. and Picard, R., *Being 'Undigital' with Digital Cameras: Extending Dynamic Range by Combining Differently Exposed Pictures*, Proc. of IST's 48th Annual Conference,, 1995. p. 422–428.

Marks, W. B., Dobelle, W. H. and MacNichol, E. F., *Visual pigments of single primate cones*, Science 143, 1964.

McCann, J. J. and Rizzi, A., *Camera and visual veiling glare in HDR images*, Journal Soc Info Display, 2007.

McCann, J. J., *Mechanism of Color Constancy*, Proc. IS&T/SID Color Imaging Conference, 2004.

McCann, J. J. and Houston, K. L., *Color Sensation, Color Perception and Mathematical Models of Color Vision*, Colour Vision, J. D. Mollon, and L. T. Sharpe, ed., Academic Press, 1983.

McCann, J. J., Parraman, C. E. and Rizzi, A., *Reflectance, Illumination, and edges in 3-D Mondrian Colour Constancy Experiments*, Sidney : Proceedings of the 2009 Association Internationale de la Couler 11th Congress, 2009.

Mitsunaga, T. and Nayar, S. K., *Radiometric Self Calibration, Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, 1999. p. 374–380. Vol. 2.

Mori, L.; Sobagaki, H.; Komatsubara, H.; Ikeda, K., *Field trials on CIE chromatic adaptation formula*, Proceedings of the CIE 22nd Session, Melbourne, 1991.

Moroney, N. and Tastl, I., A Comparison of Retinex and iCAM for Scene Rendering, Journal of Electronic Imaging, 2004.

Moroney, N.; Fairchild, M. D.; Hunt, R.W.G.; Li, C. J; Luo, M. R.; Newman, T., *The CIECAM02 color appearance model*, Scottsdale, IS&T/SID 10th Color Imaging Conference, 2002.

NASA, Langley research centre, [Online 21-12-2010]. http://dragon.larc.nasa.gov/retinex/background/retpubs.html.

Nayar, S. e Mitsunaga, T., *High Dynamic Range Imaging: Spatially Varying Pixel Exposures*, IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Vol.1, pp.472-479, 2000. p. 472-479.

Nayatani, Y.; Hashimoto, K.; Takahama, K.; Sobagaki, H., *A nonlinear color-appearance model using Estévez–Hunt–Pointer primaries*, Color Res. Appl. 12, 1987.

Neugebauer, H. E., *Die theoretischen Grundlagen des Mehrfarbendrucks* , Zeitschrift fur wissenschaftliche Photographie, Photophysik und Photochemie [translated by D. Wyble and A. Kraushaar, The theoretical Basis of Multicolor Letterpress Printing, 2005], 1937.

Newton, Isaac,*Opticks: Or, A Treatise of the Reflexions, Refractions, Inflexions and Colours of Light*, Smith and Waltford, 1704.

Oleari, C. (a cura di), *Misurare il colore. Spettrofotometria, fotometria e colorimetria. Fisiologia e percezione*, Hoepli, 1998.

Oleari, C. (a cura di), *Misurare il colore: Fisiologia della visione a colori - Fotometria - Colorimetria e norme internazionali*, Milano,,Claudio Oleari Ed., Editrice Hoepli, 2008.

Optical Society of America, Committee on Colorimetry, *The Science of Colour*, Crowell, 1953.

Pearlman, A. L., Birch, J. and Meadows, J.C., *Cerebral color blindness: an acquired defect in hue discrimination*, Ann. Neurol. 5, 1979.

Provenzi, E.; Gatta, C.; Fierro, M.; Rizzi, A., *A Spatially Variant White Patch and Gray World Method for Color Image Enhancement Driven by Local Contrast*, IEEE Transactions on Pattern Analysis and Machine Intelligence Vol. 30, No. 10, 2008.

Provenzi, E.; Fierro, M.; Rizzi, A.; De Carli, L.; Gadia, D.; Marini, D., *Random Spray Retinex: a new Retinex implementation to investigate the local properties of the model*, IEEE Transactions on Image Processing, Vol. 16, Issue 1, pp. 162-171, 2007.

Purkyně, J. E., *Neue Beiträge zur Kenntniss des Sehens in Subjectiver Hinsicht*, Berlin, Reimer, 1825.

Rahman, Z.; Jobson, D. J.; Woodell, G. A.; Hines, G. D., *Impact of multiscale retinex computation on performance of segmentation algorithms*, Visual Information Processing XIII, Proc. SPIE 5438, 2004.

Reinhard, E.; Ward, G.; Pattanaik, S.; Debevec, P., *High Dynamic Range Imaging*, Morgan Kaufmann, 2005.

Rizzi, A. and McCann, J.J., *On the behavior of spatial models of color*. S. Jose, California : IS&T/SPIE Electronic Imaging 2007, 2007.

Rizzi, A.; Gatta, C.; Taraschi, F.; Zanardini, M.; Abbiati, M., *Un algoritmo per la valutazione percettiva delle interfacce visuali*, DDD_INTERFACCE, Rivista trimestrale Disegno e Design Digitale, 2003.

Rizzi, A.; Gatta, C.; Slanzi, C.; Ciocca, G.; Schettini, R., *Unsupervised Color Film Restoration Using Adaptive Color Equalization*, Lecture Notes in Computer Science, Volume 3736, 2006.

Rizzi, A., Gatta, C. e Marini, D., *A new algorithm for unsupervised global and local color correction*, Pattern Recognition letters, Vol. 24 (11), pp. 1663-1677, 2003.

Robinson, J. O., *The Psychology of Visual Illusion*, Hutchinson, 1972.

Rolleston, R. e Balasubramanian, R., *Accuracy of Various Types of Neugebauer*, IS&T and SID's Color Imaging Conference: Transforms & Transportability of Color, 1993.

Saito, K., *Electronic image pickup device*, 08-340486 Japan, December 1996.

Schanda, J., *Colorimetry: Understanding the CIE System*, Janos Schanda, 2007.

Semani, D., Chambah, M. and Courtellmont, P., *Processing of underwater color images applied to live aquarium videos*, Journal of Robotics and Automation, 20(2), 2005.

Semmelroth, C. C., *Prediction of lightness and brightness on different backgrounds*, J. Opt. Soc. Am. 60, 1970.

Somers, D. C. and Adelson, E. H., *Junctions, transparency, and brightness*, Invest. Ophthalmol. Vis. Sci. 38, 1997.

Stevens, J.C. e Stevens, S.S., *Brightness functions: Effects of adaptation*, J. Opt. Soc. Am. 53, 1963.

von Helmholtz, Hermann, *Treatise on physiological optics* (translation of of 3rd German ed., 1909), Dover ed., 1962.

von Helmholtz, Hermann, *Wissenschaftliche Abhandlungen von Hermann von Helmholtz*, J.A. Barth, 1892.

von Kries, J., *Chromatic adaptation* [Translation: D.L. MacAdam, Sources of Color Science, MIT Press 1970], Festschrift der Albrecht-Ludwig-Universität, 1902.

Woodell, G. A.; Jobson, D. J.; Rahman, Z.; Hines, G. D., *Advanced Image Processing of Aerial Imagery*, Visual Information Processing XV, Proc. SPIE 6246, 2006.

Wright, W.D., *Why and how chromatic adaptation has been studied*, Color Res. Appl. 6, 1981.

Wyszecki, G., *Current developments in colorimetry*, AIC Color, 1973.

Young, Thomas, *On the Theory of Light and Colours*, Philosophical Transactions Royal Society, 1802.

Yule, J.A.C. and Nielsen, W.J., *The penetration of light into paper and its effect on halftone reproduction*, TAGA (Technical Association of the Graphic Arts) Proceeding, p. 65-76, 1951.

Zeki, S., *A vision of the brain*, Wiley-Blackwell, 1993.

Zeki, S., et al., *A direct demonstration of functional specialization in human visual cortex*, Journal of Neuroscience, Vol. 11, 1991.

Zeki, S., *Functional Specialization in the Visual Cortex of the Rhesus Monkey*, Nature 274, 1978.

Zuffi, S., Santini, S. and Schettini, R., *Accounting for Inks Interaction in the Yule Nielsen Spectral Neugebauer Model*, Journal of Imaging Science and Technology, 2006. Vol. Vol. 50.